
POLITECNICO DI TORINO

Collegio di Ingegneria Gestionale
Corso di Laurea Magistrale in Ingegneria Gestionale

Tesi di Laurea di II livello

# Analisi bibliografica sull'evoluzione delle tecniche di Manutenzione Predittiva: dalla statistica all'AI

Relatori:
Prof. Maurizio Galetto
Prof.ssa Elisa Verna

Candidato:
Edoardo Migliasso

Anno accademico 2025-2026

# Indice

# Capitolo 1: Introduzione

Nell'attuale scenario economico globale, caratterizzato da una competizione sempre più intensa e da mercati che richiedono prodotti di alta qualità a costi contenuti, l'efficienza e l'affidabilità dei processi produttivi non rappresentano più un semplice vantaggio competitivo, ma un requisito fondamentale per la sopravvivenza e il successo delle aziende manifatturiere. In questo contesto l'attività di manutenzione assume un ruolo centrale. Un approccio non ottimale alla manutenzione può infatti generare inefficienze che si ripercuotono negativamente sull'intera catena, compromettendo la redditività e la reputazione aziendale.

## 1.1. Contesto: L'Importanza della Manutenzione nei Sistemi Produttivi

In passato, la manutenzione era spesso vista come un ruolo di supporto, un inevitabile centro di costo per l'azienda. Oggi, nel contesto dell'Industria 4.0, questo approccio è stato superato: la manutenzione è sempre più riconosciuta come un processo strategico, un fattore chiave per la competitività di un'impresa [2]. Una strategia di manutenzione efficace, infatti, non si limita più a riparare i guasti, ma contribuisce direttamente a raggiungere gli obiettivi aziendali in termini di produttività, qualità e sicurezza. Ignorare la sua importanza significa esporre l'azienda a seri rischi operativi e finanziari, che si manifestano principalmente attraverso fermi macchina non pianificati e una riduzione della qualità del prodotto.

L'impatto economico di un guasto improvviso va ben oltre il costo immediato della riparazione. Il costo totale di un fermo macchina è infatti una somma di costi diretti, indiretti e intangibili. I costi diretti, come le spese per i ricambi e la manodopera, sono i più facili da calcolare, ma spesso rappresentano solo la punta dell'iceberg [70].

I costi indiretti sono solitamente molto più onerosi. Il principale è la perdita di produzione, che si traduce in una mancata fatturato per ogni minuto di fermo. Questo problema è ancora più grave nei sistemi di produzione *Just-In-Time* (JIT), dove l'assenza di scorte intermedie può causare un "effetto domino", paralizzando l'intera linea produttiva a causa di un singolo guasto [100]. A questo si aggiungono i ritardi nelle consegne, che possono comportare penali e la necessità di costose spedizioni urgenti. Inoltre, un macchinario che sta degradando —

come un utensile usurato o un cuscinetto che vibra in modo anomalo — produrrà pezzi difettosi. Questo genera i cosiddetti "costi della non-qualità", un concetto introdotto da Feigenbaum [29] che include sia i costi interni (scarti, rilavorazioni) sia quelli esterni (garanzie, reclami, richiami di prodotto).

Infine, i costi intangibili, sebbene difficili da quantificare, possono essere devastanti nel lungo periodo. Ritardi e prodotti difettosi danneggiano la reputazione aziendale, spingendo i clienti verso la concorrenza. Un ambiente di lavoro con guasti frequenti può anche abbassare il morale dei dipendenti e aumentare i rischi per la sicurezza, con costi legati a infortuni e a un maggior turnover del personale [73].

Per affrontare questi problemi, le pratiche di manutenzione si sono evolute nel tempo. Questo percorso può essere schematizzato in diverse fasi, ciascuna con i propri vantaggi e limiti.

1. Manutenzione Correttiva (*Run-to-Failure* - RTF): La forma più basilare di manutenzione, anche nota come "manutenzione a guasto", che consiste nell'intervenire solo dopo che il guasto si è verificato. Questa strategia massimizza l'utilizzo teorico di un componente ma porta inevitabilmente a fermi macchina non pianificati, richiedendo riparazioni in emergenza che sono logisticamente complesse e spesso più costose a causa dell'urgenza. Inoltre, i guasti catastrofici possono causare danni secondari ad altri componenti del macchinario e porre seri rischi per la sicurezza degli operatori [84].

2. Manutenzione Preventiva (*Time-Based Maintenance* - TBM): Rappresenta il primo e fondamentale passo verso un approccio proattivo. Invece di attendere il guasto, gli interventi di manutenzione (come sostituzioni, lubrificazioni o ispezioni) vengono eseguiti a intervalli di tempo predeterminati (es. ogni 6 mesi) o dopo un certo numero di ore di funzionamento o cicli produttivi. Questi intervalli sono tipicamente basati su dati storici di affidabilità del componente, come il *Mean Time Between Failures* (MTBF), o sulle raccomandazioni del costruttore. La TBM permette di pianificare i fermi macchina, trasformando le interruzioni da eventi imprevedibili a operazioni programmate, da eseguirsi preferibilmente durante i periodi di bassa produzione. Questo riduce i costi di emergenza e aumenta la disponibilità dell'impianto [110]. Tuttavia, la sua efficacia è limitata e non tiene conto delle reali condizioni operative del macchinario, che possono variare significativamente. Ciò porta a due tipi di inefficienza:

- Sovra-manutenzione (*Over-maintenance*): Se un componente viene utilizzato in condizioni meno severe del previsto, esso viene sostituito quando ha ancora una considerevole vita utile residua (*Remaining Useful Life* - RUL), generando uno spreco di risorse.

- Sotto-manutenzione (*Under-maintenance*): Se le condizioni operative sono più gravi del previsto, il guasto può verificarsi prima dell'intervento programmato, riportando l'azienda alla problematica della manutenzione correttiva.

La consapevolezza di questi limiti ha guidato la ricerca verso paradigmi manutentivi più sofisticati e "intelligenti". La transizione successiva, fondamentale per questa tesi, è il passaggio dalla manutenzione basata sul tempo a quella basata sulla condizione effettiva dell'asset. Questo apre le porte alla Manutenzione su Condizione (*Condition-Based Maintenance* - CBM) e, come sua evoluzione predittiva, alla Manutenzione Predittiva (*Predictive Maintenance* - PdM), che sfruttano i dati raccolti in tempo reale dai sensori per ottimizzare le decisioni di intervento, costituendo il fulcro dell'analisi che verrà sviluppata nei capitoli successivi [48][61].

## 1.2. Il Limite degli Approcci Reattivi: il Monitoraggio con le Carte di Controllo

Un approccio storicamente fondamentale per il monitoraggio dei processi è il Controllo Statistico di Processo (SPC), una metodologia sviluppata da Walter A. Shewhart e poi diffusa da W. Edwards Deming, che ha spostato il focus dall'ispezione del prodotto finito alla stabilità del processo produttivo [25] [101]. L'SPC fornisce uno strumento per "ascoltare" il processo e capire se il suo comportamento è normale o se sta subendo delle anomalie.

Il suo strumento principale è la carta di controllo. Questo strumento grafico non è semplicemente un diagramma temporale, ma un'applicazione rigorosa della teoria statistica

per il monitoraggio in tempo reale. L'idea di base è distinguere tra due tipi di variazione che affliggono qualsiasi processo:

1. Variazione da Cause Comuni (o Casuali): È il risultato di un numero elevato di piccole cause impossibili da isolare singolarmente (es. lievi fluttuazioni di temperatura, minime differenze nella materia prima, impercettibili usure). Un processo soggetto unicamente a cause comuni è detto "in stato di controllo statistico": il suo comportamento, pur non essendo costante, è stabile e prevedibile entro un certo range.

2. Variazione da Cause Speciali (o Assegnabili): Questa è una variabilità anomala, non casuale, che deriva da eventi specifici, identificabili e potenzialmente correggibili. Esempi classici in un contesto di lavorazione meccanica includono un cambio di lotto di materiale con caratteristiche diverse, un'errata impostazione dei parametri macchina da parte di un operatore, la rottura improvvisa di un utensile o il degrado accelerato di un componente.

La carta di controllo (*Figura 1*) traduce questa teoria in un'applicazione pratica. Su un grafico vengono riportati i valori di una caratteristica di qualità (es. il diametro di un pezzo tornito) o di un parametro di processo (es. la temperatura di un forno) campionati a intervalli regolari. Su questo grafico vengono tracciate tre linee guida: la Linea Centrale (CL), che rappresenta la media storica del processo quando è in controllo, e i Limiti di Controllo Superiore (UCL) e Inferiore (LCL). È importante notare che questi limiti non sono le tolleranze di progetto, ma sono calcolati statisticamente dai dati stessi, rappresentando la "voce del processo". Tipicamente posizionati a una distanza di ±3 deviazioni standard (σ) dalla media, un intervallo che contiene il 99.73% dei dati di un processo stabile. Ciò implica che la probabilità che un punto cada al di fuori di questi limiti per puro caso è solo dello 0.27%, un evento sufficientemente raro da essere considerato un segnale di allarme attendibile [72].
Il funzionamento della carta di controllo è quindi diagnostico. Un processo è considerato stabile finché i punti campionati fluttuano casualmente attorno alla linea centrale e all'interno dei limiti di controllo. Il sistema entra in allarme quando si verifica un'anomalia statistica:

- Violazione dei limiti di controllo: Un singolo punto cade al di fuori dell'UCL o del LCL. Questo è il segnale più forte che una causa speciale ha agito sul processo.

- Pattern non casuali (*Run Rules*): Anche se tutti i punti sono all'interno dei limiti, la loro disposizione può rivelare un problema. Le "Western Electric Rules" codificano questi pattern, come una serie di sette punti consecutivi tutti al di sopra o al di sotto della linea centrale, o sei punti in trend crescente o decrescente. Questi pattern indicano che il processo sta subendo una deriva sistematica, anche se non ancora così grave da violare i limiti di controllo.

L'SPC è uno strumento di diagnosi post-evento, non di prognosi pre-evento. Il suo allarme, sia esso una violazione dei limiti o un pattern non casuale, indica che una condizione indesiderata si è già manifestata o sta iniziando a manifestarsi. L'azione correttiva che ne consegue è una reazione a un problema che è già in atto.
In conclusione, il Controllo Statistico di Processo, pur rimanendo uno strumento indispensabile per la gestione della qualità, rappresenta il culmine di un paradigma di monitoraggio reattivo. Esso delimita il confine tra stabilità e instabilità, ma la sua azione si innesca solo al superamento di tale confine. La sfida contemporanea, spinta dalle potenzialità dell'Industria 4.0, consiste nel superare questo limite concettuale: non limitarsi a monitorare lo stato attuale, ma prevedere l'evoluzione futura dello stato di salute di un sistema, intervenendo in modo proattivo quando il processo è ancora perfettamente funzionante e "in controllo".

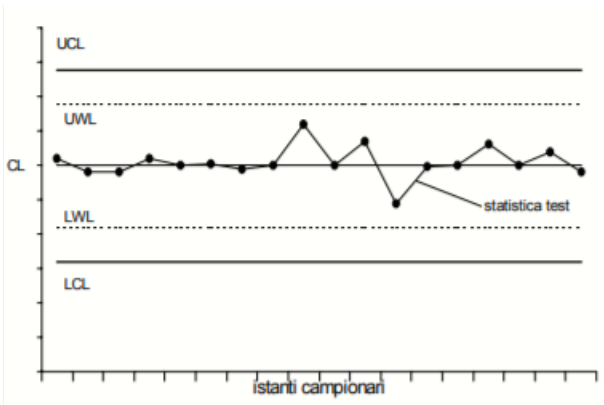


*Figura 1 – Esempio di Carta di Controllo*

## 1.3. Verso un Approccio Proattivo: Prognostica e Manutenzione Predittiva (PdM)

I limiti evidenti della manutenzione preventiva a calendario (TBM) e del monitoraggio reattivo (SPC) hanno spinto la ricerca a sviluppare un approccio superiore. L'obiettivo è superare la scelta tra l'agire troppo presto, sprecando risorse, e l'agire troppo tardi, subendo guasti catastrofici. Si cerca quindi un intervento ottimale, una sorta di "just-in-time" della manutenzione, che sia basato sulla condizione reale dell'asset ma orientato al futuro. Questo salto concettuale si basa su due discipline collegate: la prognostica e la sua applicazione pratica, la manutenzione predittiva (PdM).

La prognostica (*prognostics*) è la scienza che si occupa della previsione del degrado e della stima della vita operativa rimanente di un sistema o di un componente. Come definito da Pecht [80], la prognostica è la capacità di prevedere il tempo rimanente prima che un sistema o un componente non sia più in grado di svolgere la sua funzione prevista entro i limiti di performance desiderati. Il suo risultato principale è la stima della Vita Utile Residua (*Remaining Useful Life* - RUL), spesso espressa non come un valore, una distribuzione di probabilità per tener conto dell'incertezza. Per compiere questa stima, la prognostica si basa sull'acquisizione continua di dati da sensori installati sui sensori, che misurano parametri fisici sensibili al processo di degrado: segnali di vibrazione per macchinari rotanti, analisi termografiche, emissioni acustiche, analisi dell'olio lubrificante, assorbimenti di corrente elettrica, e molti altri. Analizzando queste serie storiche di dati, attraverso modelli avanzati, è possibile identificare i primi segnali di un guasto (*fault precursors*) e tracciare la traiettoria del degrado, prevedendo quando supererà una soglia critica [48].

È importante distinguere la Manutenzione Predittiva (PdM) dalla sua antenata, la Manutenzione su Condizione (*Condition-Based Maintenance* - CBM). La CBM ha rappresentato il primo passo verso una manutenzione "intelligente", basando le decisioni di intervento non sul tempo, ma sul superamento di una soglia di allarme predefinita su un parametro monitorato. Ad esempio, un intervento CBM viene attivato quando l'ampiezza delle vibrazioni di un cuscinetto supera un valore limite. Sebbene più efficiente della TBM, la CBM è ancora fondamentalmente una strategia reattiva alla soglia: l'allarme scatta quando il degrado ha già raggiunto un livello considerato critico, lasciando una finestra temporale per l'intervento potenzialmente molto breve.

La Manutenzione Predittiva (PdM), invece, rappresenta l'evoluzione finale di questo approccio, incorporando la prognostica nel processo decisionale. La PdM non si limita a reagire a una soglia, ma agisce sulla base di una previsione. Utilizzando i modelli prognostici, essa stima la RUL e pianifica l'intervento manutentivo in modo che avvenga in un punto ottimale prima del guasto previsto. In altre parole, invece di intervenire *quando* la vibrazione supera il limite, la PdM interviene *quando prevede che* la vibrazione supererà il limite. Questa capacità previsionale trasforma radicalmente la pianificazione della manutenzione, offrendo vantaggi strategici significativi [87]:

1. Ottimizzazione della Vita Utile e dei Costi: La PdM consente di estendere l'utilizzo dei componenti fino quasi al termine della loro vita utile effettiva, eliminando gli sprechi della TBM e riducendo drasticamente i costi dei ricambi e degli interventi non necessari.

2. Aumento della Disponibilità e Affidabilità (Uptime): Sostituendo i fermi macchina imprevisti con interventi pianificati con largo anticipo, la PdM massimizza la disponibilità operativa degli impianti. La pianificazione consente di ordinare i ricambi per tempo, schedulare il personale e programmare il fermo durante periodi di bassa domanda, minimizzando l'impatto sulla produzione [18].

3. Miglioramento della Sicurezza: Anticipare i guasti, specialmente quelli catastrofici, riduce significativamente i rischi per la sicurezza degli operatori e per l'ambiente.

L'implementazione di una tale strategia richiede un'architettura integrata nota come sistema di Prognostics and Health Management (PHM). Un sistema PHM è un framework ingegneristico che orchestra l'intero processo, dalla raccolta dati all'azione, tipicamente strutturato in sette livelli sequenziali [62]:

1. Acquisizione Dati (Data Acquisition): Installazione di sensori per monitorare i parametri rilevanti.
2. Pre-elaborazione Dati (Data Pre-processing): Pulizia dei dati da rumore e artefatti.
3. Estrazione delle Feature (Feature Extraction): Calcolo di indicatori sintetici (features) dallo stato di salute (es. RMS delle vibrazioni).
4. Diagnosi (Diagnostics): Identificazione e isolamento di un guasto quando si verifica.

5. Prognosi (Prognostics): Stima della RUL e previsione del comportamento futuro del degrado.
6. Supporto alle Decisioni (Decision Support): Raccomandazione di azioni ottimali basate sulla prognosi.
7. Presentazione (Presentation Layer): Interfaccia utente per visualizzare lo stato di salute e le raccomandazioni.

## 1.4. Obiettivi della Tesi e Domanda di Ricerca

Come abbiamo visto, il percorso evolutivo della manutenzione ci ha portati alla Manutenzione Predittiva (PdM) come paradigma di frontiera. Tuttavia, proprio per la sua importanza, il numero di studi scientifici su questo argomento è cresciuto in modo esponenziale. Oggi, il campo della PdM è diventato così vasto e frammentato che è difficile, sia per i ricercatori che per i professionisti del settore, avere una visione d'insieme chiara e operativa.

L'obiettivo principale di questa tesi è proprio quello di fare ordine in questo panorama. Attraverso una revisione sistematica e critica della letteratura, questo lavoro si propone di mappare il campo della PdM, identificando le tecniche principali, classificandole in modo logico e tracciandone l'evoluzione nel tempo. L'intento non è una semplice raccolta di articoli, ma una sintesi che aiuti a comprendere lo stato dell'arte.

L'indagine si concentrerà esclusivamente sulle applicazioni della PdM nel settore manifatturiero, con un'attenzione specifica ai processi di lavorazione meccanica per asportazione di truciolo (es. tornitura, fresatura, foratura). Tale scelta è motivata da tre fattori strategici sia scientifici che pratici.

In primo luogo, la criticità economica di questi processi è vasta, costituiscono il cuore della produzione di componenti ad alta precisione in settori industriali fondamentali. Si pensi, ad esempio, alla tornitura di un albero motore nel settore automotive, alla fresatura di una pala di turbina in superlega per il settore aerospaziale, o alla micro-foratura di un impianto di protesi in titanio nel medicale. In contesti del genere, un guasto imprevisto o un degrado non monitorato dell'utensile non causa solo un fermo macchina, ma può portare allo scarto di pezzi dal valore di migliaia di euro, a ritardi critici e, nei casi più gravi, a compromettere la sicurezza e l'affidabilità del prodotto finale.

In secondo luogo, questo settore offre una straordinaria ricchezza di dati. I moderni centri di lavoro a controllo numerico (CNC) sono sistemi equipaggiati con molteplici sensori che monitorano in tempo reale ogni aspetto del processo: accelerometri misurano le vibrazioni del mandrino, celle di carico registrano le forze di taglio, termocoppie monitorano le temperature, e i drive dei motori forniscono dati su coppia e assorbimento di corrente. Questo flusso di dati multi-sensoriale, ricco e ad alta frequenza, offre un ottimo spunto e *data-rich* per lo sviluppo e la validazione di modelli predittivi complessi, permettendo di testare le metodologie più avanzate in un contesto ricco di informazioni.
Infine, la complessità del degrado rende questi processi un eccellente banco di prova scientifico. I fenomeni di usura progressiva degli utensili e di degrado dei componenti macchina (come mandrini, cuscinetti e viti) sono intrinsecamente complessi e non lineari. L'usura di un utensile, ad esempio, non è un processo lineare ma è influenzata da complesse interazioni termo-meccaniche tra l'utensile, il pezzo e il truciolo. Questa complessità "naturale" costringe i modelli predittivi ad andare oltre le semplici previsioni, rendendo questo dominio il terreno ideale per confrontare e validare le metodologie prognostiche più avanzate, dal Machine Learning classico al Deep Learning.
Per guidare questa indagine in modo rigoroso, il lavoro sarà strutturato attorno a una domanda di ricerca (RQ) principale, articolata in tre sotto-domande (SQ), che definiscono le tappe del percorso analitico:

- RQ (Domanda Principale): Qual è lo stato dell'arte e quali sono le traiettorie evolutive delle metodologie di Manutenzione Predittiva per i processi di lavorazione meccanica, delineando la transizione dagli approcci statistici fondativi alle tecniche innovative basate su Intelligenza Artificiale e architetture Digital Twin?

Per rispondere in modo esaustivo a questa domanda, verranno analizzate le seguenti questioni secondarie:

1. SQ1: Quali sono le metodologie fondamentali e gli approcci statistici che hanno storicamente costituito le fondamenta della Manutenzione Predittiva?
   Questa sotto-domanda mira a costruire il basamento teorico della tesi, analizzando criticamente la letteratura relativa a modelli di affidabilità (es. Weibull), analisi di serie storiche e processi stocastici, al fine di elucidarne i principi operativi e,

soprattutto, i limiti intrinseci che hanno reso necessaria l'adozione di approcci più sofisticati.

2. SQ2: In che modo le diverse famiglie di algoritmi di Machine Learning e il paradigma del Digital Twin hanno rivoluzionato l'approccio alla PdM? Quali sono le architetture, le tecniche e le sfide applicative specifiche che caratterizzano queste innovazioni?
   Questo costituisce il nucleo investigativo della tesi. Si procederà a un'indagine dettagliata degli studi che impiegano apprendimento supervisionato per la stima della RUL, apprendimento non supervisionato per l'identificazione di anomalie, e reti neurali profonde per l'analisi di dati complessi (es. segnali di vibrazione grezzi).

3. SQ3: Quali sono i principali trend applicativi, le sfide ricorrenti e le future direzioni di ricerca nel campo della PdM per le lavorazioni meccaniche?
   Questa domanda finale ha lo scopo di fornire una visione critica e prospettica. Attraverso l'analisi tassonomica dei lavori selezionati, si intende sintetizzare quali problemi sono stati affrontati con maggior successo, quali sfide rimangono aperte (es. interpretabilità dei modelli, generalizzazione a diverse condizioni operative, gestione di piccoli set di dati di guasto) e quali filoni di ricerca appaiono più promettenti.

## 1.5. Struttura del Lavoro

Per rispondere a queste domande di ricerca in modo sistematico e chiaro, la tesi è strutturata in un percorso logico suddiviso in sei capitoli.
Il Capitolo 1 ha avuto la funzione di gettare le fondamenta dell'intera ricerca. Ha delineato il contesto industriale, definito la problematica strategica della manutenzione, esaminato criticamente i limiti degli approcci tradizionali e, infine, ha stabilito con precisione la motivazione, gli obiettivi, il perimetro e le domande guida che costituiscono la spina dorsale di questo studio.
Il Capitolo 2 è interamente dedicato alla spiegazione del quadro metodologico, un passaggio indispensabile per assicurare la trasparenza, il rigore scientifico e la replicabilità

dell'indagine. Verrà illustrato in dettaglio il protocollo PRISMA (*Preferred Reporting Items for Systematic Reviews and Meta-Analyses*), lo standard di riferimento nella letteratura scientifica per la conduzione di revisioni sistematiche [71]. Questo capitolo renderà esplicita ogni fase del processo: la formulazione delle stringhe di ricerca, la selezione dei database scientifici (Scopus, Web of Science, Google Scholar), e la definizione dei criteri di inclusione ed esclusione utilizzati per filtrare il vasto corpus di letteratura e giungere a un set di studi finale. La descrizione di questo framework costituisce il fondamento su cui poggia la validità di tutte le analisi e conclusioni successive.

A partire dal Capitolo 3, la tesi risponde in modo mirato alla prima sotto-domanda di ricerca (SQ1). Questo capitolo è dedicato allo stato dell'arte delle metodologie tradizionali della Manutenzione Predittiva. Si procederà a un'analisi sistematica degli approcci basati sull'ingegneria dell'affidabilità (es. analisi di Weibull per la modellazione dei guasti), sui modelli statistici per le serie temporali (es. modelli auto-regressivi integrati a media mobile, ARIMA) e sui processi stocastici (es. modelli di Markov e Hidden Markov Models). L'obiettivo sarà duplice: descrivere i principi operativi e, soprattutto, valutare le assunzioni e i limiti, in particolare per quanto riguarda la loro applicabilità a sistemi meccanici complessi, caratterizzati da non-linearità e interazioni multiple.

Il Capitolo 4 affronta la seconda sotto-domanda di ricerca (SQ2), esplorando il paradigma innovativo che oggi definisce la frontiera della PdM. Questo capitolo analizzerà l'impatto trasformativo delle tecniche di Machine Learning, esaminando le tre principali famiglie di apprendimento: supervisionato (per la stima della RUL), non supervisionato (per l'identificazione di anomalie e la clusterizzazione degli stati operativi) e il Deep Learning (per l'analisi diretta di dati grezzi e complessi). In parallelo, si esaminerà il paradigma del Digital Twin.

Il Capitolo 5 presenterà la sintesi e l'analisi aggregata dei risultati emersi dalla revisione. Qui verrà sviluppata una tassonomia multi-dimensionale per classificare in modo organico la letteratura selezionata, utilizzando criteri quali l'approccio metodologico, l'algoritmo specifico, il processo meccanico di applicazione e il tipo di dati utilizzati.

Infine, il Capitolo 6 rappresenta il termine del percorso di ricerca. Questo capitolo conclusivo offrirà una sintesi dei risultati, fornendo una risposta alla domanda di ricerca principale (RQ). Si procederà con una discussione critica degli sviluppi teorici e pratici dello studio, una valutazione dei suoi limiti e, in modo particolare, si indentificherà una visione sulle direzioni future, identificando le lacune (*research gaps*) nella conoscenza attuale e

proponendo una potenziale agenda per la ricerca futura nel dominio della Manutenzione Predittiva avanzata.

# Capitolo 2: Metodologia della Revisione Sistematica della Letteratura (PRISMA)

Per scrivere una tesi solida è fondamentale spiegare in modo chiaro e trasparente *come* si è arrivati a determinati risultati, per garantirne il rigore e la validità. Questo capitolo ha proprio questo scopo: descrivere nel dettaglio il metodo che ho seguito per condurre questa ricerca.
Invece di fare una semplice rassegna di articoli, ho scelto di utilizzare una metodologia di Revisione Sistematica della Letteratura. Questo approccio non si limita a "raccontare" ciò che è stato scritto su un argomento, ma segue un protocollo rigoroso per identificare, selezionare e analizzare in modo critico tutti gli studi pertinenti a una specifica domanda di ricerca. L'obiettivo è quello di costruire una mappa completa e il più possibile imparziale dello stato dell'arte, basata su un processo che, in teoria, chiunque potrebbe replicare per arrivare a conclusioni simili.
Per strutturare questo processo e seguire uno standard riconosciuto a livello internazionale, ho adottato il protocollo PRISMA (Preferred Reporting Items for Systematic Reviews and Meta-Analyses). Questo framework fornisce una guida e una checklist precise, che mi hanno aiutato a garantire che nessun passaggio fondamentale della revisione venisse tralasciato.

## 2.1. Il Protocollo PRISMA (Preferred Reporting Items for Systematic Reviews and Meta-Analyses)

Per garantire che questa revisione fosse completa e verificabile, ho scelto di non seguire un approccio narrativo tradizionale, che può essere soggettivo, ma di basare il mio lavoro su una metodologia strutturata. A tal fine, questo lavoro si fonda sul protocollo PRISMA, acronimo di Preferred Reporting Items for Systematic Reviews and Meta-Analyses.
Il PRISMA rappresenta l'attuale "gold standard" internazionale per la conduzione e il reporting di revisioni sistematiche. Sebbene le sue origini risiedano nel campo medico-clinico, dove la necessità di sintetizzare in modo rigoroso le evidenze scientifiche è di primaria importanza, i suoi principi di completezza lo hanno reso il framework di riferimento in innumerevoli altre discipline.
Utilizzare il protocollo PRISMA significa trasformare la revisione della letteratura da un'arte interpretativa a una scienza metodica. Si segue un percorso predefinito e documentato, che

costituisce un chiaro "audit trail" (traccia di verifica). Questo percorso permette a qualsiasi altro ricercatore di comprendere, valutare e, potenzialmente, replicare ogni passo, fondando così la validità dello studio su basi solide e difendibili.
Il framework PRISMA si articola su due strumenti operativi complementari, come delineato nella documentazione ufficiale "PRISMA 2020 Statement":

- La Checklist di Reporting: Si tratta di una guida completa, articolata in 27 punti fondamentali, che assicura la completezza informativa del manoscritto finale. Questa checklist funge da strumento di autovalutazione per il ricercatore, garantendo che nessun aspetto metodologico cruciale venga tralasciato.

- Il Diagramma di Flusso a Quattro Fasi: Questo è l'elemento più riconoscibile e operativo del protocollo. Fornisce una rappresentazione visiva e quantitativa dell'intero processo di selezione degli studi, illustrando il "viaggio" di ogni articolo dalla sua identificazione iniziale fino alla sua eventuale inclusione nell'analisi finale. Le quattro fasi sequenziali, che verranno meticolosamente seguite in questa tesi, costituiscono un processo di filtraggio progressivo:

    1. Identificazione (*Identification*): È la fase di avvio, in cui si "getta una rete ampia" per catturare il maggior numero possibile di studi potenzialmente pertinenti. Ciò avviene attraverso ricerche sistematiche in banche dati scientifiche prescelte, utilizzando stringhe di ricerca attentamente progettate per bilanciare la sensibilità (la capacità di trovare tutti gli studi rilevanti) e la specificità (la capacità di minimizzare il "rumore" di fondo, ovvero gli studi irrilevanti).

    2. *Screening*: In questa fase, il vasto corpus grezzo viene raffinato. Il primo passo consiste nella rimozione di tutti i record duplicati. Successivamente, i record unici rimanenti sono sottoposti a un primo, rapido vaglio basato esclusivamente sulla lettura del titolo e dell'abstract. Questo passaggio è un filtro efficiente che permette di escludere un gran numero di studi manifestamente non pertinenti con un dispendio di tempo contenuto.

3. Eleggibilità (*Eligibility*): Questa è la fase più intensiva e critica del processo. Gli articoli che hanno superato lo screening iniziale vengono recuperati per una lettura approfondita e completa del testo integrale (*full-text*). È in questo momento che i criteri di inclusione ed esclusione vengono applicati in modo rigoroso e non ambiguo. Ogni studio viene valutato attentamente per la sua pertinenza tematica, la sua adeguatezza metodologica e il rispetto di tutti gli altri criteri stabiliti. Le ragioni specifiche per l'esclusione di ogni articolo in questa fase vengono meticolosamente registrate, garantendo la massima trasparenza.

4. Inclusione (*Inclusion*): Gli studi che superano con successo la valutazione di eleggibilità costituiscono il corpus finale di evidenze. Questo insieme di articoli rappresenta il risultato del processo di filtraggio sistematico e costituisce il fondamento, la "roccia solida", su cui si baserà l'intera analisi qualitativa, la tassonomia, l'analisi comparativa e la sintesi critica.

L'applicazione di un framework così strutturato in un dominio come la Manutenzione Predittiva fornisce l'ancora metodologica necessaria per navigare in un mare di informazioni. Garantisce che le conclusioni tratte non siano il frutto di una navigazione casuale, ma il risultato di una rotta pianificata, verificabile e scientificamente robusta [83].

## 2.2. Strategia di Ricerca e Fonti Dati

La fase di Identificazione, che inaugura il protocollo PRISMA, è un momento di cruciale importanza strategica per l'intera revisione. La validità dei risultati finali dipende in modo diretto dalla completezza e dal rigore con cui viene condotta la ricerca iniziale. Un approccio sistematico in questa fase è essenziale per costruire un corpus di studi che sia rappresentativo dello stato dell'arte, limitando il rischio di esclusioni involontarie che potrebbero alterare l'analisi. La strategia qui adottata si basa su due pilastri fondamentali e interconnessi: la selezione mirata delle banche dati bibliografiche e la costruzione logica e iterativa delle stringhe di ricerca.

### 2.2.1. Selezione delle Banche Dati Bibliografiche

Consapevole che nessuna singola fonte può garantire una copertura totale della produzione scientifica, si è optato per un approccio multi-database. La scelta è ricaduta su una fonte primaria, Scopus, supportata da fonti complementari per massimizzare l'esaustività.
La decisione di prediligere Scopus come database primario per questa revisione è motivata da una serie di considerazioni strategiche che lo rendono particolarmente adatto a un'indagine in ambito ingegneristico e tecnologico.
Scopus, gestito dall'editore Elsevier, è una delle più vaste banche dati al mondo di letteratura scientifica sottoposta a peer-review. La sua rilevanza per questa tesi non deriva solo dalla sua dimensione, ma dalle sue specifiche caratteristiche funzionali:

- Contenuto Curato e di Qualità: A differenza dei motori di ricerca accademici generici, Scopus indicizza contenuti provenienti da editori che devono superare rigorosi criteri di qualità. Questo fornisce una prima, fondamentale scrematura, garantendo che i risultati provengano da fonti affidabili.

- Copertura Multidisciplinare con Focus Tecnologico: Scopus offre un'eccellente copertura delle discipline ingegneristiche, informatiche e delle scienze dei materiali. Un aspetto di particolare valore è la sua vasta indicizzazione di atti di conferenze (*conference proceedings*), che in settori a rapida evoluzione come il Machine Learning e l'Industria 4.0 rappresentano un canale di diffusione scientifica tanto importante quanto le riviste tradizionali, spesso anticipandone i contenuti.

- Interfaccia di Ricerca Avanzata: La piattaforma permette la costruzione di query di ricerca complesse e precise. Supporta l'uso di operatori booleani (AND, OR, NOT) per combinare o escludere concetti, l'uso di wildcard (caratteri jolly come l'asterisco *) per catturare varianti di una parola, e la ricerca per frasi esatte (tramite virgolette "). Questa flessibilità è essenziale per tradurre una domanda di ricerca complessa in una query efficace, come verrà mostrato nel paragrafo successivo.

- Analisi Citazionale: Ogni articolo in Scopus è collegato a una rete di citazioni, permettendo di vedere sia i lavori che cita (*references*) sia i lavori che lo hanno citato (*cited by*). Questa funzionalità è preziosa per la tecnica dello snowballing, che

consiste nell'esplorare le bibliografie di articoli chiave per trovare ulteriori studi pertinenti.

Per queste ragioni, Scopus è stato identificato come lo strumento più potente ed efficiente per condurre la ricerca principale.

### 2.2.2. Costruzione della Stringa di Ricerca (Query)

La definizione della stringa di ricerca è il cuore operativo della fase di identificazione. Il suo obiettivo è trovare un equilibrio ottimale tra sensibilità (la capacità di includere tutti gli studi pertinenti) e specificità (la capacità di escludere quelli non pertinenti). Per raggiungere tale equilibrio, è stato seguito un approccio strutturato, scomponendo la domanda di ricerca nei suoi tre blocchi concettuali fondamentali.

- Blocco 1 – Dominio (Il "Cosa"): L'argomento centrale della Manutenzione Predittiva e i suoi concetti correlati.

- Blocco 2 – Tecniche (Il "Come"): Le metodologie analitiche, sia tradizionali che innovative.

- Blocco 3 – Contesto Applicativo (Il "Dove"): Il settore industriale e i processi specifici di interesse.

All'interno di ogni blocco, i termini e i loro sinonimi sono stati collegati con l'operatore OR per ampliare la ricerca. I tre blocchi sono stati poi uniti con l'operatore AND per assicurare che ogni articolo risultante fosse pertinente a tutte e tre le dimensioni della ricerca.
La query finale, ottimizzata per la sintassi di Scopus e applicata ai campi TITLE-ABS-KEY (Titolo, Abstract, Keyword), è la seguente:

TITLE-ABS-KEY

( ("predictive maintenance" OR “prognostics” OR "condition-based maintenance" OR "fault diagnosis" OR “phm” OR "remaining useful life" OR “rul”)
AND
("machine learning" OR "artificial intelligence" OR "deep learning" OR "digital twin" OR “statistic” OR “model”)
AND
(“manufacturing” OR “machining” OR “milling” OR “turning” OR “drilling” OR "production system" OR “industry”) )

Questa ricerca è stata eseguita senza restrizioni temporali, per consentire un'analisi dell'evoluzione storica del campo. L'esecuzione di questa query ha generato il corpus di record iniziale, che costituisce l'input per le successive fasi di filtraggio descritte nel resto di questo capitolo.

## 2.3. Criteri di Inclusione ed Esclusione degli Studi

Dopo aver generato, tramite la strategia di ricerca descritta, un corpus iniziale di studi potenzialmente rilevanti, il passo successivo del protocollo PRISMA consiste nel filtrare sistematicamente questo vasto insieme. Lo scopo di questa fase è isolare unicamente quei lavori che sono direttamente pertinenti, metodologicamente solidi e allineati con le domande di ricerca di questa tesi. Questo processo di filtraggio non può essere arbitrario; al contrario, deve basarsi su un insieme di criteri di inclusione ed esclusione definiti con la massima precisione a priori, ovvero prima di iniziare l'esame effettivo degli articoli.
La definizione a priori di queste "regole del gioco" è un caposaldo irrinunciabile di una revisione sistematica. Essa funge da barriera contro il bias di selezione, un rischio cognitivo per cui un ricercatore potrebbe, anche in modo non intenzionale, favorire gli studi che confermano le proprie ipotesi o escludere quelli che le contraddicono. Stabilire criteri chiari e non ambigui garantisce che ogni decisione di inclusione o esclusione sia oggettiva, tracciabile e difendibile, rafforzando così la validità interna dell'intero studio [52].

### 2.3.1. Criteri di Inclusione (CI)

Per essere considerato idoneo e quindi essere incluso nell'analisi qualitativa e quantitativa finale, uno studio doveva soddisfare simultaneamente e senza eccezioni tutti i seguenti criteri:

- CI1 - Pertinenza Tematica Centrale: Lo studio deve avere come oggetto primario di indagine la Manutenzione Predittiva o concetti a essa intrinsecamente legati. Questo include la prognostica (*prognostics*), la diagnosi dei guasti (*fault diagnosis*) come precursore della prognosi, i sistemi integrati di gestione della salute (*Prognostics and Health Management - PHM*), o la stima quantitativa della Vita Utile Residua (*Remaining Useful Life - RUL*).

- CI2 - Focus Metodologico Esplicito: L'articolo deve andare oltre una semplice discussione concettuale. Deve descrivere, proporre, applicare, validare o confrontare in modo dettagliato una o più tecniche metodologiche. Queste possono spaziare dagli approcci statistici tradizionali (es. modelli di regressione, analisi di affidabilità) alle più moderne tecniche di Intelligenza Artificiale, come gli algoritmi di Machine Learning (es. Support Vector Machines, Random Forests) e Deep Learning (es. Reti Neurali Convoluzionali, LSTM), fino alle architetture basate sul paradigma del Digital Twin.

- CI3 - Contesto Applicativo Specifico: Lo studio deve essere chiaramente ambientato nel settore manifatturiero. Verrà data massima priorità e considerazione agli studi la cui applicazione o caso di studio riguardi specificamente processi di lavorazione meccanica per asportazione di truciolo (es. tornitura, fresatura, foratura, rettifica), in quanto cuore tecnologico del perimetro di questa tesi. Saranno considerati anche studi su altri processi produttivi (es. stampaggio, saldatura), purché chiaramente inseriti in un contesto manifatturiero.

- CI4 - Standard di Pubblicazione Accademica: Per garantire un elevato standard di qualità scientifica e di validità dei risultati, saranno inclusi esclusivamente lavori che hanno superato un processo di revisione paritaria (*peer-review*). Questo criterio

limita la selezione a due tipologie principali di pubblicazioni: articoli pubblicati su riviste scientifiche internazionali (*journal articles*) e contributi presentati in atti di conferenze accademiche riconosciute (*conference papers*).

- CI5 - Lingua di Pubblicazione: Per assicurare la piena e accurata comprensione del contenuto scientifico e metodologico, e data la sua posizione di lingua franca nella comunicazione scientifica globale, l'articolo deve essere scritto in lingua inglese.

### 2.3.2. Criteri di Esclusione (CE)

Analiticamente, uno studio è stato escluso dal corpus finale se soddisfaceva anche solo uno dei seguenti criteri, indipendentemente dal fatto che potesse soddisfare alcuni dei criteri di inclusione:

- CE1 - Irrilevanza Tematica o Semantica: Articoli che, pur contenendo le keyword di ricerca, le utilizzano in un contesto semantico radicalmente diverso.

- CE2 - Contesto Applicativo Non Pertinente: Studi la cui applicazione è focalizzata su settori diversi da quello manifatturiero.

- CE3 - Assenza di Contenuto Metodologico Sostanziale: Articoli che si configurano come editoriali, discussioni generali, opinioni, interviste, o survey ad altissimo livello che si limitano a citare le tecniche senza descriverle né applicarle. Sono esclusi anche gli studi che presentano un'architettura concettuale senza alcuna forma di implementazione o validazione, anche simulata.

- CE4 - Tipologia di Pubblicazione Non Ammessa ai Fini dell'Analisi: Per mantenere l'omogeneità e lo standard qualitativo del campione, sono stati esclusi libri interi (sebbene un capitolo specifico, se autonomo e peer-reviewed, possa essere valutato), recensioni di libri, white paper commerciali, standard tecnici, brevetti, tesi di laurea e trattazioni di dottorato.

- CE5 - Testo Completo Non Reperibile: Un criterio di natura pragmatica. Se, dopo un ragionevole sforzo, non è stato possibile accedere al testo integrale dell'articolo attraverso le risorse elettroniche del Politecnico, le richieste all'autore o altri canali di accesso legali, lo studio è stato necessariamente escluso, in quanto una valutazione basata solo sull'abstract sarebbe incompleta e inaffidabile.

- CE6 - Ridondanza dei Contenuti (Studi Duplicati): Articoli che presentano una sovrapposizione sostanziale di metodologia, dati e risultati con un altro studio già incluso. Un caso frequente è quello di un articolo preliminare presentato a una conferenza, successivamente esteso e pubblicato in una versione più completa su una rivista. In tale situazione, per evitare di "inquinare" l'analisi con dati ridondanti, viene inclusa esclusivamente la versione su rivista, in quanto considerata più matura e completa.

| Criteri di Inclusione | Razionale |
|---|---|
| **CI1: Pertinenza Tematica** | Assicura che lo studio sia centrato sull'argomento principale della tesi (Manutenzione Predittiva, Prognostica, RUL). |
| **CI2: Focus Metodologico** | Garantisce che l'articolo fornisca dettagli sostanziali sulle tecniche analitiche, che sono l'oggetto primario dell'analisi. |
| **CI3: Contesto Applicativo** | Mantiene la ricerca strettamente focalizzata sul perimetro industriale (manifatturiero) e tecnologico (lavorazioni meccaniche) definito. |
| **CI4: Tipologia di Pubblicazione** | Filtra per qualità e validità scientifica, includendo solo lavori validati dalla comunità accademica tramite peer-review. |
| **CI5: Lingua** | Assicura la piena, accurata e non ambigua comprensione del contenuto scientifico da parte del ricercatore. |
| **Criteri di Esclusione** | **Razionale** |
| **CE1: Irrilevanza Tematica/Semantica** | Elimina il "rumore" di fondo generato da omonimie o dall'uso delle keyword in contesti non pertinenti. |
| **CE2: Contesto Applicativo Errato** | Evita la dispersione dell'analisi su settori (es. civile, aerospaziale, medico) che esulano dallo scopo della tesi. |
| **CE3: Mancanza di Dettaglio Metodologico** | Esclude lavori puramente descrittivi o concettuali, non sufficientemente dettagliati per un'analisi metodologica approfondita. |
| **CE4: Tipologia di Pubblicazione Non Ammessa** | Mantiene l'omogeneità e l'elevato standard qualitativo del corpus di studi finale, escludendo la letteratura grigia dall'analisi formale. |
| **CE5: Testo Completo Non Reperibile** | Criterio di natura pragmatica per escludere studi che non possono essere analizzati nella loro interezza e in modo affidabile. |
| **CE6: Duplicazione/Ridondanza dei Contenuti** | Evita di analizzare e conteggiare più volte la stessa ricerca, garantendo l'unicità e l'indipendenza dei dati raccolti. |

*Tabella 1 - Riepilogo dei Criteri di Inclusione ed Esclusione*

## 2.4. Processo di Selezione ed Estrazione dei Dati

Dopo aver definito la strategia di ricerca, il passo successivo è stato quello di applicarla per selezionare gli studi pertinenti e, successivamente, estrarne le informazioni chiave. Questo processo si è articolato in due fasi: la selezione degli articoli e la successiva estrazione strutturata dei dati.

Il processo di selezione degli studi ha seguito le fasi di Screening ed Eleggibilità previste dal protocollo PRISMA, come illustrato nel diagramma di flusso (*Figura 2*). Il corpus iniziale contava circa 2.150 record. Come primo passo, questi record sono stati importati nel software di gestione bibliografica Mendeley, che ha permesso di identificare e rimuovere 550 studi duplicati.

Successivamente, i 1.600 record unici rimanenti sono stati sottoposti a uno screening preliminare basato sulla lettura del titolo e dell'abstract. In questa fase, sono stati esclusi 1.275 articoli chiaramente non pertinenti, applicando i criteri di esclusione (CE1, CE2, CE4). Ad esempio, articoli come "Predictive models for stock market trends" o "Maintenance of concrete bridges" sono stati immediatamente scartati in quanto fuori dal perimetro tematico e applicativo. Nel dubbio sulla pertinenza di un articolo, si è preferito includerlo nella fase successiva per un'analisi più approfondita.

I 300 articoli che hanno superato questo primo screening sono stati quindi recuperati per un'analisi completa del testo integrale (*full-text*). Ogni articolo è stato letto interamente per verificare il pieno rispetto di tutti i criteri di inclusione (CI1-CI5) e l'assenza di criteri di esclusione (CE1-CE6). Questa analisi ha permesso di valutare aspetti non evidenti dall'abstract, come la robustezza della metodologia o la coerenza dei risultati. A seguito di questa lettura approfondita, sono stati esclusi ulteriori 180 articoli, mantenendo una documentazione dettagliata delle ragioni di esclusione per garantire la trasparenza del processo.

Una volta definito il set finale di 120 articoli, si è proceduto con il processo di estrazione dei dati. Per garantire coerenza e comparabilità, è stata progettata e utilizzata una scheda di estrazione dati strutturata, una pratica fondamentale nelle revisioni sistematiche [118]. Per ogni articolo incluso, sono state registrate le seguenti informazioni, definite in base alle domande di ricerca della tesi:

- Dati Anagrafici e Bibliografici: ID Unico per tracciabilità, Autore/i, Anno di pubblicazione, Titolo e Fonte.

- Caratterizzazione Metodologica:

    - Approccio Generale: Classificazione in "Tradizionale/Statistico", "Machine Learning (classico)", "Deep Learning" o "Ibrido".
    - Algoritmo/Tecnica Specifica: Il nome esatto della tecnica utilizzata (es. "ARIMA", "SVM con kernel RBF", "CNN 1D", "Modello di Weibull").
    - Obiettivo Primario: Lo scopo del modello (es. "Diagnosi/Rilevamento guasti", "Prognosi/Stima della RUL").
    - Dati di Input: La tipologia di dati usati dal modello (es. "Segnali di vibrazione", "Dati storici di guasto").

- Caratterizzazione del Contesto Applicativo:

    - Processo Specifico: Il processo meccanico oggetto di studio (es. "Tornitura di acciai legati", "Fresatura di alluminio").
    - Metodologia di Validazione: Il tipo di dati usati per la validazione (es. "Dati sperimentali da banco prova", "Dati da linea di produzione reale", "Dataset pubblico di riferimento").

- Sintesi Qualitativa: Un breve riassunto del contributo scientifico principale e dei risultati quantitativi più significativi riportati (es. "accuratezza del 98% nella classificazione", "errore medio del 5% nella stima della RUL").

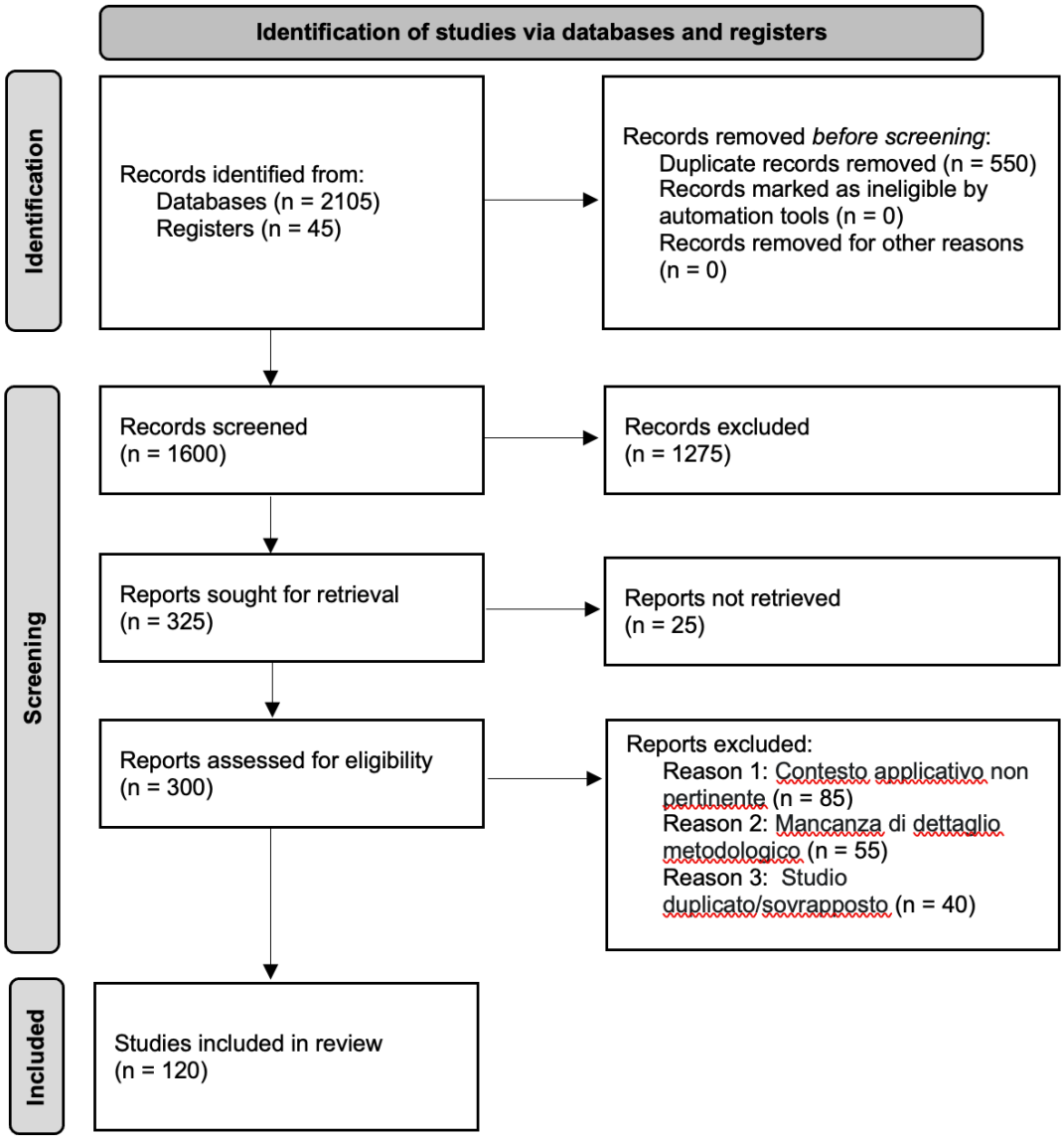


*Figura 2 – Diagramma di Flusso PRISMA*

# Capitolo 3: Fondamenti e Approcci Tradizionali alla Manutenzione Predittiva

Con il termine "tradizionali" si intende quell'insieme di approcci, prevalentemente basati su modelli statistici, stocastici e di affidabilità, che hanno costituito la spina dorsale della disciplina prima dell'avvento delle tecniche di Machine Learning e Intelligenza Artificiale.
Analizzare a fondo questi approcci non è un semplice esercizio storico su tecniche superate. Al contrario, è un passaggio fondamentale per capire l'evoluzione della disciplina, per almeno tre ragioni. Innanzitutto, molte di queste tecniche sono ancora oggi molto utilizzate nell'industria, specialmente quando i dati sono pochi o quando è richiesta una forte interpretabilità del modello. In secondo luogo, i principi su cui si basano (come l'analisi statistica e probabilistica) costituiscono le fondamenta su cui si sono sviluppate anche le tecniche più moderne. Infine, e questo è l'aspetto più importante per la mia tesi, è proprio analizzando i limiti di questi modelli tradizionali che si capisce perché la ricerca si sia dovuta spingere verso paradigmi più complessi come il Machine Learning.
Per questo motivo, l'analisi che seguirà in questo capitolo avrà un duplice scopo. Da un lato, sarà descrittiva, per spiegare come funzionano queste tecniche. Dall'altro, sarà critica, per metterne in luce i limiti e le debolezze.
Il capitolo inizierà definendo alcuni concetti chiave del Prognostics and Health Management (PHM), come la Vita Utile Residua (RUL). Verrà introdotta anche la distinzione fondamentale tra gli approcci basati sulla fisica del guasto (*physics-based*) e quelli basati sui dati (*data-driven*), in cui rientrano la maggior parte delle tecniche che analizzeremo.
Successivamente, l'analisi si concentrerà su tre principali famiglie di metodologie tradizionali:

1. Approcci basati sull'Ingegneria dell'Affidabilità: Si esamineranno i modelli che descrivono la probabilità di guasto nel tempo, come la distribuzione di Weibull, utilizzati per stimare l'affidabilità di componenti e sistemi sulla base di dati storici di guasto.

2. Approcci basati sull'Analisi delle Serie Storiche: Verranno analizzate le tecniche che modellano i dati di sensori (es. vibrazioni, temperatura) come sequenze temporali, cercando di estrapolarne il trend futuro per prevedere il superamento di una soglia

critica. In questo contesto, verrà discusso il modello ARIMA (Auto-Regressive Integrated Moving Average) come esempio paradigmatico.

3. Approcci basati su Processi Stocastici: Infine, si esploreranno i modelli che rappresentano il degrado di un sistema come una transizione probabilistica tra diversi stati di salute, come le Catene di Markov e le loro evoluzioni.

## 3.1. Concetti Fondamentali: Prognostics and Health Management (PHM)

Per implementare la Manutenzione Predittiva è necessario un framework ingegneristico strutturato, noto come Prognostics and Health Management (PHM). Il PHM è l'architettura concettuale e tecnologica che orchestra l'intero processo, dalla raccolta dati all'azione, al fine di valutare lo stato di salute di un sistema e prevederne il degrado futuro.

Sviluppatosi in settori critici come l'aerospaziale, il PHM si sta oggi diffondendo nel manifatturiero grazie alle tecnologie dell'Industria 4.0 (IoT, Big Data). La sua essenza risiede in un cambio di prospettiva fondamentale: dal paradigma diagnostico, che risponde alla domanda reattiva *"Cosa non funziona adesso?"*, al paradigma prognostico, che risponde alla domanda proattiva *"Cosa si guasterà nel futuro e quando?"*.

Un'architettura PHM trasforma i dati grezzi dei sensori in decisioni di manutenzione attraverso una catena del valore dell'informazione, che include tipicamente l'acquisizione e la pre-elaborazione dei dati, l'estrazione di feature (indicatori di salute), la diagnosi dello stato attuale e, infine, la prognosi del comportamento futuro, che è l'oggetto di analisi di questo capitolo.

### 3.1.1. Il Cuore della Prognostica

Al centro di ogni sforzo prognostico, come fulcro concettuale e obiettivo quantificabile, si trova la stima della Vita Utile Residua, internazionalmente nota con l'acronimo RUL (Remaining Useful Life). La RUL rappresenta la quantificazione predittiva del tempo operativo che un componente, un sottosistema o un intero impianto può ancora sostenere prima di non essere più in grado di adempiere alla sua funzione designata secondo le

specifiche di performance richieste. Formalmente, la RUL di un sistema all'istante di tempo attuale, *t*, è definita come la differenza tra il tempo previsto di fine vita, e il tempo attuale stesso. Sebbene la sua definizione matematica appaia semplice, la sua stima pratica è uno dei problemi più complessi e sfidanti dell'ingegneria moderna, poiché implica una proiezione nel futuro di un sistema fisico soggetto a innumerevoli variabili.

È cruciale, in primo luogo, definire con precisione il concetto di "fine vita" (EoL). L'End-of-Life non coincide necessariamente con un guasto catastrofico o con la rottura fisica del componente. Piuttosto, esso rappresenta il raggiungimento di uno stato di guasto funzionale, un punto in cui il sistema, pur potendo ancora operare, non soddisfa più i requisiti minimi di qualità, sicurezza o efficienza. Questo stato viene tipicamente definito tramite il superamento di una soglia di fallimento (Failure Threshold - FT) da parte di uno o più indicatori di salute (Health Indicators - HI). Un indicatore di salute è una grandezza quantitativa, estratta dai dati dei sensori. Ad esempio, per un utensile da taglio, l'HI potrebbe essere la rugosità superficiale del pezzo lavorato e la FT il valore massimo di rugosità tollerato dal cliente; per un cuscinetto, l'HI potrebbe essere il valore RMS (Root Mean Square) del segnale di vibrazione e la FT il limite di sicurezza definito dalle normative ISO. La definizione di queste soglie è essa stessa un'attività ingegneristica non banale, che richiede una profonda conoscenza del processo e dei suoi requisiti operativi. Pertanto, la stima della RUL si traduce nel problema di prevedere l'istante di tempo futuro in cui la traiettoria dell'indicatore di salute intersecherà la soglia di fallimento predefinita.

Il principale ostacolo scientifico nella stima della RUL è la gestione dell'incertezza, che pervade ogni aspetto del problema prognostico. Ignorare o sottostimare l'incertezza porta a previsioni "fragili", che possono essere pericolosamente ottimistiche o eccessivamente conservative. È possibile identificare diverse fonti primarie di incertezza. In primo luogo, vi è l'incertezza del modello, che nasce dalla discrepanza intrinseca tra il modello matematico utilizzato per la prognosi e il reale processo fisico. Qualsiasi modello, per quanto sofisticato, è un'astrazione che introduce semplificazioni e assunzioni. A ciò si aggiunge l'incertezza della misura, derivante dal fatto che i dati dei sensori, che alimentano il modello, sono sempre affetti da rumore stocastico, bias sistematici ed errori di calibrazione. La terza e forse più critica fonte è l'incertezza delle condizioni operative future. Un modello prognostico viene addestrato su dati passati, ma la previsione riguarda un futuro le cui condizioni di carico, temperatura e utilizzo possono variare in modo imprevedibile, influenzando drasticamente la velocità di degrado. Infine, esiste un'incertezza intrinseca dei processi e dello stato iniziale, poiché anche due componenti nominalmente identici presenteranno

sempre micro-variazioni a livello di materiali e geometria, e il loro stato di salute iniziale non sarà mai perfettamente identico o noto.

Data la presenza ineludibile di queste fonti di incertezza, presentare una stima della RUL come un singolo valore puntuale e deterministico è una pratica scientificamente incompleta e operativamente limitante. Un approccio molto più rigoroso e informativo consiste nel rappresentare la RUL in forma probabilistica, tipicamente attraverso una distribuzione di densità di probabilità (Probability Density Function - PDF). Una PDF della RUL non si limita a fornire la stima più probabile (la moda della distribuzione), ma caratterizza l'intera gamma di possibili valori futuri e la loro verosimiglianza, quantificando l'incertezza associata (la varianza o la larghezza della distribuzione). Questa rappresentazione probabilistica permette di calcolare intervalli di confidenza (es. "esiste una probabilità del 95% che la RUL sia compresa tra 80 e 120 ore"), che trasformano la previsione da un'affermazione secca a uno strumento strategico per la gestione del rischio. Un responsabile della manutenzione può così decidere se intervenire in anticipo per minimizzare il rischio di guasto, o attendere per massimizzare l'utilizzo del componente, basando la propria decisione su una valutazione quantitativa della probabilità di fallimento [103].

Per valutare e confrontare in modo oggettivo l'efficacia di diversi algoritmi prognostici, la comunità scientifica, spinta in particolare dai lavori pionieristici del NASA Ames Research Center, ha sviluppato un insieme di metriche di performance standardizzate. Queste metriche vanno oltre la semplice misurazione dell'errore puntuale e valutano la qualità complessiva del profilo prognostico nel tempo. Tra le più influenti vi sono il Prognostic Horizon (PH), che misura la "precocità" di un algoritmo, ovvero quanto tempo prima del guasto effettivo esso inizia a fornire previsioni stabili e accurate; l'α-λ Performance, che valuta se le previsioni della RUL rimangono confinate all'interno di un "cono di accuratezza" che si restringe man mano che ci si avvicina al momento del guasto; e la Relative Accuracy (RA), che calcola l'errore di previsione percentuale in specifici istanti di tempo. Una metrica aggregata particolarmente utile è la Cumulative Relative Accuracy (CRA), che media l'accuratezza relativa lungo la traiettoria di previsione, ma ponderando maggiormente gli errori commessi in prossimità del guasto. Questo riflette il fatto che un errore di previsione commesso molto in anticipo è meno critico di un errore commesso quando la decisione di intervento è imminente.

In sintesi, la stima della RUL è il fine ultimo e il cuore pulsante della prognostica. Non è un semplice calcolo, ma un processo inferenziale complesso che richiede la modellazione del degrado, la gestione rigorosa dell'incertezza e la valutazione trasparente delle performance.

### 3.1.2. Le Macro-Famiglie di Approcci Prognostici

Si identificano due macro-famiglie di approcci principali, quasi antitetiche nella loro logica di partenza: gli approcci basati sulla fisica del guasto (*Physics-of-Failure*) e quelli guidati dai dati (*Data-Driven*). Da questa dicotomia fondamentale, emerge una terza via, quella degli approcci ibridi, che si propone come una sintesi sinergica per superare i limiti intrinseci di ciascuna delle due famiglie originarie.

Il primo paradigma, quello degli approcci basati sulla fisica del guasto (PoF), adotta un approccio deduttivo, partendo dai "primi principi". La sua premessa fondamentale è che il degrado di un sistema fisico non sia un evento casuale, ma il risultato di processi governati da leggi fondamentali della fisica, della chimica e della meccanica. L'obiettivo di un modello PoF è quindi quello di catturare queste leggi in un insieme di equazioni matematiche esplicite, che descrivono la relazione causa-effetto tra le sollecitazioni operative (i carichi, le temperature, le pressioni) e l'accumulo di un danno specifico (la crescita di una cricca, la corrosione, l'usura di una superficie). Questo approccio mira a modellare i meccanismi del degrado. Di conseguenza, la costruzione di un tale modello richiede una profonda conoscenza a priori del sistema, dei materiali che lo compongono e delle interazioni fisico-chimiche che ne determinano il comportamento nel tempo. È un approccio "white-box", in cui ogni parametro e ogni equazione ha un significato fisico preciso e interpretabile.

Diametralmente opposto è il paradigma degli approcci guidati dai dati (Data-Driven). Questi adottano una logica induttiva, quasi fenomenologica, che non presume una conoscenza approfondita dei meccanismi di guasto sottostanti. La sua premessa è che, indipendentemente dalla complessità dei processi fisici interni, il degrado di un sistema si manifesti esternamente attraverso cambiamenti misurabili nei dati raccolti dai sensori. L'obiettivo di un modello data-driven è quindi quello di apprendere, direttamente e unicamente dai dati storici, le correlazioni e i pattern latenti che legano l'evoluzione dei segnali dei sensori allo stato di salute del sistema. Invece di modellare i *meccanismi*, questi approcci modellano i sintomi del degrado. Il sistema viene trattato come una "scatola nera" (*black-box*) o "grigia" (*grey-box*), il cui comportamento futuro viene inferito estrapolando le tendenze osservate nel passato. Questo paradigma è diventato dominante nell'era dell'Industria 4.0, grazie alla sua capacità di sfruttare la massiccia quantità di dati oggi disponibile.

Riconoscendo che entrambe queste filosofie presentano vantaggi e svantaggi complementari, la frontiera della ricerca si è progressivamente mossa verso lo sviluppo di approcci ibridi. Questo terzo paradigma nasce dalla consapevolezza che né la sola conoscenza fisica (spesso incompleta) né i soli dati (spesso rumorosi o insufficienti) sono in grado di fornire una soluzione universalmente ottimale. L'approccio ibrido si propone quindi come un paradigma di sintesi, che mira a orchestrare una sinergia tra i due mondi. L'idea è di utilizzare la conoscenza fisica per fornire una struttura di base al modello, una sorta di "scheletro" che ne garantisca la coerenza e la generalizzabilità, e di utilizzare i dati in tempo reale per "calibrare", correggere e adattare questo scheletro alla realtà specifica del singolo asset monitorato. Questo permette di ottenere un modello che sia più robusto, accurato e affidabile di quanto potrebbero esserlo i suoi componenti presi singolarmente.

La scelta tra un approccio PoF, data-driven o ibrido è una decisione strategica che dipende da un'attenta valutazione di diversi fattori specifici del problema in esame. Il primo e più fondamentale fattore di scelta risiede nel trade-off tra conoscenza del sistema e disponibilità di dati. In contesti dove la fisica del guasto è ben compresa e modellabile (es. sistemi semplici e critici) ma i dati di guasto sono rari o inesistenti (es. componenti di un satellite), un approccio PoF è spesso l'unica via percorribile. Al contrario, per sistemi estremamente complessi (es. un'intera linea di produzione) dove la modellazione fisica sarebbe proibitiva ma i dati operativi sono abbondanti, un approccio data-driven diventa la scelta naturale.

Un secondo fattore critico è il requisito di interpretabilità versus performance. I modelli PoF, essendo "white-box", offrono una trasparenza totale, un requisito spesso non negoziabile in settori certificati o dove le decisioni di manutenzione hanno implicazioni di sicurezza elevate. I modelli data-driven, specialmente quelli basati su Deep Learning, possono raggiungere performance predittive superiori in problemi complessi, ma spesso al costo di essere "black-box", rendendo difficile spiegare il "perché" di una certa previsione.

Infine, considerazioni su costi, tempo e capacità di generalizzazione giocano un ruolo chiave. Lo sviluppo di un modello PoF richiede un elevato investimento iniziale in termini di tempo e competenze di dominio. I modelli data-driven richiedono investimenti nella raccolta, pulizia e archiviazione dei dati. Inoltre, la capacità di un modello di generalizzare, ovvero di fornire previsioni accurate anche per condizioni operative diverse da quelle su cui è stato sviluppato, è tipicamente superiore negli approcci PoF e ibridi, mentre rappresenta una delle sfide principali per i modelli puramente data-driven.

### 3.1.3. L'Approccio Basato sulla Fisica del Guasto (PoF)

L'approccio basato sulla fisica del guasto, noto in letteratura come Physics-of-Failure (PoF) o più genericamente come approccio *model-based*, rappresenta il paradigma prognostico più classico e, in un certo senso, più ambizioso dal punto di vista ingegneristico. La sua filosofia fondamentale è intrinsecamente deduttiva: essa postula che il degrado e il guasto di un componente non siano fenomeni stocastici da osservare passivamente, ma processi deterministici o quasi-deterministici governati da leggi fondamentali della fisica, della chimica e della meccanica dei materiali. L'obiettivo ultimo di un approccio PoF non è quindi quello di trovare correlazioni nei dati, ma di comprendere e modellare matematicamente i meccanismi di danno sottostanti, come la fatica, la corrosione, l'usura, la propagazione di cricche o il degrado termico. Questo approccio tratta il sistema come una "scatola bianca" (*white-box*), un sistema il cui funzionamento interno è, almeno in linea di principio, completamente trasparente, noto e modellabile.

Il processo di costruzione di un modello PoF è un'attività ingegneristica complessa e multi-stadio. Il primo passo consiste nell'identificazione dei meccanismi di guasto dominanti. Attraverso analisi come la FMECA (Failure Mode, Effects, and Criticality Analysis) e una profonda conoscenza del dominio, l'ingegnere deve determinare quali processi fisici sono i principali responsabili del degrado del componente nelle sue specifiche condizioni operative. Successivamente, si procede alla formulazione del modello matematico. Questa fase richiede la traduzione delle leggi fisiche pertinenti in un insieme di equazioni, spesso differenziali, che legano le sollecitazioni operative (i *load*, come carichi meccanici, cicli termici, correnti elettriche) all'evoluzione di una variabile di danno interna. Un esempio paradigmatico in meccanica è la Legge di Paris-Erdogan, che modella la velocità di crescita di una cricca di fatica in funzione del range del fattore di intensificazione degli sforzi.

Una volta formulato, il modello deve essere parametrizzato e calibrato. Le equazioni contengono costanti e parametri che dipendono dalle proprietà specifiche dei materiali, dalla geometria del componente e dalle condizioni ambientali. La determinazione di questi parametri richiede spesso test di laboratorio dedicati o l'utilizzo di dati da manuali di ingegneria dei materiali. Infine, il modello viene integrato in un ciclo prognostico. Questo ciclo parte da una stima dello stato di danno attuale del componente (che può essere zero per un componente nuovo o un valore misurato per un componente in servizio) e, data una previsione delle future condizioni operative, integra numericamente nel tempo le equazioni di degrado per simulare l'evoluzione del danno. La stima della RUL viene quindi calcolata

come il tempo necessario affinché la variabile di danno simulata raggiunga un valore di soglia critico predefinito.

Il suo principale punto di forza è senza dubbio l'elevata interpretabilità. Essendo ogni componente del modello legato a un significato fisico, i risultati sono trasparenti e le cause di un guasto previsto possono essere ricondotte a meccanismi specifici. Questa caratteristica è non negoziabile in settori safety-critical, come quello aerospaziale o biomedicale, dove le decisioni di manutenzione devono essere giustificate e certificate. Inoltre, se il modello fisico cattura accuratamente la realtà, può raggiungere un'elevatissima accuratezza predittiva. Un altro vantaggio cruciale è la sua relativa indipendenza da grandi dataset di guasto storici. I modelli PoF possono essere sviluppati e utilizzati anche per componenti nuovi o estremamente affidabili, per i quali non esistono dati di "run-to-failure", rendendoli strumenti preziosi già in fase di progettazione (*prognostics-by-design*) per confrontare diverse alternative progettuali. Infine, questi modelli possiedono, in teoria, un'eccellente capacità di generalizzazione: essendo fondati su leggi fisiche universali, sono in grado di fornire previsioni valide anche per profili di missione o condizioni operative diverse da quelle viste in precedenza, a patto che rimangano all'interno del dominio di validità del modello fisico stesso.

Tuttavia, questa stessa ambizione di modellare la realtà fisica è anche la fonte delle più significative limitazioni dell'approccio PoF, che ne restringono l'applicabilità pratica. La complessità e il costo di sviluppo rappresentano la barriera più alta. La creazione di un modello fisico ad alta fedeltà è una sfida formidabile che richiede tempo, risorse e, soprattutto, una profonda e rara expertise multidisciplinare. Per molti sistemi industriali complessi, i meccanismi di guasto non sono completamente compresi, o sono talmente interconnessi da rendere la loro modellazione matematica proibitiva. Una seconda, critica debolezza è l'incapacità di modellare l'imprevisto. Un modello PoF è, per sua natura, "cieco" a qualsiasi fenomeno che non sia stato esplicitamente incluso nelle sue equazioni. Non può catturare effetti di degrado emergenti, interazioni complesse e non lineari tra più meccanismi di guasto (es. corrosione sotto sforzo), o guasti indotti da cause esterne come un errore di montaggio o un danno da impatto.

Inoltre, la difficoltà di una parametrizzazione accurata può compromettere la validità del modello. Le proprietà dei materiali possono variare da lotto a lotto e le condizioni al contorno reali possono differire da quelle idealizzate in laboratorio. Infine, il costo computazionale può essere un ostacolo, specialmente per modelli che si basano su simulazioni complesse come quelle agli elementi finiti (FEM), rendendone difficile

l'implementazione in sistemi embedded o per applicazioni che richiedono una risposta in tempo reale.

In sintesi, l'approccio basato sulla fisica del guasto non deve essere visto come una metodologia superata, ma come uno strumento specialistico di straordinaria potenza, il "gold standard" della prognostica laddove la fisica è nota e l'interpretabilità è sovrana. La sua applicazione rimane fondamentale nella progettazione di sistemi ad alta affidabilità e nell'analisi di guasto fondamentale. Ciononostante, la sua impraticabilità in una vasta gamma di sistemi industriali complessi, la cui fisica è sconosciuta o troppo intricata da modellare, ha creato un vuoto metodologico evidente.

### 3.1.4. L'Approccio Guidato dai Dati

In netto contrasto con la filosofia deduttiva e basata sui primi principi dell'approccio Physics-of-Failure (PoF), l'approccio guidato dai dati (Data-Driven) rappresenta un cambiamento paradigmatico fondamentale, fondato su una logica prettamente induttiva. La premessa fondamentale di questo paradigma non è la comprensione a priori dei meccanismi fisici di degrado, ma l'assunto che, indipendentemente dalla loro complessità interna, i processi di degrado di un sistema si manifestino esternamente attraverso cambiamenti misurabili e pattern identificabili nei dati raccolti dai sensori. Invece di modellare le cause del guasto, l'approccio data-driven si concentra sul modellare i suoi sintomi. Il sistema fisico viene trattato come una "scatola nera" (*black-box*) o, nei casi più sofisticati, "grigia" (*grey-box*), il cui comportamento futuro viene inferito estrapolando le tendenze e le correlazioni apprese unicamente dai dati storici [103].

L'ascesa e l'attuale predominio di questo paradigma nella letteratura PHM non sono casuali, ma sono la diretta conseguenza della convergenza di tre potenti driver tecnologici che caratterizzano l'era dell'Industria 4.0: la crescita di sensoristica a basso costo, che permette di monitorare una vasta gamma di parametri fisici; lo sviluppo di infrastrutture per la gestione di Big Data, capaci di archiviare e processare enormi volumi di dati di serie temporali; e, soprattutto, i progressi esponenziali negli algoritmi di Intelligenza Artificiale e Machine Learning, che forniscono gli strumenti analitici per estrarre valore da questi dati [61].

Il flusso di lavoro tipico di un approccio data-driven per la prognostica si articola in due fasi principali: una fase di addestramento (training) offline e una fase di test (o inferenza) online. Durante la fase di addestramento, il modello viene "istruito" utilizzando un dataset storico che contiene dati di sensori raccolti durante l'intero ciclo di vita di più esemplari dello stesso componente, idealmente fino al loro guasto (*run-to-failure data*). Questo processo di apprendimento permette al modello di associare specifici pattern nei dati (le *features*) a corrispondenti stati di salute o valori di RUL. Una volta addestrato, il modello viene implementato online. Durante la fase di test, esso riceve in tempo reale i dati dal componente in funzione, li elabora e fornisce una stima della RUL attuale.

Gli approcci data-driven si sono evoluti in due ondate principali, che verranno analizzate in dettaglio nel corso di questa tesi. La prima ondata è quella dei Modelli Statistici Tradizionali, che costituisce il cuore di questo capitolo. Questa famiglia include tecniche che si basano su principi statistici consolidati per modellare i dati. Esempi includono l'analisi di regressione per mappare direttamente le feature alla RUL, i modelli di serie temporali come ARIMA (Auto-Regressive Integrated Moving Average) per estrapolare trend di degrado, e modelli di affidabilità che si basano su distribuzioni statistiche di guasto, come la distribuzione di Weibull. Questi modelli sono spesso considerati "grey-box" perché, pur essendo data-driven, si basano su assunzioni statistiche esplicite che ne rendono, in parte, interpretabile il comportamento.

La seconda e più recente ondata è quella dei Modelli di Intelligenza Artificiale e Machine Learning, questa famiglia impiega algoritmi molto più flessibili e potenti, capaci di apprendere relazioni non lineari e complesse senza necessità di assunzioni a priori. Include algoritmi "classici" come le Reti Neurali Artificiali (ANN), le Support Vector Machines (SVM) e i Random Forest, che hanno dimostrato grande efficacia in problemi di classificazione (diagnosi) e regressione (stima della RUL) [18]. Più recentemente, l'avvento del Deep Learning ha introdotto architetture ancora più sofisticate, come le Reti Neurali Convoluzionali (CNN), ideali per analizzare dati strutturati come le immagini spettrali dei segnali di vibrazione, e le Reti Neurali Ricorrenti (RNN), in particolare le varianti Long Short-Term Memory (LSTM) e Gated Recurrent Unit (GRU), che sono state specificamente progettate per modellare sequenze temporali e sono diventate lo stato dell'arte per la stima della RUL basata su dati di sensori.

Il principale punto di forza che ha decretato il successo degli approcci data-driven è la loro applicabilità quasi universale. Essi possono essere implementati per qualsiasi sistema da cui sia possibile raccogliere dati di buona qualità, anche quando la fisica del guasto è

sconosciuta, troppo complessa da modellare o quando il sistema è affetto da molteplici meccanismi di degrado interagenti. La loro rapidità di sviluppo, data la disponibilità di librerie software open-source (es. Scikit-learn, TensorFlow, PyTorch) e di dati sufficienti, può essere notevolmente superiore a quella richiesta per la creazione di un modello PoF. Inoltre, questi modelli hanno la straordinaria capacità di scoprire pattern e correlazioni latenti nei dati, relazioni che potrebbero non essere evidenti nemmeno a un esperto di dominio, portando a nuove intuizioni sul processo di degrado.

Tuttavia, questo potere e questa flessibilità hanno un costo e presentano delle sfide significative. Il "tallone d'Achille" di ogni modello data-driven è la sua assoluta dipendenza dai dati. La celebre massima "garbage in, garbage out" è qui più valida che mai. La performance di questi modelli è intrinsecamente legata alla quantità, qualità e rappresentatività del dataset di addestramento. Essi richiedono dataset di grandi dimensioni che coprano un'ampia gamma di condizioni operative e, soprattutto, che contengano un numero sufficiente di esempi di degrado completo fino al guasto. L'acquisizione di tali dati *run-to-failure* è spesso l'ostacolo più grande in un contesto industriale, poiché è costosa, richiede tempo e, per ragioni di sicurezza e produzione, le aziende tendono a evitare di far funzionare i macchinari fino alla rottura [48].

Una seconda, profonda sfida è la scarsa interpretabilità della maggior parte dei modelli di Machine Learning avanzati. Modelli come le reti neurali profonde sono spesso descritti come "black-box" perché, sebbene possano fornire previsioni estremamente accurate, è estremamente difficile comprendere la logica interna che li ha portati a una specifica decisione. Questa mancanza di trasparenza può rappresentare una barriera insormontabile per la loro adozione in settori critici e può minare la fiducia degli operatori e dei responsabili della manutenzione [1].

Infine, sussiste un costante rischio di scarsa generalizzazione. Un modello addestrato meticolosamente su dati raccolti da un macchinario in specifiche condizioni operative potrebbe fallire drasticamente se tali condizioni dovessero cambiare (un problema noto come *concept drift* o *domain shift*). Garantire che un modello data-driven sia robusto e in grado di generalizzare a nuove condizioni operative o a macchinari leggermente diversi è uno dei principali filoni di ricerca attuali.

In conclusione, l'approccio guidato dai dati ha rivoluzionato il campo della prognostica, offrendo strumenti potenti per affrontare problemi di una complessità inaccessibile ai modelli basati sulla fisica. Tuttavia, il suo successo non è garantito e dipende criticamente

dalla disponibilità di dati di alta qualità e dalla consapevolezza delle sue limitazioni intrinseche in termini di interpretabilità e generalizzazione.

### 3.1.5. L'Approccio Ibrido

L'analisi critica delle due macro-famiglie di approcci prognostici — basati sulla fisica del guasto (PoF) e guidati dai dati (data-driven) — rivela un quadro di punti di forza e di debolezza quasi perfettamente complementari. Da un lato, i modelli PoF offrono interpretabilità, rigore scientifico e capacità di generalizzazione, ma soffrono di complessità di sviluppo e sono "ciechi" a fenomeni non modellati. Dall'altro, i modelli data-driven vantano flessibilità, applicabilità universale e la capacità di scoprire pattern complessi, ma sono afflitti da una forte dipendenza dai dati, da problemi di interpretabilità e da rischi di scarsa robustezza al di fuori del loro dominio di addestramento. Questa complementarità ha naturalmente spinto la frontiera della ricerca PHM verso un terzo paradigma, quello degli approcci ibridi, che non si propone come un'alternativa, ma come una sintesi sinergica dei due precedenti.

La filosofia fondamentale dell'approccio ibrido è che né la sola conoscenza fisica a priori (spesso incompleta o idealizzata) né i soli dati (spesso rumorosi, scarsi o raccolti in condizioni non rappresentative) sono sufficienti per costruire un sistema prognostico veramente robusto, accurato e affidabile in contesti industriali reali. L'obiettivo di un modello ibrido è quindi quello di orchestrare una fusione intelligente tra questi due mondi, utilizzando la conoscenza fisica per guidare, regolarizzare e dare una struttura al modello, e, allo stesso tempo, sfruttando i dati in tempo reale per calibrare, adattare e correggere le previsioni di tale modello [34]. Questo paradigma mira a ottenere il "meglio dei due mondi": la trasparenza e la generalizzabilità dei modelli PoF, unite alla precisione e alla capacità di adattamento dei modelli data-driven.

Sebbene non esista una tassonomia universalmente accettata, è possibile identificare alcune tipologie di ibridazione ricorrenti [4] [22]:

1. Ibridazione Sequenziale (Physics-informed Data-driven): In questa configurazione, il modello fisico e il modello data-driven operano in sequenza. Un'implementazione comune vede il modello fisico utilizzato per generare un dataset sintetico di alta qualità, simulando diverse traiettorie di degrado in varie condizioni operative. Questo

dataset, ricco e completo, viene poi utilizzato per addestrare un modello data-driven (es. una rete neurale). Questo approccio è particolarmente utile quando i dati reali di *run-to-failure* sono scarsi o inesistenti. Il modello fisico agisce come un "generatore di conoscenza" che permette al modello data-driven di apprendere da una realtà simulata ma fisicamente coerente, superando il problema della scarsità dei dati.

2. Ibridazione Parallela (Fusion-based): In questa architettura, i modelli PoF e data-driven operano in parallelo, fornendo ciascuno una stima indipendente della RUL. Un "modulo di fusione" (*fusion module*), che può essere basato su tecniche come la media pesata, la logica fuzzy o un ulteriore modello di Machine Learning, combina queste stime multiple per produrre una previsione finale più robusta. Il peso assegnato a ciascun modello può essere statico o dinamico, ad esempio dando più importanza al modello fisico nelle fasi iniziali del degrado (dove i dati sono poco informativi) e aumentando progressivamente il peso del modello data-driven man mano che i sintomi del degrado diventano più evidenti nei dati dei sensori.

3. Ibridazione Integrata (Physics-guided Data-driven): Questa è forse la forma di ibridazione più profonda e promettente. Invece di trattare i due modelli come entità separate, si cerca di integrare la conoscenza fisica direttamente all'interno della struttura o del processo di addestramento del modello data-driven. Un esempio è l'uso di tecniche di filtraggio bayesiano, che rappresentano lo stato dell'arte per questa tipologia di fusione. In questo schema, il modello fisico viene utilizzato per definire l'equazione di processo (o di stato), che descrive come lo stato di salute del sistema si prevede che evolva nel tempo. I dati dei sensori in tempo reale vengono invece utilizzati per definire l'equazione di misura, che mette in relazione lo stato di salute interno (non direttamente osservabile) con le grandezze misurate. Algoritmi come il Filtro di Kalman (per sistemi lineari) e le sue estensioni non lineari (EKF, UKF) sono stati ampiamente utilizzati. Tuttavia, il Filtro Particellare (Particle Filter) è emerso come lo strumento d'elezione per problemi di prognostica complessi, grazie alla sua capacità di gestire sistemi altamente non lineari e distribuzioni di probabilità non gaussiane [78]. Il Filtro Particellare rappresenta la distribuzione di probabilità dello stato di salute tramite un insieme di "particelle" pesate, che vengono propagate nel tempo usando il modello fisico (fase di predizione) e poi ri-pesate e ri-

campionate sulla base dei nuovi dati dei sensori (fase di aggiornamento). Questo permette un aggiornamento ricorsivo e robusto della stima della RUL, che tiene conto sia della fisica del sistema sia dell'evidenza fornita dai dati. Un'altra frontiera dell'ibridazione integrata è quella della Physics-Informed Neural Networks (PINN), in cui le equazioni differenziali che governano il modello fisico vengono incorporate direttamente nella funzione di costo (loss function) di una rete neurale, obbligando la rete a trovare una soluzione che non solo si adatti ai dati, ma che rispetti anche le leggi fisiche note [86].

Il vantaggio primario degli approcci ibridi è la loro capacità di mitigare le debolezze intrinseche dei modelli puri. L'integrazione della fisica aiuta a superare il problema della scarsità dei dati e migliora la capacità di generalizzazione dei modelli data-driven, evitando previsioni fisicamente implausibili. D'altro canto, l'aggiornamento basato sui dati permette al modello di correggere le imprecisioni del modello fisico, di adattarsi a condizioni operative reali e di catturare effetti non modellati, migliorando drasticamente l'accuratezza della previsione finale. Questa sinergia porta a una maggiore robustezza e affidabilità del sistema prognostico nel suo complesso.
In conclusione, il paradigma ibrido non deve essere visto come una semplice terza opzione, ma come la naturale evoluzione e la sintesi matura della disciplina prognostica. Esso riconosce che la conoscenza ingegneristica e l'evidenza empirica non sono in competizione, ma sono due fonti di informazione complementari che, se fuse in modo intelligente, possono portare a sistemi di monitoraggio e previsione di una potenza e un'affidabilità superiori. La crescente complessità dei sistemi industriali moderni e la richiesta di decisioni di manutenzione sempre più critiche e ottimizzate rendono lo sviluppo e l'adozione di questi approcci ibridi una delle direzioni più promettenti e strategiche per la ricerca e l'applicazione industriale futura.

## 3.2. Metodologie Basate sulla Statistica e l'Affidabilità

Prima dell'attuale ondata di innovazione legata al Machine Learning, la Manutenzione Predittiva si basava già su un approccio guidato dai dati, anche se con una filosofia molto diversa. Questa "prima ondata" di metodologie non utilizzava algoritmi complessi capaci di apprendere da soli, ma si affidava alla teoria della probabilità, alla statistica e all'ingegneria

dell'affidabilità. Questi approcci, che in questa tesi definisco "tradizionali", sono il primo tentativo sistematico di modellare e prevedere il degrado e i guasti applicando principi statistici a dati storici.

Analizzarli non è un semplice esercizio storico. Al contrario, è un passaggio fondamentale per capire l'evoluzione della disciplina, perché non solo costituiscono le basi su cui si sono sviluppate le tecniche più recenti, ma sono ancora oggi strumenti validi e molto utilizzati, specialmente quando i dati sono pochi (*small data*) o quando l'interpretabilità del modello è un requisito fondamentale [103].

Per orientarsi in questo campo, è essenziale fare una prima distinzione basata sul tipo di dati disponibili, perché questo determina quale famiglia di modelli statistici sia più adatta.

Da un lato, ci sono i dati di evento (*event data*), noti anche come dati di vita o lifetime data. Questa tipologia di dati non descrive l'evoluzione continua di un sistema, ma registra unicamente gli istanti di tempo in cui si verifica un evento di interesse, tipicamente un guasto. I dati si presentano quindi nella forma di una collezione di tempi al guasto (es. "il componente A si è guastato dopo 1250 ore, il componente B dopo 1380 ore, ecc."). Questi dati possono essere "completi", se si è osservato il guasto di ogni unità del campione, o, più frequentemente, "censurati", se l'osservazione termina prima che tutte le unità si siano guastate (ad esempio, perché lo studio è terminato o perché alcune unità sono state rimosse per altre ragioni). Questa tipologia di dati, per sua natura aggregata e focalizzata sull'evento finale, è il carburante per i modelli classici dell'ingegneria dell'affidabilità e dell'analisi di sopravvivenza.

Dall'altro lato, vi sono i dati di monitoraggio della condizione (*condition monitoring data*). Questi dati, resi sempre più accessibili dalla sensoristica moderna, consistono in serie temporali continue o quasi-continue di misurazioni di parametri fisici (es. vibrazioni, temperatura, pressione, segnali acustici, assorbimento di corrente). A differenza dei dati di evento, i dati di monitoraggio della condizione non registrano solo il "quando" del guasto, ma descrivono il "come" il sistema si è evoluto nel tempo, fornendo una traccia, un "elettrocardiogramma", del suo processo di degrado. L'analisi di queste traiettorie di dati, spesso dopo un'opportuna fase di estrazione di indicatori di salute (*Health Indicators*), è il dominio dei modelli di regressione, dei modelli di analisi di serie storiche e dei modelli basati su processi stocastici [48].

### 3.2.1. Approcci basati sull'Ingegneria dell'Affidabilità e Analisi di Sopravvivenza

All'inizio della gestione scientifica della manutenzione, quando il monitoraggio continuo delle condizioni era ancora relegato al regno della fantascienza tecnologica, la necessità di prendere decisioni razionali e non aneddotiche sulla sostituzione dei componenti ha dato vita alla disciplina dell'ingegneria dell'affidabilità. Questa branca dell'ingegneria, nata dalle esigenze impellenti dei settori militare e aerospaziale durante la Seconda Guerra Mondiale e la Guerra Fredda, dove il fallimento di un singolo componente poteva determinare il fallimento di un'intera missione, rappresenta il primo e più fondamentale tentativo di applicare il rigore della statistica matematica per comprendere, quantificare e, in una certa misura, prevedere i fenomeni di guasto. La sua filosofia è intrinsecamente legata a quella dell'analisi di sopravvivenza (*survival analysis*), sua disciplina gemella in ambito biostatistico, e condivide con essa l'obiettivo primario: non prevedere il destino di un singolo individuo, ma caratterizzare il comportamento di un'intera popolazione nel tempo.
Invece di concentrarsi sulla traiettoria di degrado di un singolo componente, l'analisi di affidabilità adotta una prospettiva aggregata, quasi demografica. Il suo input fondamentale sono i dati di evento (*event data*), ovvero i tempi al guasto osservati su un campione rappresentativo della popolazione di interesse. Questi dati possono essere "completi", se si è osservato il guasto di ogni unità del campione, o, più frequentemente nel mondo reale, "censurati", ovvero incompleti a causa della fine dello studio o della rimozione di unità per ragioni diverse dal guasto. L'obiettivo non è una prognosi in tempo reale, ma una caratterizzazione statistica del processo di guasto della popolazione, al fine di rispondere a domande di natura strategica, logistica e finanziaria.
Per tradurre i dati di guasto in conoscenza ingegneristica, l'analisi di affidabilità si fonda su un insieme di funzioni matematiche interconnesse, che insieme forniscono un ritratto completo del comportamento della popolazione. Definendo $T$ come la variabile casuale continua che rappresenta il tempo al guasto, le funzioni fondamentali sono:

- La Funzione di Densità di Probabilità (PDF), $f(t)$, che descrive la distribuzione dei guasti nel tempo. L'area sottesa alla curva $f(t)$ tra due istanti di tempo, rappresenta la probabilità che un componente scelto a caso dalla popolazione si guasti in quell'intervallo.

- La Funzione di Ripartizione (CDF), *F(t)*, che è l'integrale della PDF e rappresenta la probabilità cumulativa di guasto. *F(t)* esprime la probabilità che un'unità si sia già guastata entro il tempo *t*. È una funzione monotona non decrescente che va da 0 a 1.

- La Funzione di Affidabilità (Reliability Function), *R(t)*, definita come il complemento a uno della CDF, ovvero *R(t) = 1 - F(t)*. Questa è forse la metrica più intuitiva e utilizzata, poiché esprime la probabilità che un'unità sia ancora funzionante ("sopravvissuta") al tempo *t*.

- La Funzione di Rischio (Hazard Function), *h(t)* o *λ(t)*. Questo è il concetto più sofisticato e potente. Formalmente definita come *h(t) = f(t) / R(t)*, essa quantifica la propensione istantanea al guasto al tempo *t*, condizionata alla sopravvivenza fino a quell'istante. Non è una probabilità, ma un tasso di guasto istantaneo. L'analisi della forma della funzione di rischio è uno strumento diagnostico potentissimo: un andamento decrescente suggerisce mortalità infantile, uno costante suggerisce guasti casuali, e uno crescente suggerisce un processo di invecchiamento e usura, come descritto dal celebre modello concettuale della curva a vasca da bagno (*bathtub curve*).

Per dare una forma matematica a queste funzioni, i ricercatori hanno proposto diverse distribuzioni di probabilità (Esponenziale, Lognormale, Normale). Tuttavia, nessuna ha raggiunto la popolarità e l'applicabilità universale della distribuzione di Weibull. Proposta dall'ingegnere e matematico svedese Waloddi Weibull nel 1951, la sua forza risiede in una flessibilità senza pari, che le permette di adattarsi a una vasta gamma di dati di vita e di rappresentare tutti e tre i regimi della curva a vasca da bagno [119].
Nella sua forma più comune a due parametri, la distribuzione è definita da:

- Un parametro di forma (β, beta), anche detto "pendenza di Weibull". Questo parametro è adimensionale e determina la forma della distribuzione e, di conseguenza, la natura del processo di guasto. La sua interpretazione è di cruciale importanza strategica. Se $\beta < 1$, la funzione di rischio è decrescente, indicando una mortalità infantile. In questo regime, i componenti deboli tendono a guastarsi presto, e quelli che sopravvivono sono più robusti. La manutenzione preventiva basata sul tempo è controproducente. Se $\beta = 1$, la funzione di rischio è costante, e la

distribuzione di Weibull si riduce alla distribuzione Esponenziale. Questo è il modello dei guasti puramente casuali. La manutenzione preventiva non offre alcun vantaggio. Se $\beta > 1$, la funzione di rischio è crescente, indicando un processo di usura e invecchiamento. Questo è l'unico scenario in cui una politica di sostituzione preventiva può essere statisticamente giustificata. Maggiore è il valore di $\beta$, più il processo di usura è rapido e più i guasti si concentrano attorno a un'età specifica.

- Un parametro di scala ($\eta$, eta), anche detto vita caratteristica. Questo parametro ha la stessa unità di misura del tempo e rappresenta il tempo al quale il 63.2% della popolazione si è guastato. Funge da "stiramento" della distribuzione lungo l'asse dei tempi.

Esiste anche una versione a tre parametri, che aggiunge un parametro di posizione ($\gamma$, gamma), che rappresenta un periodo di vita iniziale garantito durante il quale la probabilità di guasto è nulla. Questo è utile per modellare fenomeni come la fatica, che richiedono un certo numero di cicli prima che il danno inizi ad accumularsi.

La stima dei parametri di Weibull a partire dai dati di guasto storici è il passo operativo fondamentale. Il metodo più intuitivo è quello grafico, basato sulle carte di probabilità di Weibull (*Weibull probability plots*). Attraverso una trasformazione logaritmica degli assi, questa carta ha la proprietà di linearizzare i dati che seguono una distribuzione di Weibull. I dati di guasto vengono plottati su questa carta e una retta di regressione viene adattata ai punti. La pendenza di questa retta fornisce una stima del parametro di forma $\beta$, mentre la sua intercetta è legata al parametro di scala $\eta$. Sebbene potente dal punto di vista visivo e diagnostico, questo metodo è meno rigoroso di approcci puramente statistici.

Il metodo di stima considerato il "gold standard" è la Stima di Massima Verosimiglianza (*Maximum Likelihood Estimation* - MLE). Questo metodo, pur essendo computazionalmente più intensivo, cerca i valori dei parametri ($\beta$, $\eta$, $\gamma$) che massimizzano la probabilità (la "verosimiglianza") di aver osservato proprio il set di dati di guasto a disposizione. La MLE fornisce stime con proprietà statistiche desiderabili (consistenza, efficienza asintotica) e permette di calcolare intervalli di confidenza per i parametri stimati, quantificando l'incertezza dovuta alla limitatezza del campione [68].

Una volta stimati i parametri, è possibile utilizzare il modello di Weibull per calcolare una miriade di metriche di affidabilità e per ottimizzare le strategie di manutenzione, ad esempio

determinando l'intervallo di sostituzione preventiva che minimizza il costo totale per unità di tempo.

L'analisi di Weibull classica si basa su un'assunzione fondamentale: che la popolazione di componenti sia omogenea e che l'unico fattore che influenza il guasto sia il tempo. Nella realtà industriale, questa assunzione è spesso troppo semplicistica. La vita di un componente può essere significativamente influenzata da una serie di fattori, o covariate, come le condizioni operative (temperatura, carico), le caratteristiche del materiale (fornitore, lotto di produzione) o le pratiche di manutenzione. Per incorporare l'effetto di queste covariate sull'affidabilità, la statistica ha sviluppato potenti modelli di regressione per dati di sopravvivenza.

Il più celebre, elegante e influente di questi è il modello a rischi proporzionali di Cox, introdotto da Sir David Cox in un articolo rivoluzionario del 1972. La genialità del modello di Cox risiede nel suo essere semi-parametrico. Esso scompone la funzione di rischio di un individuo *i* in due parti. La prima parte,  è la funzione di rischio di base (*baseline hazard*), una funzione arbitraria e non parametrica che descrive come il rischio cambia nel tempo per un individuo "di riferimento" (con tutte le covariate pari a zero). La seconda parte, il termine esponenziale, è la parte parametrica e modella come le $k$ covariate (le $x$) modificano moltiplicativamente questo rischio di base. I coefficienti $\beta$, stimati dal modello, sono i log-hazard ratios e quantificano l'impatto di ciascuna covariata: un $\beta$ positivo indica che un aumento unitario della covariata aumenta il rischio, mentre un $\beta$ negativo indica l'opposto [23]. Questo modello è straordinariamente potente perché permette di valutare l'impatto di diverse condizioni operative sull'affidabilità senza dover fare alcuna assunzione sulla forma della distribuzione di guasto sottostante.

Il punto di forza inattaccabile degli approcci basati sull'analisi di affidabilità risiede nella loro solida base teorica e nella loro comprovata utilità come strumento di pianificazione strategica. Essi sono la scienza attuariale dei componenti meccanici ed elettronici. Forniscono gli strumenti per ottimizzare le politiche di sostituzione di intere flotte, per gestire in modo scientifico le scorte di magazzino, per prezzare contratti di servizio e per definire clausole di garanzia basate su una solida quantificazione del rischio a livello di popolazione.

Tuttavia, il loro limite fondamentale, concettuale e insormontabile, è proprio quello di essere una scienza della popolazione, e non dell'individuo. La loro logica è analoga a quella delle tavole di mortalità usate dalle compagnie di assicurazione sulla vita. Una tavola attuariale può prevedere con straordinaria accuratezza l'aspettativa di vita media di un uomo di 50 anni,

non fumatore e residente in una certa area geografica. Tuttavia, quella stessa tavola è completamente cieca di fronte alla salute specifica di quel singolo uomo. Non sa se egli ha un'ipertensione non diagnosticata o uno stile di vita particolarmente stressante (dati di condizione) che lo mettono a rischio imminente, informazioni che la statistica di popolazione, per sua stessa definizione, non può e non intende catturare.

Allo stesso modo, un'analisi di Weibull, per quanto sofisticata, può prevedere con precisione il comportamento medio di una flotta di cuscinetti, ma è strutturalmente incapace di distinguere tra un singolo cuscinetto che ha operato per 1900 ore in condizioni ideali e un altro che, dopo le stesse 1900 ore, è sull'orlo del guasto a causa di sovraccarichi e contaminazione. Per il modello di affidabilità, entrambi sono semplicemente "cuscinetti di 1900 ore". Questo "peccato originale", questa cecità alla condizione individuale e alla storia operativa unica di ogni singolo asset, è la debolezza intrinseca che ha creato la necessità impellente di sviluppare un paradigma prognostico differente, basato non più solo sui dati di evento, ma sui ricchi flussi di dati provenienti dal monitoraggio della condizione in tempo reale.

### 3.2.2. Approcci basati sull'Analisi delle Serie Storiche

Con la crescente disponibilità di dati provenienti da sistemi di monitoraggio della condizione (*Condition Monitoring Systems*), la comunità scientifica e ingegneristica ha potuto compiere un salto paradigmatico fondamentale, superando i limiti dei modelli di affidabilità basati sulla popolazione. L'avvento di flussi di dati continui, specifici per ogni singolo asset, ha permesso di spostare il focus dall'analisi del "quando" un componente si guasta in media, all'analisi del "come" un componente specifico si sta degradando nel tempo. Questo ha dato vita a una famiglia di approcci prognostici fondati sull'analisi delle serie temporali, una branca della statistica che si occupa di analizzare sequenze di dati ordinati nel tempo per estrarne pattern significativi, comprenderne la struttura sottostante e, in ultima analisi, prevederne l'evoluzione futura.

La filosofia di base di questi approcci è tanto intuitiva quanto potente: si postula che il complesso processo di degrado fisico di un componente, pur essendo governato da leggi fisiche intricate, lasci una "impronta" osservabile e quantificabile nell'evoluzione dei segnali misurati dai sensori. Il problema prognostico viene così riformulato come un problema

di previsione di serie temporali (*time series forecasting*). Il flusso di lavoro tipico è un processo a due stadi. Nel primo stadio, noto come estrazione delle feature (*feature extraction*), i dati grezzi dei sensori, spesso ad alta frequenza e rumorosi (come i segnali di vibrazione o le emissioni acustiche), vengono processati per calcolare uno o più indicatori di salute (*Health Indicators* - HI). Un HI è una grandezza sintetica, una *feature*, progettata per avere due proprietà desiderabili: primo, essere sensibile al processo di degrado e, secondo, mostrare un andamento il più possibile monotono (costantemente crescente o decrescente) man mano che il degrado avanza. La scelta e la progettazione di buoni indicatori di salute è un'arte ingegneristica di per sé, e esempi classici includono il valore RMS (*Root Mean Square*) di un segnale di vibrazione, la sua kurtosi (che misura la "puntutezza" della sua distribuzione, sensibile agli impatti), la skewness o altri momenti statistici.
Nel secondo stadio, la sequenza di valori dell'HI nel tempo viene trattata come una serie storica e un modello statistico viene costruito per catturarne la dinamica e prevederne i valori futuri. La stima della RUL viene quindi calcolata come il tempo necessario affinché la traiettoria estrapolata dell'HI intersechi una soglia di fallimento predefinita e ingegneristicamente determinata.
All'interno del vasto e variegato arsenale di tecniche per l'analisi delle serie storiche, la famiglia di modelli che ha raggiunto uno status quasi canonico, grazie alla sua solidità teorica, alla sua flessibilità e, soprattutto, alla sua metodologia di applicazione strutturata, è quella dei modelli ARIMA (*Auto-Regressive Integrated Moving Average*). La sua formalizzazione nel lavoro seminale e monumentale di George Box e Gwilym Jenkins nel 1970 non ha semplicemente introdotto un nuovo modello, ma ha definito un'intera filosofia di modellazione iterativa, nota come metodologia Box-Jenkins, che ha dominato il campo per decenni [12].
Un modello ARIMA non è un'entità monolitica, ma una sintesi flessibile di tre componenti che mirano a catturare aspetti differenti e complementari della dinamica di una serie temporale. Una comprensione approfondita di ciascuna componente è essenziale:

- La Componente Auto-Regressiva (AR(p)): Questa componente si fonda sull'idea intuitiva che il futuro di una serie dipenda dal suo passato recente. Un modello AR(p) esprime il valore attuale della serie come una regressione lineare sui suoi *p* valori passati. In sostanza, la componente AR cattura l'autocorrelazione della serie, la sua "memoria" o la sua inerzia. Per un indicatore di degrado, questo riflette la natura cumulativa del danno: lo stato di degrado attuale è intrinsecamente legato a quello

immediatamente precedente. Il parametro *p* definisce l'ordine, ovvero la profondità di questa memoria.

- La Componente a Media Mobile (MA(*q*)): Questa componente, concettualmente più sottile, non modella la dipendenza dai valori passati della serie, ma piuttosto la dipendenza dagli errori di previsione passati. Questa componente è particolarmente efficace nel modellare eventi o "shock" casuali e non prevedibili il cui effetto non si esaurisce istantaneamente ma si propaga nel tempo, influenzando le osservazioni successive. In un contesto di degrado, può rappresentare l'effetto transitorio di sovraccarichi operativi, fluttuazioni di temperatura o altri disturbi esterni.

- La Componente Integrata (I(*d*)): Questa componente è, in realtà, un'operazione di pre-elaborazione fondamentale, che affronta un requisito teorico cruciale della metodologia Box-Jenkins: la stazionarietà. Una serie temporale è detta stazionaria (in senso debole) se le sue proprietà statistiche, in particolare la sua media e la sua varianza, sono costanti nel tempo. I modelli ARMA sono definiti per serie stazionarie. Tuttavia, le serie temporali degli indicatori di degrado sono, per loro stessa natura, quasi sempre non stazionarie, poiché esibiscono un trend sistematico crescente o decrescente che riflette l'accumulo di danno. Questa operazione, simile al calcolo di una derivata discreta, ha lo scopo di rimuovere il trend. Se il trend è lineare, una singola differenziazione ($d$=1) è sufficiente. Se il trend è quadratico, potrebbero essere necessarie due differenziazioni ($d$=2). La serie differenziata, ora stazionaria, può essere quindi modellata con una struttura ARMA. Il termine "integrato" deriva dal fatto che, per tornare a fare previsioni sulla scala originale, l'output del modello ARMA deve essere "integrato" (sommato) *d* volte.

Il processo di costruzione di un modello ARIMA, o metodologia Box-Jenkins, è un ciclo iterativo di identificazione, stima e verifica diagnostica. L'identificazione degli ordini *p* e *q* è una fase critica che si basa sull'analisi visuale dei grafici della Funzione di Autocorrelazione (ACF) e della Funzione di Autocorrelazione Parziale (PACF) della serie resa stazionaria. La stima dei coefficienti del modello viene poi eseguita tramite metodi statistici rigorosi come la stima di massima verosimiglianza. Infine, la verifica diagnostica consiste nell'analizzare i residui del modello: se il modello ha catturato tutta la struttura presente nei dati, i residui

dovrebbero essere indistinguibili da un rumore bianco. In caso contrario, il ciclo ricomincia con un nuovo modello.

Il vantaggio innegabile e rivoluzionario dei modelli di serie storiche rispetto all'analisi di affidabilità è la loro capacità di fornire una prognosi individualizzata e dinamica. La previsione non si basa più su medie di popolazione, ma sui dati specifici del singolo componente, e può essere aggiornata continuamente man mano che nuove misurazioni diventano disponibili. La loro natura di "scatola grigia", fondata su una teoria statistica trasparente e ben compresa, li rende inoltre più interpretabili di molti approcci di machine learning "black-box".

Nel corso degli anni, il framework ARIMA è stato esteso per gestire dinamiche più complesse. I modelli SARIMA (Seasonal ARIMA) includono componenti aggiuntive per modellare la stagionalità, un fenomeno comune in molti dati industriali. I modelli ARIMAX permettono di includere l'effetto di variabili esogene (o predittori esterni) nel modello.

Nonostante questa flessibilità, i modelli ARIMA soffrono di limitazioni strutturali profonde, che ne circoscrivono l'efficacia in molti problemi di prognostica del mondo reale. La loro debolezza più fondamentale è che sono intrinsecamente e per costruzione modelli lineari. Essi assumono che la relazione tra i valori passati e futuri della serie, così come l'effetto delle eventuali variabili esogene, sia di natura lineare. Questa assunzione è spesso una forzatura della realtà. I processi di degrado meccanico sono notoriamente non lineari: possono esibire fasi di stabilità (plateau), seguite da accelerazioni esponenziali del degrado all'avvicinarsi del guasto (ad esempio, nella propagazione di cricche sotto fatica). I modelli ARIMA, per loro stessa costruzione matematica, faticano a catturare queste transizioni di fase e queste dinamiche complesse, portando a previsioni che possono essere sistematicamente errate, specialmente nelle fasi più critiche del degrado [66].

In secondo luogo, i modelli ARIMA classici sono univariati: modellano l'evoluzione di un singolo indicatore di salute alla volta. Questo approccio, sebbene semplice da implementare, è sub-ottimale in quanto ignora le preziose informazioni contenute nelle interazioni e nelle correlazioni incrociate tra più flussi di dati provenienti da sensori diversi. In un sistema meccanico complesso, il degrado raramente si manifesta in un singolo segnale. Più spesso, è un fenomeno multidimensionale: un aumento delle vibrazioni è spesso accompagnato da un aumento della temperatura, da un cambiamento nello spettro acustico e da un aumento dell'assorbimento di corrente del motore. Un modello prognostico robusto dovrebbe essere in grado di fondere queste informazioni per ottenere una visione olistica dello stato di salute.

Sebbene esistano estensioni multivariate come i modelli VARIMA (*Vector Auto-Regressive Integrated Moving Average*), la loro complessità pratica e teorica cresce esponenzialmente con il numero di serie considerate. Il numero di parametri da stimare diventa rapidamente ingestibile (*curse of dimensionality*), e l'identificazione della struttura del modello diventa estremamente ardua [81].

Infine, essendo modelli basati sull'estrapolazione di pattern passati, la loro affidabilità e accuratezza tendono a diminuire drasticamente per orizzonti di previsione lunghi. Le previsioni ARIMA si fondano sull'assunzione critica che la struttura statistica del processo osservato nel passato rimanga invariata nel futuro. Qualsiasi cambiamento non previsto nelle condizioni operative o nella dinamica del degrado (un concept drift) può invalidare completamente il modello e portare a previsioni grossolanamente errate. Questo si riflette matematicamente nel rapido allargamento degli intervalli di confidenza delle previsioni, che per orizzonti temporali estesi possono diventare così ampi da non avere più alcun valore informativo pratico per una pianificazione della manutenzione a medio-lungo termine.

In conclusione, i modelli di analisi di serie storiche come ARIMA rappresentano un passo evolutivo fondamentale, introducendo il concetto di prognosi individualizzata basata sui dati di condizione. Tuttavia, le loro assunzioni strutturali di linearità e la loro natura prevalentemente univariata li rendono inadeguati a catturare appieno la complessità multidimensionale e non lineare dei processi di degrado reali.

### 3.2.3. Approcci basati su Processi Stocastici

Una terza famiglia di approcci statistici tradizionali affronta il problema della prognostica da una prospettiva differente e concettualmente molto elegante. Invece di tentare di adattare una funzione continua alla traiettoria di un indicatore di salute, come fanno i modelli di serie storiche, questi approcci operano una discretizzazione dello spazio degli stati di degrado. Il degrado non è più visto come un processo continuo, ma come una sequenza di transizioni probabilistiche tra un numero finito di stati di salute qualitativi. Questi sono i modelli basati su processi stocastici, e il loro archetipo, nonché il punto di partenza per una ricca famiglia di sviluppi più complessi, è la Catena di Markov (*Markov Chain* - MC).

La filosofia di base di questi modelli è quella di astrarre il complesso processo fisico di degrado, riducendolo a un sistema più semplice e matematicamente trattabile. Si postula che

un componente, in ogni istante di tempo discreto, possa trovarsi in uno e uno solo tra $N$ stati di salute predefiniti e mutualmente esclusivi. Questi stati possono essere definiti in modo qualitativo (ad esempio, per un cuscinetto: Stato 1: "Sano", Stato 2: "Pitting iniziale", Stato 3: "Spalling moderato", Stato 4: "Spalling esteso", Stato 5: "Guasto imminente") o in modo quantitativo, suddividendo il range di un indicatore di salute in intervalli discreti. Il comportamento del sistema, ovvero il suo "viaggio" attraverso questi stati, è interamente governato da un insieme di probabilità di transizione, che catturano la dinamica intrinseca del processo di degrado.

Una Catena di Markov a tempo discreto e a stati finiti è un processo stocastico definito da due elementi fondamentali: un insieme finito di stati $S$ e una matrice di probabilità di transizione. La Matrice di Probabilità di Transizione (*Transition Probability Matrix* - TPM), indicata con $P$, è una matrice quadrata di dimensione $N \ x \ N$ in cui ogni elemento $P$ rappresenta la probabilità che il sistema, trovandosi attualmente nello stato *s(i)*, transiti allo stato *s(j)* nel prossimo passo temporale. Ogni riga della matrice rappresenta una distribuzione di probabilità, e quindi la somma dei suoi elementi deve essere uguale a 1. In un contesto di degrado, si assume tipicamente che il processo sia progressivo, ovvero che un componente non possa "ringiovanire". Questo si traduce in una matrice $P$ triangolare superiore o a banda, dove la probabilità di tornare a uno stato più sano è zero. Lo stato di "Guasto", *s(N)*, è tipicamente modellato come uno stato assorbente, uno stato da cui è impossibile uscire.

Il fondamento matematico che rende le Catene di Markov così trattabili è la fondamentale e potente proprietà di Markov, che postula l'assenza di memoria del processo. Formalmente, la probabilità di transizione verso uno stato futuro dipende unicamente dallo stato presente ed è condizionatamente indipendente dall'intera sequenza di stati che ha portato il sistema a trovarsi in quella condizione.

Nonostante la loro eleganza matematica, le Catene di Markov semplici si scontrano con un limite pratico insormontabile nella stragrande maggioranza delle applicazioni di condition monitoring. Esse si basano sull'assunzione critica che lo stato di salute del sistema sia direttamente e perfettamente osservabile in ogni istante di tempo. Questa è un'assunzione quasi sempre irrealistica. Nella realtà, non possiamo "vedere" direttamente lo stato interno di "usura lieve" di un ingranaggio; possiamo solo misurare le vibrazioni che esso produce, o analizzare le particelle metalliche nel suo olio lubrificante. Queste misurazioni sono manifestazioni esterne, rumorose, incomplete e indirette dello stato di salute latente e non osservabile del componente.

Per superare questa fondamentale limitazione e rendere i modelli stocastici applicabili a dati di sensori reali, la ricerca ha sviluppato una delle estensioni più potenti e influenti della statistica moderna: il Modello Nascosto di Markov (*Hidden Markov Model* - HMM). L'HMM, come il suo nome suggerisce in modo eloquente, introduce un livello di astrazione cruciale, un "velo" tra l'osservatore e il sistema. Esso postula che la sequenza degli stati di salute, che segue ancora una Catena di Markov, sia nascosta (hidden) e non direttamente accessibile. Ciò che possiamo osservare è una sequenza di misurazioni provenienti dai sensori, che sono una manifestazione probabilistica dello stato nascosto sottostante. L'HMM è quindi un modello doppiamente stocastico: non solo la transizione tra gli stati nascosti è un processo probabilistico, ma anche la relazione tra ogni stato nascosto e le osservazioni che esso genera è di natura probabilistica.
Un HMM è quindi definito da tre insiemi di parametri, convenzionalmente indicati con la notazione compatta $\lambda = (A, B, \pi)$:

1. La matrice delle probabilità di transizione tra gli stati (A): Analoga alla matrice *P* della Catena di Markov, governa la dinamica della sequenza di stati nascosti.

2. L'insieme delle distribuzioni di probabilità di emissione (B): Questo è il nuovo, cruciale ingrediente. Per ogni stato nascosto *i*, esiste una distribuzione di probabilità *b* che definisce la verosimiglianza di osservare un certo valore *o* dai sensori quando il sistema si trova in quello stato. Se le osservazioni sono discrete (es. "bassa", "media", "alta" vibrazione), le probabilità di emissione sono descritte da una matrice. Più comunemente, per dati di sensori continui, la probabilità di emissione di ogni stato è modellata da una funzione di densità di probabilità (PDF) parametrica, tipicamente una distribuzione Gaussiana o, per maggiore flessibilità, una mistura di Gaussiane (*Gaussian Mixture Model* - GMM).

3. Il vettore delle probabilità iniziali degli stati ($\pi$): Definisce la probabilità che il sistema inizi il suo percorso da ciascuno degli *N* stati nascosti.

La teoria degli HMM, la cui trattazione classica è il tutorial seminale di Rabiner [85], fornisce algoritmi efficienti basati sulla programmazione dinamica per risolvere tre problemi fondamentali che rendono il modello operativamente utile. Il Problema della Valutazione,

risolto dall'algoritmo Forward-Backward, calcola la probabilità di una data sequenza di osservazioni secondo il modello. Il Problema della Decodifica, risolto dall'algoritmo di Viterbi, trova la sequenza più probabile di stati nascosti che ha generato una data sequenza di osservazioni; questo corrisponde al problema della diagnosi in tempo reale. Infine, il Problema dell'Apprendimento, risolto dall'algoritmo iterativo di Baum-Welch, stima i parametri del modello ($A$, $B$ e $\pi$) che massimizzano la probabilità delle sequenze di dati di addestramento.

Per la prognostica, il flusso di lavoro prevede l'addestramento di un HMM su dati storici di *run-to-failure*. Una volta addestrato, il modello cattura la "firma" statistica di ogni stato di degrado. Data una nuova sequenza di osservazioni da un componente in servizio, si può usare l'algoritmo di Viterbi per inferire il suo stato di salute nascosto più probabile. A questo punto, si possono usare le probabilità di transizione della matrice $A$ per calcolare la probabilità di raggiungere lo stato di guasto in un certo numero di passi futuri, fornendo così una stima probabilistica della RUL [11] [26].

Il punto di forza principale che rende i modelli stocastici, e in particolare gli HMM, così attraenti risiede nella loro solida e coerente base probabilistica. Essi forniscono un framework nativo ed elegante per gestire l'incertezza, sia quella intrinseca al processo di degrado (modellata dalle transizioni stocastiche) sia quella legata alla misurazione dei dati (modellata dalle emissioni probabilistiche). La loro capacità di distinguere concettualmente tra lo stato di salute interno e latente del sistema e le sue manifestazioni esterne e rumorose è un notevole passo avanti in termini di realismo modellistico rispetto agli approcci che lavorano direttamente sui dati osservati.

La flessibilità del framework ha permesso lo sviluppo di numerose estensioni per superarne i limiti. I Modelli Semi-Nascosti di Markov (HSMM) affrontano direttamente il limite della proprietà di Markov. In un HMM classico, la durata di permanenza in uno stato segue implicitamente una distribuzione geometrica, il che è spesso irrealistico. In un HSMM, la durata di permanenza in ogni stato è modellata da una distribuzione di probabilità esplicita e arbitraria, permettendo di catturare in modo più fedele la fisica del degrado. Gli HMM Auto-regressivi e gli HMM Accoppiati (Coupled HMMs) sono stati proposti per modellare le interdipendenze tra più serie temporali, estendendo il framework a contesti multivariati.

Nonostante la loro eleganza e potenza, anche questi modelli presentano debolezze intrinseche che ne limitano l'applicabilità. La proprietà di Markov, anche nelle sue forme estese, rimane un'assunzione forte. Molti processi di degrado fisico possiedono una "memoria" a lungo termine che non è facilmente catturabile da modelli basati su stati. Una

sfida significativa e spesso sottovalutata è la definizione a priori della topologia del modello, in particolare il numero di stati discreti. Questa scelta, spesso compiuta in modo euristico o basata sull'esperienza, ha un impatto profondo sulla struttura e sulle performance del modello, ma non esiste una procedura univoca e oggettiva per determinarla in modo ottimale. Un numero troppo basso di stati potrebbe non riuscire a catturare le diverse fasi del degrado, mentre un numero troppo alto potrebbe portare a problemi di overfitting e a una stima dei parametri instabile. Infine, sebbene gli algoritmi di addestramento siano ben definiti, possono essere computazionalmente intensivi per sequenze di dati molto lunghe e sono suscettibili di convergere a ottimi locali piuttosto che globali, rendendo i risultati sensibili alle condizioni di inizializzazione dei parametri.

In conclusione, i modelli basati su processi stocastici offrono un framework probabilistico potente, flessibile e concettualmente ricco per la diagnosi e la prognosi. La loro capacità di modellare stati latenti e di gestire l'incertezza li rende un passo avanti significativo rispetto ai modelli di serie storiche più semplici. Tuttavia, le loro assunzioni intrinseche, in particolare la proprietà di Markov e la necessità di discretizzare lo spazio degli stati, rappresentano dei vincoli che hanno motivato la ricerca di modelli ancora più flessibili e agnostici, capaci di apprendere direttamente da dati continui e ad alta dimensionalità senza imporre una struttura a stati predefinita. Questa ricerca di maggiore flessibilità è una delle principali forze che spingono verso l'adozione delle tecniche di Machine Learning.

### 3.2.4. Sintesi Comparativa degli Approcci Tradizionali

Dopo aver analizzato in dettaglio le tre principali famiglie di approcci tradizionali, è utile riassumerne e confrontarne le caratteristiche in modo schematico. La seguente tabella (*Tabella 2*) è stata creata proprio con questo scopo: fornire una sintesi comparativa immediata per mettere a fuoco i trade-off associati a ciascuna metodologia.

La tabella confronta i Modelli di Affidabilità, i Modelli di Serie Storiche e i Processi Stocastici lungo una serie di dimensioni chiave: dalla filosofia di base che li anima e il tipo di dati di cui hanno bisogno, fino ai loro principali punti di forza e ai limiti che li caratterizzano. Questa visione d'insieme non solo aiuta a consolidare quanto discusso finora, ma evidenzia anche un punto fondamentale.

Come si può notare, sebbene ciascuna di queste famiglie offra strumenti potenti per problemi specifici, esse sono tutte accomunate da una serie di assunzioni e limitazioni strutturali.

| **Dimensione di Confronto** | **Modelli di Affidabilità (es. Weibull, Cox)** | **Modelli di Serie Storiche (es. ARIMA)** | **Processi Stocastici (es. HMM)** |
|---|---|---|---|
| **Filosofia di Base** | Analisi statistica di una popolazione di componenti per modellare la probabilità di guasto nel tempo. | Estrapolazione del trend futuro di un singolo indicatore di salute (HI) di un asset individuale. | Modellazione del degrado come una sequenza di transizioni probabilistiche tra stati di salute discreti e latenti. |
| **Tipo di Dati Richiesti** | Dati di evento (tempi al guasto), anche in piccole quantità (*small data*) e con dati "censurati". | Serie temporale di un indicatore di salute (HI) campionato a intervalli regolari. | Sequenze di osservazioni da sensori, tipicamente associate a cicli di vita completi (*run-to-failure*). |
| **Obiettivo Primario** | Pianificazione strategica della manutenzione (es. intervalli di sostituzione), gestione delle scorte, definizione di garanzie. | Prognosi a breve-medio termine della RUL per un singolo asset. | Diagnosi dello stato di salute attuale (decodifica) e stima probabilistica della RUL. |
| **Punti di Forza Principali** | - Interpretabilità (il parametro β di Weibull ha un chiaro significato fisico).<br>- Efficace con pochi dati.<br>- Solida base teorica. | - Prognosi individualizzata e dinamica.<br>- Relativamente semplice da implementare.<br>- Buona interpretabilità ("scatola grigia"). | - Gestione nativa dell'incertezza.<br>- Distinzione tra stato latente e osservazioni rumorose.<br>- Framework probabilistico coerente. |
| **Limiti Fondamentali** | - Cecità alla condizione individuale (modello di popolazione, non di individuo).<br>- Non utilizza dati di condition monitoring. | - Assunzione di linearità e stazionarietà (spesso irrealistica).<br>- Prevalentemente univariato.<br>- Inaffidabile su orizzonti lunghi. | - Assunzione di assenza di memoria (proprietà di Markov).<br>- Necessità di discretizzare gli stati (scelta soggettiva).<br>- Complessità computazionale. |
| **Esempio di Applicazione Pratica** | Determinare l'intervallo ottimale per sostituire una flotta di 100 pompe identiche in un impianto. | Prevedere l'evoluzione dell'usura di un singolo utensile da taglio nel prossimo turno di lavoro. | Identificare lo stadio di degrado attuale (es. "sano", "pitting", "spalling") di un cuscinetto e stimare il tempo al guasto. |

*Tabella 2 - Sintesi Comparativa degli Approcci Tradizionali*

## 3.3. Limiti degli Approcci Tradizionali e Problemi Aperti

Come abbiamo visto, gli approcci statistici tradizionali — dall'ingegneria dell'affidabilità ai modelli di serie storiche e ai processi stocastici — hanno costituito per decenni la base della prognostica quantitativa. Hanno segnato un passaggio fondamentale da una manutenzione basata sull'intuizione a una gestione fondata su principi scientifici, fornendo gli strumenti per ragionare in modo rigoroso su probabilità, rischio e previsione.

Tuttavia, con l'aumentare della complessità dei sistemi industriali e la crescente disponibilità di dati, i limiti di questi modelli sono diventati sempre più evidenti. Questa sezione si propone di analizzare in modo critico queste debolezze, che non sono semplici difetti tecnici, ma veri e propri problemi aperti che i modelli tradizionali faticano a risolvere. Essi sono accomunati da una serie di assunzioni e limitazioni strutturali che, pur essendo state utili in passato, oggi ne limitano il potenziale. Sono, in un certo senso, strumenti molto raffinati ma progettati per un mondo di dati scarsi (*small data*) e di relazioni relativamente semplici [103].

È proprio analizzando queste sfide che si comprende la necessità di un nuovo paradigma. La domanda a cui la ricerca moderna cerca di rispondere non è più solo "come possiamo migliorare l'accuratezza del nostro modello ARIMA?", ma piuttosto "esiste una nuova classe di modelli, basata su principi diversi, capace di superare i limiti del passato?". Questo nuovo paradigma deve essere in grado di gestire la non linearità, di lavorare con dati provenienti da molti sensori e di sfruttare appieno le enormi quantità di dati generate dall'Industria 4.0 [61].

### 3.3.1. Il Dogma della Linearità e della Stazionarietà

Una delle limitazioni più importanti degli approcci statistici classici, in particolare dei modelli di serie storiche come ARIMA, è la loro assunzione di linearità. Questi modelli, per come sono costruiti, ipotizzano che la relazione tra il valore attuale di un indicatore di salute e la sua storia passata possa essere rappresentata da una funzione lineare. Sebbene questa ipotesi renda i modelli facili da interpretare e matematicamente gestibili, spesso non riflette la realtà dei processi di degrado meccanico, che sono quasi sempre non lineari.

I meccanismi di guasto che governano la vita dei componenti meccanici sono, nella stragrande maggioranza dei casi, fenomeni non lineari. Consideriamo, ad esempio, un processo di fatica dei materiali, una delle principali cause di guasto nei sistemi soggetti a carichi ciclici. La teoria della meccanica della frattura ci insegna che questo processo non segue affatto una progressione lineare. Esso si articola tipicamente in tre fasi distinte, con dinamiche diverse. La prima è una fase di nucleazione della cricca, che può costituire la maggior parte della vita utile del componente. Durante questa lunga fase, il danno si accumula a livello microstrutturale, ma i segnali esterni misurabili (come le vibrazioni o le emissioni acustiche) possono rimanere quasi del tutto invariati, mostrando solo un rumore di fondo stazionario. Segue una fase di propagazione stabile, in cui la cricca, una volta formatasi, cresce in modo quasi lineare ad ogni ciclo di carico, seguendo leggi come quella di Paris-Erdogan. In questa fase, un indicatore di salute ben progettato potrebbe mostrare un trend approssimativamente lineare. Infine, all'avvicinarsi della dimensione critica della cricca, quando la sezione resistente residua del componente si riduce drasticamente, il processo entra in una fase di frattura instabile, caratterizzata da una crescita esponenziale e catastrofica che porta al collasso del componente in un intervallo di tempo molto breve.

Un modello lineare come ARIMA, quando applicato a un indicatore di salute che riflette questo processo a tre fasi, si troverebbe in grave difficoltà. Esso tenterebbe di adattare un'unica funzione lineare a un fenomeno che è, per sua natura, una successione di regimi con dinamiche diverse. Il risultato sarebbe un modello che "media" il comportamento complessivo, fallendo nel catturare le cruciali transizioni di fase. In particolare, sottostimerebbe drasticamente l'accelerazione del degrado nelle fasi finali e più critiche, quelle in cui una previsione accurata è più necessaria. Questo si tradurrebbe in stime della RUL pericolosamente ottimistiche, che potrebbero portare a un guasto imprevisto, vanificando l'intero scopo del sistema prognostico [66]. Allo stesso modo, fenomeni come l'usura adesiva o abrasiva sono governati da complesse interazioni di superficie e spesso presentano comportamenti a soglia (transizione da usura lieve a severa) e relazioni non lineari con le condizioni operative come il carico, la velocità relativa e le proprietà del lubrificante. L'imposizione di un modello lineare su tali fenomeni non è un'approssimazione ragionevole, ma una potenziale fonte di errori grossolani e sistematici.

Legata a questa problematica è l'altrettanto rigida assunzione di stazionarietà. Come discusso, la teoria statistica alla base dei modelli ARMA richiede che la serie temporale analizzata sia stazionaria, ovvero che le sue proprietà statistiche, in particolare la sua media e la sua varianza, siano costanti nel tempo. Sebbene tecniche come la differenziazione siano

efficaci nel rimuovere trend semplici, esse sono molto meno adeguate a gestire cambiamenti più complessi e strutturali nella dinamica della serie, che sono la norma piuttosto che l'eccezione in contesti industriali reali.

La realtà di un impianto produttivo è dinamica e non stazionaria. Un macchinario su una linea di produzione non opera in condizioni di laboratorio controllate e costanti. Può essere chiamato a lavorare a regimi di velocità e carico differenti a seconda del pezzo da produrre (es. sgrossatura vs. finitura); può processare lotti di materie prime con caratteristiche meccaniche e di lavorabilità leggermente diverse; può essere soggetto a profili di missione che cambiano in base alla domanda di mercato; o può operare in un ambiente la cui temperatura e umidità variano nel corso della giornata e delle stagioni. Ognuno di questi cambiamenti nelle condizioni operative esterne, può alterare la dinamica stessa del processo di degrado. Un esempio pratico è quello di una fresatrice CNC utilizzata per lavorare sia alluminio (un materiale tenero) che acciaio inossidabile (un materiale duro). Un modello prognostico tradizionale, addestrato su dati raccolti durante la lavorazione dell'alluminio, apprenderebbe un certo tasso di usura dell'utensile. Se la produzione passa all'acciaio inossidabile, le forze di taglio e le temperature aumentano drasticamente. Il modello, basandosi sulla sua esperienza passata, continuerebbe a prevedere un degrado lento, portando a una stima della RUL dell'utensile grossolanamente ottimistica. L'utensile si romperebbe molto prima del previsto, potenzialmente danneggiando il pezzo e il mandrino. Questo fallimento non è dovuto a un errore del modello, ma alla sua incapacità strutturale di gestire un cambiamento di regime operativo, ovvero un cambiamento fondamentale nelle statistiche del processo. Questo fenomeno, noto come *concept drift*, rappresenta una delle sfide più grandi e insidiose per l'implementazione di sistemi prognostici online [32].

I modelli statistici tradizionali sono strutturalmente fragili di fronte a questi cambiamenti. Quando le condizioni operative cambiano, il modello addestrato sui dati passati, che assume implicitamente che il futuro assomiglierà statisticamente al passato, diventa progressivamente disallineato dalla nuova realtà del processo. Le sue previsioni iniziano a deviare e la sua accuratezza degrada, a volte in modo silenzioso e a volte in modo catastrofico. Essi mancano di un meccanismo intrinseco per adattarsi dinamicamente a un ambiente operativo che è non stazionario e mutevole.

In sintesi, il "dogma" della linearità e l'assunzione di stazionarietà, che hanno reso i modelli statistici tradizionali così eleganti, trattabili e comprensibili, si rivelano essere la loro più grande debolezza. Questa profonda discrepanza tra le assunzioni del modello e la realtà del processo fisico è una delle ragioni primarie che ha spinto la ricerca scientifica verso

paradigmi di modellazione intrinsecamente più flessibili, non parametrici e potenti, come quelli offerti dal Machine Learning, che sono nativamente progettati per apprendere e rappresentare funzioni non lineari arbitrarie e, nelle loro varianti online, per adattarsi a contesti non stazionari.

### 3.3.2. La "Maledizione della Dimensionalità" e la Visione a Tunnel

La seconda e forse ancora più invalidante limitazione che interessa le metodologie statistiche tradizionali è la loro difficoltà nel gestire efficacemente dati multidimensionali. La stragrande maggioranza degli approcci classici, dai modelli ARIMA alle Catene di Markov, è stata concepita, sviluppata e validata in un'epoca di scarsità di dati, un'era in cui il monitoraggio di un sistema complesso si limitava spesso alla misurazione periodica di un singolo, o al massimo di una manciata, di parametri chiave. Di conseguenza, queste tecniche sono, nella loro forma più comune, più utilizzata e più compresa, intrinsecamente univariate. Sono state progettate per modellare la dinamica di una singola serie temporale, ovvero di un singolo indicatore di salute, alla volta. Questo approccio, sebbene possa apparire metodologicamente pulito e semplice da implementare, impone una "visione a tunnel" sul problema prognostico, una prospettiva fondamentalmente riduzionista che, per sua stessa natura, è cieca a una delle fonti di informazione più potenti e sottili: la sinergia informativa che emerge dalle interazioni, dalle dipendenze e dalle complesse correlazioni incrociate tra i dati provenienti da molteplici e diversi sensori.

In un sistema industriale moderno, un singolo macchinario critico non è un'entità semplice e isolata, ma un organismo complesso, un sistema ciber-fisico il cui stato di salute è costantemente monitorato da una rete eterogenea di sensori. Questa rete agisce come un sistema nervoso artificiale, fornendo un flusso continuo di dati che, nel loro insieme, costituiscono un ritratto multidimensionale del suo stato operativo. Vibrazioni lungo tre assi cartesiani, temperature in punti strategici come i cuscinetti e l'avvolgimento del motore, pressioni del circuito idraulico, portate del fluido lubrificante, assorbimenti di corrente dei motori elettrici, emissioni acustiche ad alta frequenza e analisi spettrali non sono semplicemente variabili indipendenti; sono proiezioni diverse di un unico e complesso stato fisico sottostante. È quindi altamente probabile, se non scientificamente certo, che un processo di degrado incipiente non si manifesti come un'anomalia chiara e isolata in un

singolo segnale, ma piuttosto come un pattern distribuito, sottile e correlato che emerge dall'intero spazio multidimensionale dei dati.

Un buon esempio per capire questo limite è il processo di degrado di un ingranaggio. Un guasto iniziale, come una piccola cricca su un dente, si manifesta attraverso molti segnali contemporaneamente. Potremmo notare un leggero aumento delle vibrazioni in alcune frequenze specifiche, un piccolo ma costante aumento della temperatura dell'olio a causa del maggiore attrito, e forse anche una variazione quasi impercettibile nel consumo di corrente del motore. Allo stesso tempo, un'analisi dell'olio potrebbe rivelare più particelle metalliche, e il rumore prodotto potrebbe arricchirsi di nuovi suoni impulsivi.

Se analizzassimo ciascuno di questi segnali da solo, con un modello separato per ognuno, avremmo una visione frammentata del problema. Potremmo perdere l'informazione più importante, perché l'indizio più forte del guasto non è in una singola osservazione, ma nel fatto che tutti questi piccoli segnali stiano accadendo insieme e stiano evolvendo in modo correlato. La capacità di fondere queste informazioni provenienti da più sensori, nota come multi-sensor data fusion, è oggi considerata una delle chiavi per ottenere sistemi di diagnosi e prognosi veramente robusti [63].

Sebbene la teoria statistica abbia, nel corso degli anni, sviluppato estensioni multivariate dei modelli tradizionali, come i modelli VARIMA (Vector Auto-Regressive Integrated Moving Average), la loro applicazione pratica si scontra rapidamente e violentemente con un ostacolo fondamentale e quasi insormontabile della statistica e del Machine Learning, un fenomeno noto con l'espressione di "maledizione della dimensionalità" (*curse of dimensionality*). Questo termine, coniato da Richard Bellman negli anni '60 nel contesto della programmazione dinamica, descrive una serie di fenomeni paradossali che emergono quando si lavora in spazi ad alta dimensionalità [8].

All'aumentare del numero di variabili, il volume dello spazio dei dati cresce esponenzialmente, rendendo i dati sempre più "sparsi" e le nozioni statistiche di vicinanza meno significative.

Più concretamente, il numero di parametri da stimare esplode. Consideriamo un esempio pratico: monitorare un sistema con soli 10 sensori (un numero molto modesto per un macchinario moderno) utilizzando un modello VARIMA. Se volessimo modellare le interazioni tra questi 10 segnali, anche con un modello molto semplice che consideri solo il ritardo temporale precedente, dovremmo stimare una matrice di coefficienti di 10x10 (100 parametri) e una matrice di covarianza di 10x10 (altri 55 parametri unici). Con soli 10 sensori, il modello avrebbe già 155 parametri da apprendere. Se i sensori fossero 50, il

numero di parametri supererebbe le migliaia. Questa crescita esplosiva ha due conseguenze nefaste. La prima è l'instabilità statistica: con un numero limitato di dati, le stime di così tanti parametri diventano estremamente incerte. La seconda, e ancora più grave, è l'aumento quasi inevitabile del rischio di *overfitting*. Il modello, avendo troppi gradi di libertà, cesserà di essere un modello del processo fisico per diventare un modello del rumore specifico del dataset di addestramento. Finirà per "imparare a memoria" le correlazioni spurie, mostrando performance eccellenti su quei dati ma perdendo ogni capacità di generalizzare su dati nuovi. Il modello tradisce la sua stessa vocazione predittiva per diventare un mero esercizio di interpolazione iper-parametrica.

La complessità dell'identificazione della struttura del modello, della stima dei suoi innumerevoli parametri e della sua rigorosa validazione diagnostica ha relegato i modelli statistici multivariati classici al ruolo di curiosità accademiche piuttosto che a quello di strumenti pratici e scalabili per l'industria. Essi sono, in sostanza, strumenti nati e ottimizzati per un mondo di dati "bassi e larghi" (poche variabili, moltissime osservazioni), un mondo che sta rapidamente scomparendo. L'era dell'Industria 4.0 ci presenta sempre più spesso problemi "alti e stretti" (moltissime variabili/sensori, spesso con un numero limitato di cicli di vita completi da cui imparare).

In conclusione, l'approccio prevalentemente univariato dei modelli statistici tradizionali costituisce una limitazione concettuale profonda e insuperabile. Li rende incapaci, per costruzione, di sfruttare la ricchezza informativa che deriva dalla fusione di dati multisensoriali, costringendoli a una visione parziale, frammentata e sub-ottimale del complesso processo di degrado. La “maledizione della dimensionalità”, d'altra parte, impedisce alle loro estensioni multivariate di essere una soluzione pratica, robusta e scalabile a questo problema.

### 3.3.3. La Servitù dell'Estrazione Manuale delle Feature

Oltre ai limiti legati alle assunzioni matematiche, c'è una terza sfida, di natura più pratica, che riguarda quasi tutti gli approcci tradizionali e anche gran parte del Machine Learning "classico": la loro forte dipendenza da un processo noto come estrazione delle feature (*feature engineering*).

Questi modelli, infatti, raramente lavorano direttamente sui dati grezzi acquisiti dai sensori. I dati grezzi, come i segnali di vibrazione, sono spesso molto rumorosi e contengono un'enorme quantità di informazioni, la maggior parte delle quali irrilevante. L'informazione utile sul degrado è presente, ma è come "annegata" in questo mare di dati. Per poter essere utilizzati da un modello, questi dati grezzi devono essere prima trasformati in un insieme di indicatori di salute (*Health Indicators* - HI), o *feature*. Una feature è, in sostanza, un valore sintetico calcolato a partire dai dati grezzi, progettato per essere informativo e sensibile al processo di degrado.
Questo processo di feature engineering non è un semplice passaggio tecnico di pre-elaborazione. È, a tutti gli effetti, un'attività quasi artigianale che richiede una profonda conoscenza del sistema e delle tecniche di elaborazione dei segnali. L'ingegnere deve progettare e calcolare manualmente un insieme di feature che possano, si spera, catturare l'essenza del processo di degrado. Questo processo può attingere da diverse "scuole di pensiero" analitiche, spesso usate in combinazione:

- Analisi nel dominio del tempo: È l'approccio più diretto, che calcola momenti statistici e descrittori della forma d'onda del segnale. Include metriche come la media, la deviazione standard, il valore RMS (Root Mean Square, che è legato all'energia del segnale), la skewness (che misura l'asimmetria della distribuzione e può indicare la presenza di impatti unidirezionali) e la kurtosi.

- Analisi nel dominio della frequenza: Questo approccio si basa sull'applicazione della Trasformata di Fourier (FFT) per convertire il segnale dal dominio del tempo al dominio della frequenza, rivelando le sue componenti spettrali. Le feature estratte possono essere l'ampiezza di specifiche componenti di frequenza, come le frequenze caratteristiche di guasto dei cuscinetti (BPFI, BPFO, BSF, FTF) o le frequenze di ingranamento e le loro bande laterali (*sidebands*), che sono indicative di modulazioni causate da un danno. L'energia contenuta in determinate bande di frequenza è un'altra feature comune e potente [89].

- Analisi nel dominio tempo-frequenza: Riconoscendo che i segnali di degrado sono spesso non stazionari (il loro contenuto spettrale cambia nel tempo), queste tecniche avanzate analizzano simultaneamente il segnale in entrambi i domini. La Trasformata di Fourier a tempo breve (*Short-Time Fourier Transform* -

STFT) fornisce una prima approssimazione, ma soffre di un trade-off fisso tra risoluzione temporale e spettrale (il principio di indeterminazione di Gabor). La Trasformata Wavelet (*Wavelet Transform* - WT) supera questo limite utilizzando "ondine" di analisi di durata variabile, fornendo un'eccellente risoluzione temporale per le alte frequenze (ideale per rilevare eventi transitori e impulsivi) e un'eccellente risoluzione in frequenza per le basse frequenze. L'energia dei coefficienti wavelet a diverse scale di scomposizione è diventata una delle fonti più ricche per l'estrazione di feature prognostiche [82].

Sebbene il feature engineering sia stato per decenni, e per certi versi lo sia ancora, un passaggio indispensabile e cruciale per il successo di qualsiasi sistema di diagnosi e prognosi, esso presenta notevoli e profondi svantaggi che ne fanno un vero e proprio collo di bottiglia.

In primo luogo, è un processo altamente soggettivo e dipendente dall'esperienza. La qualità e la pertinenza delle feature estratte, e di conseguenza la performance finale dell'intero sistema prognostico, non dipendono tanto dalla potenza dell'algoritmo statistico a valle, quanto dall'abilità, dall'intuizione e dalla conoscenza a priori dell'ingegnere che le progetta. La scelta delle feature più informative è spesso un processo lungo, basato su tentativi ed errori, che richiede una rara, costosa e difficilmente trasferibile combinazione di expertise multidisciplinare. Questo introduce un elemento di "arte" non quantificabile in un processo che dovrebbe tendere alla massima scientificità e oggettività, rendendo i risultati difficilmente generalizzabili e la replicabilità degli studi una sfida costante.

In secondo luogo, le feature progettate manualmente sono, per loro natura, specifiche per un dato problema e spesso fragili. Le feature che funzionano brillantemente per diagnosticare un guasto allo statore di un motore elettrico potrebbero rivelarsi completamente inutili per prevedere l'usura di un utensile in una macchina di fresatura. Ogni nuovo macchinario, ogni nuovo tipo di guasto, ogni nuova condizione operativa potrebbe richiedere un nuovo e laborioso ciclo di progettazione, implementazione e validazione delle feature. Questo limita drasticamente la trasferibilità e la scalabilità dell'approccio, rendendo proibitivo lo sviluppo di soluzioni prognostiche "chiavi in mano" per flotte di macchinari eterogenei o per sistemi che operano in condizioni mutevoli.

In terzo luogo, e questo è forse il limite scientifico ed epistemologico più profondo, il feature engineering manuale impone una visione a priori, guidata dall'uomo, su quali siano le informazioni rilevanti contenute nei dati. L'ingegnere, basandosi sulla sua conoscenza e sugli

strumenti matematici a sua disposizione, seleziona un insieme finito di trasformazioni e calcoli, proiettando i dati grezzi ad alta dimensionalità in uno spazio a bassa dimensionalità di sua creazione. Questo processo, per quanto informato e ben intenzionato, comporta intrinsecamente il rischio fondamentale di scartare informazioni preziose o, peggio, di non riuscire a identificare e formulare matematicamente i pattern più predittivi. Questi pattern potrebbero essere rappresentazioni molto complesse, non lineari, multidimensionali e non intuitive dei dati grezzi, che sfuggono alla nostra comprensione e ai nostri strumenti analitici tradizionali. In sostanza, si corre il rischio che l'intelligenza umana, invece di essere un volano, diventi un filtro, un collo di bottiglia che impedisce all'algoritmo di apprendimento di scoprire le relazioni più sottili e nascoste che governano il processo di degrado.

Questo paradigma, in cui una parte critica dell'"intelligenza" del sistema è delegata all'ingegnere, rappresenta un ostacolo sia operativo che scientifico. Da questa consapevolezza nasce la necessità di approcci capaci di apprendere le feature più informative direttamente dai dati grezzi, in modo automatico e ottimizzato per il compito, una delle principali promesse del Deep Learning.

### 3.3.4. Le Gabbie Concettuali dei Modelli Stocastici

I modelli basati su processi stocastici, e in particolare i Modelli Nascosti di Markov (HMM), rappresentano un notevole passo avanti rispetto ai modelli di serie storiche. La loro capacità di distinguere tra uno stato di salute interno (che non possiamo vedere) e i segnali esterni che possiamo misurare, e di gestire l'incertezza in modo probabilistico, li rende strumenti molto potenti. Tuttavia, anche questi modelli si basano su alcune assunzioni fondamentali che, pur rendendoli matematicamente gestibili, ne limitano l'applicazione pratica di fronte a processi di degrado complessi. Possiamo pensare a queste assunzioni come a delle "gabbie concettuali".

La prima e più fondamentale di queste assunzioni, il vero e proprio pilastro su cui si regge l'intero edificio delle Catene di Markov e dei loro derivati, è la proprietà di Markov. Questa proprietà, nota anche come "assenza di memoria", postula che la probabilità di transizione a uno stato futuro sia completamente determinata dallo stato presente, e sia condizionatamente indipendente dall'intera sequenza di stati che ha portato il sistema a trovarsi in quella condizione. Matematicamente, questo si traduce nell'assunzione che il futuro sia

stocasticamente indipendente dal passato, dato il presente. Questa ipotesi è una drastica semplificazione che permette di ridurre la dinamica temporale del processo a una semplice e potente algebra delle matrici, ma è una semplificazione che, in molti contesti ingegneristici, si scontra frontalmente con la fisica dei meccanismi di degrado cumulativo.

Processi come la fatica dei materiali, l'usura abrasiva e adesiva, o la corrosione sotto sforzo sono fenomeni che possiedono una memoria intrinseca e non trascurabile. La velocità del degrado futuro in questi sistemi dipende non solo dallo stato di danno attuale (ad esempio, la lunghezza di una cricca o la profondità di un solco di usura), ma anche dall'intera storia di carico e di stress a cui il componente è stato sottoposto. Fenomeni microstrutturali come l'incrudimento (*strain hardening*), che aumenta la resistenza del materiale a seguito di deformazione plastica, o l'addolcimento ciclico (*cyclic softening*), che la riduce, modificano le proprietà meccaniche del componente nel tempo. Queste modifiche, a loro volta, influenzano la sua successiva risposta al degrado. Un modello markoviano, per sua stessa definizione, è "cieco" a questa storia cumulativa. Esso non può distinguere tra un componente che ha raggiunto un certo stato di degrado attraverso un processo lento e graduale a basso carico e un altro che ha raggiunto lo stesso stato apparente attraverso brevi ma intensi periodi di sovraccarico, che potrebbero aver indotto danni latenti più profondi. Questa mancanza di memoria può portare a una rappresentazione grossolanamente inaccurata della dinamica del degrado, specialmente in sistemi che operano sotto profili di carico variabili e complessi [26].

Per mitigare questa evidente limitazione, la ricerca ha sviluppato estensioni più sofisticate, come i Modelli Semi-Nascosti di Markov (*Hidden Semi-Markov Models* - HSMM). In un HMM classico, l'assunzione di Markov implica che la durata di permanenza in uno stato segua implicitamente una distribuzione geometrica, che ha una funzione di rischio costante. Questo è spesso irrealistico, poiché suggerisce che, una volta entrato in uno stato, il componente ha la stessa probabilità di uscirne in ogni istante successivo, indipendentemente da quanto tempo ci sia già rimasto. Un HSMM supera questo limite modellando esplicitamente la distribuzione della durata di permanenza in ogni stato con una funzione di probabilità arbitraria e più flessibile (es. Gamma, Lognormale, Weibull). Questo permette di catturare in modo più fedele la fisica del degrado, ad esempio modellando il fatto che più a lungo un componente rimane in uno stato di "usura lieve", più è probabile che transiti a uno di "usura moderata". Tuttavia, sebbene gli HSMM risolvano il problema della memoria legata alla durata, non risolvono completamente il problema della dipendenza dalla storia

del processo nel suo complesso, e lo fanno al costo di una maggiore complessità del modello e di algoritmi di inferenza significativamente più onerosi.

La seconda "gabbia concettuale", non meno restrittiva, dei modelli stocastici è la loro necessità di una discretizzazione a priori dello spazio degli stati di salute. Il degrado, nella stragrande maggioranza dei sistemi fisici, è un processo intrinsecamente continuo. Un indicatore di salute, come il livello RMS delle vibrazioni o la dimensione di una cricca, varia in modo continuo nel tempo. I modelli basati su processi stocastici, invece, impongono una visione del mondo in cui questo fenomeno continuo viene forzatamente proiettato, e quindi quantizzato, su un insieme finito e discreto di stati. La decisione su quanti stati discreti utilizzareper rappresentare il continuum del degrado è una scelta di modellazione critica, fondamentale e, purtroppo, spesso altamente soggettiva e priva di una solida giustificazione teorica.

Questa scelta non è un dettaglio tecnico, ma una decisione con profonde conseguenze. Un numero troppo basso di stati porterà a un modello eccessivamente grossolano, con una "risoluzione" insufficiente per catturare le diverse e sottili fasi del degrado. Ad esempio, raggruppare un'ampia gamma di comportamenti, dalla prima comparsa di un difetto fino a un danno significativo, in un unico stato di "usura moderata" potrebbe mascherare transizioni dinamiche importanti all'interno di quella fase, ritardando la diagnosi e rendendo la prognosi imprecisa. Al contrario, un numero troppo alto di stati può portare a una serie di problemi altrettanto gravi. In primo luogo, aumenta drasticamente il numero di parametri del modello da stimare (la dimensione della matrice di transizione cresce quadraticamente con il numero di stati), richiedendo dataset di addestramento enormi e spesso irrealistici per poter stimare in modo affidabile tutte le possibili probabilità di transizione. In secondo luogo, un numero eccessivo di stati aumenta il rischio di overfitting, in cui il modello si adatta eccessivamente alle peculiarità e al rumore specifico dei dati di addestramento, perdendo la capacità di generalizzare. Infine, un numero elevato di stati può rendere l'interpretazione fisica di ciascuno stato più difficile e ambigua, vanificando in parte il vantaggio di interpretabilità del modello.

Nonostante l'importanza cruciale di questa scelta di progettazione, non esiste una procedura univoca e oggettiva per determinare il numero ottimale di stati. Spesso, questa decisione viene presa in modo euristico, basandosi sull'esperienza del modellista, o attraverso procedure di confronto tra modelli basate su criteri informativi come il Criterio Informativo Bayesiano (BIC) o il Criterio Informativo di Akaike (AIC). Questi criteri cercano un compromesso tra la bontà di adattamento del modello ai dati e la sua complessità (il numero

di parametri), ma non garantiscono di trovare la rappresentazione che sia più fedele al processo fisico sottostante [16].
In conclusione, i modelli basati su processi stocastici, pur essendo molto potenti dal punto di vista probabilistico, si basano su assunzioni strutturali che ne limitano la flessibilità. La proprietà di Markov impone una visione "senza memoria", mentre la necessità di discretizzare gli stati impone una griglia artificiale su un fenomeno continuo. Queste limitazioni non li rendono obsoleti, ma li rendono più adatti a problemi specifici dove una rappresentazione a stati discreti ha senso. La ricerca di modelli capaci di lavorare con dati continui e di apprendere dipendenze temporali a lungo raggio è una delle principali ragioni che hanno spinto la comunità scientifica verso le metodologie del Machine Learning, in particolare verso le reti neurali ricorrenti, che sono state progettate proprio per superare queste "gabbie concettuali".

### 3.3.5. Verso il Machine Learning

Questi approcci, pur rappresentando una pietra miliare insostituibile nella storia della gestione scientifica della manutenzione e nell'evoluzione della prognostica quantitativa, si rivelano essere strumenti concettualmente e matematicamente inadeguati a fronteggiare la scala, la dimensionalità e la complessità dei sistemi industriali moderni e della rivoluzione dei dati che li accompagna. Questo paragrafo finale si propone quindi di sintetizzare queste sfide, consolidando l'argomentazione per cui la transizione verso il Machine Learning non è una mera scelta di convenienza o una rincorsa alla moda tecnologica del momento, ma una necessità scientifica e ingegneristica per il progresso della disciplina.
Il primo e forse più fondamentale limite che abbiamo sviscerato è l'incapacità intrinseca dei modelli tradizionali di catturare e rappresentare la complessa dinamica non lineare che governa la stragrande maggioranza dei processi di degrado fisico. L'imposizione di un'assunzione di linearità, come nei modelli ARIMA, o di una dinamica a stati discreti con transizioni a memoria breve, come nei modelli markoviani, si scontra frontalmente con la realtà di fenomeni come la fatica, l'usura e la corrosione, che esibiscono comportamenti complessi, transizioni di fase non lineari, accelerazioni esponenziali e dipendenze temporali a lungo raggio. Questa profonda discrepanza tra la semplicità ipotizzata dal modello e la complessità del fenomeno fisico non è un'inezia accademica, ma si traduce in una

rappresentazione inaccurata del degrado, che può portare a previsioni della RUL sistematicamente errate, inaffidabili e, nel peggiore dei casi, pericolosamente ottimistiche [66].

La seconda sfida cruciale, che diventa ogni giorno più pressante, è l'inadeguatezza strutturale di questi modelli nel gestire dati multidimensionali. L'approccio prevalentemente univariato degli strumenti classici li costringe a una visione frammentata e riduzionista del processo di degrado, una sorta di "cecità selettiva" che li porta a ignorare la ricchezza informativa che risiede nelle interazioni e nelle correlazioni tra molteplici flussi di dati provenienti da sensori eterogenei. Le loro estensioni multivariate, d'altro canto, si scontrano violentemente con la "maledizione della dimensionalità", un muro computazionale e statistico che ne rende l'applicazione pratica infattibile a causa dell'esplosione del numero di parametri e del conseguente, quasi inevitabile, rischio di overfitting. In un'era in cui i sistemi sono sempre più densamente sensorizzati e la fusione di informazioni multisensoriali (*multi-sensor data fusion*) è riconosciuta come la chiave per una diagnosi e una prognosi robuste, questa incapacità di operare nativamente ed efficacemente in spazi ad alta dimensionalità rappresenta un grave e insuperabile ostacolo [63].

Il terzo collo di bottiglia, di natura più operativa ma non meno limitante, è la servitù epistemologica nei confronti del feature engineering manuale. La necessità di affidarsi all'esperienza, all'intuizione e all'abilità artigianale di un ingegnere umano per progettare e selezionare a mano gli indicatori di salute dai dati grezzi introduce soggettività, limita la scalabilità e, soprattutto, preclude la possibilità di scoprire pattern predittivi complessi e non intuitivi che potrebbero essere nascosti nelle profondità dei dati. Questo processo non solo è laborioso e costoso, ma rappresenta un limite fondamentale alla capacità di scoperta automatica della conoscenza, che dovrebbe essere il cuore di un sistema di apprendimento moderno. Si corre il rischio che l'intelligenza umana, invece di potenziare, limiti il potenziale dell'analisi, imponendo i propri bias e la propria limitata comprensione su un fenomeno che potrebbe essere molto più complesso.

Infine, le assunzioni strutturali forti e rigide imposte da questi modelli creano una griglia concettuale che mal si adatta alla natura intrinsecamente dinamica, mutevole e continua dei sistemi industriali reali. La realtà operativa è caratterizzata da condizioni non stazionarie e da concept drift, fenomeni per i quali i modelli tradizionali, per loro stessa costruzione, non possiedono meccanismi di adattamento efficaci e robusti [32].

Questi limiti, presi nel loro insieme, non hanno lo scopo di invalidare il valore storico o il potenziale utilizzo contestuale di questi approcci. Essi rimangono strumenti potenti e utili

per problemi ben posti, di bassa dimensionalità e con dinamiche relativamente semplici, dove la loro interpretabilità può essere un vantaggio decisivo. Tuttavia, essi evidenziano con una forza quasi assiomatica la necessità di un nuovo paradigma prognostico, un salto di qualità metodologico che sia nativamente equipaggiato per affrontare queste sfide.

La comunità scientifica e ingegneristica ha individuato la risposta a questa pressante necessità nell'ampio, potente e flessibile arsenale di tecniche offerte dal Machine Learning. Gli algoritmi di Machine Learning, e in particolare le architetture di Deep Learning, sono, per loro natura e costruzione, progettati per superare molte di queste limitazioni. Sono approssimatori di funzioni universali, intrinsecamente capaci di modellare complesse e arbitrarie relazioni non lineari direttamente dai dati. Sono costruiti per operare in spazi ad alta dimensionalità, utilizzando potenti tecniche di regolarizzazione (come il dropout o il weight decay) per combattere l'overfitting. Le architetture più avanzate, come le Reti Neurali Convoluzionali e Ricorrenti, possono eseguire il feature learning end-to-end, apprendendo automaticamente rappresentazioni gerarchiche e significative dei dati direttamente dai segnali grezzi, liberando l'ingegnere dalla servitù del feature engineering manuale. Infine, sono modelli agnostici e non parametrici, che non impongono rigide assunzioni a priori sulla struttura dei dati o sulla dinamica del processo, ma lasciano che siano i dati stessi a "parlare".

La transizione verso questo nuovo paradigma, pertanto, non è una mera scelta stilistica o una rincorsa all'ultima moda tecnologica. È una risposta logica, coerente e scientificamente necessaria ai problemi lasciati irrisolti dal paradigma precedente. Il Machine Learning non offre una "panacea" — introduce a sua volta nuove e complesse sfide, come la necessità di enormi quantità di dati di addestramento etichettati e il profondo problema dell'interpretabilità dei suoi modelli "black-box" — ma fornisce un insieme di strumenti concettualmente più potenti e metodologicamente più adatti a estrarre conoscenza e valore predittivo dalla complessità del mondo reale. La disamina approfondita di questo nuovo e promettente paradigma, con tutte le sue potenzialità, le sue diverse famiglie di algoritmi e le sue sfide, costituirà l'oggetto di analisi del capitolo successivo, che esplorerà come il Machine Learning stia, di fatto, rivoluzionando il campo della Manutenzione Predittiva e aprendo orizzonti prima inimmaginabili.

# Capitolo 4: Le Metodologie Innovative: Machine Learning e Digital Twin

Il presente capitolo si addentra nel cuore pulsante di questo nuovo paradigma, rispondendo in modo diretto, sistematico e approfondito alla seconda e più corposa domanda di ricerca di questa tesi (SQ2). L'obiettivo è esplorare le metodologie innovative che stanno attualmente ridefinendo la frontiera della Manutenzione Predittiva, segnando una transizione che è tanto tecnologica quanto filosofica.

Il passaggio al Machine Learning non rappresenta un semplice cambiamento di strumenti o l'adozione di un algoritmo più "potente". È una rivoluzione filosofica nel modo di approcciare la modellazione. Se i modelli tradizionali richiedevano al ricercatore di specificare a priori, sulla base della propria conoscenza, la forma funzionale della relazione tra le variabili (es. una relazione lineare, una distribuzione di Weibull), gli algoritmi di Machine Learning sono, per loro natura, approssimatori di funzioni universali. Sono progettati per essere "agnostici": non impongono rigide assunzioni sulla distribuzione statistica dei dati o sulla linearità dei fenomeni, ma utilizzano la loro elevata capacità rappresentativa (spesso garantita da milioni di parametri interni) per apprendere, direttamente e unicamente dai dati, qualsiasi relazione complessa e non lineare che possa esistere tra l'input (i dati dei sensori) e l'output (lo stato di salute o la RUL). Questa flessibilità intrinseca li rende strumenti ideali per modellare i processi di degrado fisico, con tutte le loro non linearità, le interazioni complesse e i comportamenti a soglia che i modelli classici faticavano a rappresentare.

## 4.1. L'Intelligenza Artificiale come Driver di Innovazione per l'Industria 4.0

Per comprendere appieno la portata e la natura di questo nuovo paradigma, incarnato dal Machine Learning, è indispensabile, in primo luogo, contestualizzare questa transizione all'interno della più vasta e profonda narrazione della Quarta Rivoluzione Industriale, nota a livello globale come Industria 4.0, e riconoscere il ruolo centrale, propulsivo e quasi ontologico che l'Intelligenza Artificiale (IA) gioca in questo nuovo ecosistema socio-tecnico.

Il termine "Industria 4.0", coniato per la prima volta in Germania nel 2011, descrive la quarta grande fase di trasformazione del settore manifatturiero, una fase che promette di essere tanto dirompente quanto le precedenti. Se la Prima Rivoluzione Industriale è stata innescata dalla meccanizzazione tramite l'energia del vapore, la Seconda dall'elettrificazione e dall'introduzione della produzione di massa, e la Terza dall'informatica e dall'automazione dei processi tramite i Controllori a Logica Programmabile (PLC) e i sistemi SCADA, la Quarta Rivoluzione si distingue per una caratteristica unica e fondamentale. Non è trainata da una singola tecnologia disruptive, ma dalla convergenza e dalla fusione sinergica di un insieme di tecnologie abilitanti che abbattono la storica barriera tra il mondo fisico degli atomi (le macchine, i prodotti, gli operatori) e il mondo digitale dei bit (l'informazione, i dati, i modelli). Questa fusione dà vita a quello che viene definito Sistema Ciber-Fisico (*Cyber-Physical System* - CPS), un sistema in cui meccanismi fisici sono strettamente e ricorsivamente controllati e monitorati da algoritmi basati su computer, creando un ciclo di feedback continuo tra il mondo fisico e la sua rappresentazione computazionale [60].

In questo ecosistema, l'Internet of Things (IoT), con la sua rete pervasiva di sensori, attuatori e dispositivi connessi, agisce come il sistema nervoso sensoriale della fabbrica, percependo in tempo reale ogni aspetto del processo produttivo. I Big Data, descritti dalle loro celebri "V" (Volume, Velocità, Varietà, Veridicità, Valore), rappresentano il flusso sanguigno di questo sistema, il diluvio di informazioni generato dai sensori. Il Cloud Computing fornisce l'infrastruttura scalabile e on-demand per l'archiviazione e l'elaborazione di questi dati, agendo come una sorta di memoria esterna e di potenza di calcolo quasi infinita. Tecnologie come la Manifattura Additiva (stampa 3D) e la Realtà Aumentata offrono nuove e potenti modalità di interazione e di materializzazione delle informazioni digitali nel mondo fisico.

Tuttavia, considerare queste tecnologie come un semplice elenco di componenti isolati sarebbe un errore concettuale profondo. Esse sono, in realtà, gli organi interconnessi di un unico, grande sistema. Ed è proprio in questo contesto che emerge il ruolo insostituibile e centrale dell'Intelligenza Artificiale. Se l'IoT è il sistema nervoso sensoriale che percepisce, e il cloud è la memoria che immagazzina, l'IA ne costituisce il cervello cognitivo. È l'entità capace di processare, correlare, comprendere e, soprattutto, apprendere da queste informazioni per trasformarle da dati grezzi in conoscenza, da conoscenza in previsioni e da previsioni in decisioni intelligenti e azioni ottimizzate. Senza l'IA, l'Industria 4.0 sarebbe un gigante con sensi acutissimi e una memoria prodigiosa, ma privo di intelligenza; sarebbe un sistema sommerso da un'alluvione di dati che non è in grado di interpretare, un "data-rich, information-poor syndrome". I Big Data sono una risorsa potenziale, un giacimento di

petrolio grezzo, ma è l'IA, con i suoi algoritmi di Machine Learning, a essere la raffineria che li trasforma in carburante per l'innovazione e la competitività [67].

Il contributo dell'IA va oltre la semplice analisi dei dati; essa abilita un salto qualitativo fondamentale, una transizione che definisce l'essenza stessa dell'Industria 4.0: il passaggio dall'automazione all'autonomia. L'Industria 3.0, la rivoluzione dell'informatica, ha portato all'automazione, ovvero alla capacità di macchine e robot di eseguire compiti predefiniti in modo rapido, preciso e instancabile, ma seguendo una programmazione rigida e deterministica. Un robot industriale tradizionale è un esempio perfetto di automazione: esegue la sua sequenza di movimenti con una precisione sovrumana, ma è completamente incapace di adattarsi a un cambiamento imprevisto nel suo ambiente. L'Industria 4.0, grazie all'IA, mira all'autonomia: la capacità di sistemi, macchine e interi processi di percepire il loro contesto, di apprendere dall'esperienza (dai dati), e di prendere decisioni per adattare e ottimizzare il proprio comportamento in tempo reale, senza un intervento umano diretto o con una supervisione minima. La Manutenzione Predittiva è l'incarnazione perfetta di questo passaggio: si abbandona un piano di manutenzione *automatico* basato su un calendario statico, per abbracciare un sistema *autonomo* che decide quando e come intervenire sulla base di una previsione costantemente aggiornata dello stato di salute futuro dell'asset.

Per navigare con chiarezza in questo campo, è essenziale padroneggiare la gerarchia dei termini. Intelligenza Artificiale (IA) è il termine ombrello, la vasta e storica disciplina scientifica, il cui sogno, nato formalmente con la conferenza di Dartmouth del 1956, è quello di creare macchine capaci di esibire comportamenti che riterremmo intelligenti se fossero eseguiti da un essere umano [92]. All'interno dell'IA, si colloca il Machine Learning (ML), un sottocampo che ha guadagnato una trazione esplosiva negli ultimi decenni. La sua essenza, magnificamente catturata da Tom Mitchell [69], è quella di dare ai computer la capacità di apprendere senza essere stati esplicitamente programmati. Un programma apprende da un'esperienza E rispetto a una classe di compiti T e una misura di performance P, se la sua performance nel compito T, misurata da P, migliora con l'esperienza E. Invece di scrivere regole rigide, si fornisce all'algoritmo una grande quantità di dati esemplificativi e un obiettivo, e l'algoritmo "impara" da solo le regole, i pattern e le funzioni complesse che collegano l'input all'output desiderato. A sua volta, all'interno del Machine Learning, un sottocampo specifico ha causato una vera e propria rivoluzione, quasi un cambio di paradigma all'interno del paradigma: il Deep Learning. Il Deep Learning si basa sull'uso di Reti Neurali Artificiali profonde (*Deep Neural Networks* - DNN), ovvero architetture ispirate alla struttura gerarchica della corteccia visiva umana, con un numero molto elevato

di strati di neuroni artificiali. La sua caratteristica distintiva e rivoluzionaria è la capacità di eseguire il representation learning end-to-end: non richiede un processo di feature engineering manuale, ma apprende autonomamente una gerarchia di rappresentazioni (o feature) sempre più complesse e astratte, direttamente dai dati grezzi. Ai livelli più bassi impara feature semplici, e ai livelli più alti combina queste feature per creare concetti via via più astratti, ottimizzando l'intera gerarchia per il compito di previsione finale [59].

In conclusione, l'Intelligenza Artificiale non è semplicemente "una" delle tante tecnologie abilitanti dell'Industria 4.0; ne è il motore cognitivo, il catalizzatore che permette a tutte le altre tecnologie di raggiungere il loro pieno potenziale e di integrarsi in un sistema coerente e intelligente. È l'IA che trasforma il diluvio di dati dell'IoT in previsioni azionabili e in diagnosi precise. È l'IA che conferisce autonomia e adattabilità ai robot. È l'IA che ottimizza le complesse catene logistiche in tempo reale. Ed è l'IA che potenzia la creatività umana nel processo di progettazione.

## 4.2. Tecniche di Machine Learning per la Manutenzione Predittiva

È indispensabile "aprire il cofano" di questa intelligenza per analizzare in dettaglio gli strumenti specifici, gli algoritmi, che ne rendono possibile l'applicazione pratica e ne determinano l'efficacia trasformativa. Questi strumenti appartengono al vasto, potente e in continua e febbrile evoluzione dominio del Machine Learning (ML). Il Machine Learning, come sottocampo dell'IA, non è semplicemente una nuova tecnica statistica, ma rappresenta un cambiamento fondamentale nella relazione tra l'uomo, il computer e i dati. Esso si concentra sullo sviluppo di algoritmi capaci di apprendere pattern, relazioni e funzioni complesse direttamente dai dati, superando la necessità di una programmazione esplicita basata su regole fisse e deterministiche, che per decenni ha dominato l'informatica e l'ingegneria dei sistemi.

Nel contesto della Manutenzione Predittiva, l'adozione del Machine Learning non è un mero aggiornamento tecnologico o l'aggiunta di un nuovo, più potente, strumento alla cassetta degli attrezzi dell'ingegnere. È, a tutti gli effetti, la risposta diretta e scientificamente necessaria ai limiti strutturali dei modelli statistici tradizionali, che sono stati analizzati criticamente e in profondità nel capitolo precedente. Laddove i modelli classici richiedevano

rigide assunzioni sulla distribuzione dei dati, sulla linearità delle relazioni e sulla stazionarietà dei processi, e faticavano a gestire la complessità dei dati multidimensionali, gli algoritmi di ML sono stati progettati per eccellere proprio in questi scenari. La loro forza risiede nella loro capacità intrinseca di agire come approssimatori di funzioni universali, capaci di apprendere, con sufficienti dati e capacità computazionale, qualsiasi relazione non lineare arbitraria che possa esistere tra l'input e l'output [45] [24].

L'applicazione del ML alla prognostica trasforma radicalmente l'approccio alla modellazione, spostandolo da uno di modellazione statistica esplicita a uno di apprendimento da esempi. Il fulcro del lavoro dell'ingegnere e del data scientist si sposta dalla derivazione di complesse equazioni matematiche basate su una conoscenza a priori, alla preparazione, pulizia, arricchimento e strutturazione di dataset di alta qualità. L'obiettivo non è più quello di postulare una forma funzionale per il degrado e di stimarne i parametri, ma piuttosto quello di raccogliere dati storici che siano il più possibile rappresentativi del comportamento di un macchinario e di utilizzarli per "addestrare" un modello. Questo processo di addestramento, tipicamente un'ottimizzazione iterativa che minimizza una funzione di costo, permette al modello di scoprire autonomamente le complesse e spesso non intuitive relazioni latenti nei dati. La speranza, e l'obiettivo ultimo, è che la conoscenza così appresa possa essere generalizzata per fare previsioni accurate su nuovi dati futuri, mai visti durante la fase di training. La capacità di generalizzazione, ovvero la performance del modello su dati inediti, diventa così la metrica ultima e più onesta del suo successo, la vera misura della sua intelligenza appresa [35].

Il panorama degli algoritmi di Machine Learning è estremamente vasto e può essere classificato secondo diversi paradigmi di apprendimento, ciascuno adatto a un diverso tipo di problema e a un diverso tipo di disponibilità di dati. Per gli scopi della Manutenzione Predittiva, due di questi paradigmi sono di importanza centrale e predominante, poiché mappano direttamente i principali compiti della diagnosi e della prognosi: l'Apprendimento Supervisionato (*Supervised Learning*) e l'Apprendimento Non Supervisionato (*Unsupervised Learning*). Un terzo paradigma, il Reinforcement Learning (Apprendimento per Rinforzo), in cui un "agente" impara a compiere azioni in un ambiente per massimizzare una ricompensa cumulativa, sebbene estremamente potente e di crescente interesse, trova applicazioni più di nicchia in questo specifico contesto. Si orienta maggiormente all'ottimizzazione dinamica delle politiche di manutenzione e controllo (ad esempio, "dato lo stato di salute previsto, qual è l'azione ottimale da intraprendere ora per estendere la vita

residua del componente?"), un problema di teoria delle decisioni che esula dall'obiettivo primario di diagnosi e stima della RUL qui trattato [109].

### 4.2.1. Apprendimento Supervisionato

L'Apprendimento Supervisionato (*Supervised Learning*) rappresenta il paradigma di Machine Learning più studiato, più compreso e, in molte applicazioni industriali di successo, più potente. La sua filosofia è radicata in un concetto tanto semplice quanto fondamentale: l'apprendimento da esempi con un "insegnante". La metafora della "supervisione" è calzante: l'algoritmo non esplora i dati alla cieca, ma viene addestrato su un dataset "etichettato" (*labeled data*), in cui un "supervisore" (tipicamente un esperto umano o il risultato di un esperimento controllato) ha fornito la "risposta corretta" per ogni campione di dati. L'obiettivo dell'algoritmo, durante la fase di addestramento, è quindi quello di produrre una funzione di mappatura generale, $f$, che possa associare nel modo più accurato possibile un dato vettore di input, $X$ (che rappresenta le feature estratte dai sensori), a un dato output, $y$ (che rappresenta lo stato di salute o la RUL). Il processo di apprendimento consiste in un'ottimizzazione iterativa, spesso basata su algoritmi di discesa del gradiente, in cui i parametri interni del modello vengono sistematicamente aggiustati per minimizzare una funzione di costo (*cost function*), una misura matematica della discrepanza tra le previsioni del modello e le etichette reali [40].

Una volta che il modello è stato addestrato, la sua vera prova del nove è la sua capacità di generalizzare. Un modello utile non è quello che ricorda perfettamente gli esempi di addestramento, ma quello che è in grado di fare previsioni accurate su nuovi dati mai visti prima, dimostrando di aver catturato la relazione sottostante e non semplicemente di aver "imparato a memoria" il rumore e le peculiarità del training set. La lotta contro l'overfitting (l'eccessivo adattamento ai dati di training a scapito della generalizzazione) è, di fatto, una delle sfide centrali e ricorrenti di tutto il Machine Learning supervisionato.

Nel contesto della Manutenzione Predittiva, la creazione di un dataset etichettato di alta qualità è spesso l'operazione più costosa, lunga e complessa, e costituisce il principale collo di bottiglia per l'applicazione di queste tecniche. Le etichette possono essere di due tipi principali, che a loro volta definiscono i due compiti fondamentali dell'apprendimento supervisionato: la classificazione e la regressione.

Il primo compito, la classificazione, si verifica quando l'etichetta di output è una variabile categorica e discreta. Nel nostro contesto, questo si traduce direttamente nel problema della diagnosi dei guasti. L'obiettivo è addestrare un modello classificatore a riconoscere la "classe" di appartenenza di un nuovo campione di dati. Il dataset di addestramento, in questo caso, consisterà in segmenti di segnale (o le relative feature) a cui è stata associata un'etichetta di classe come "Sano", "Guasto alla Pista Esterna di un Cuscinetto", "Guasto alla Pista Interna", "Danno a un Dente di Ingranaggio" o "Squilibrio del Rotore". La capacità di un modello di eseguire questa classificazione in modo accurato e tempestivo permette di identificare non solo la presenza di un guasto, ma anche la sua tipologia e la sua localizzazione, fornendo informazioni cruciali per pianificare un intervento di riparazione mirato.

Il secondo compito, la regressione, si verifica quando l'etichetta di output è un valore numerico continuo. Questo compito si mappa perfettamente sul problema della prognosi e, più specificamente, sulla stima quantitativa della RUL. In questo caso, il dataset di addestramento deve contenere, per ogni istante di tempo, le feature estratte dai sensori e il valore esatto della vita utile residua in quell'istante. La creazione di un tale dataset è estremamente onerosa, poiché richiede tipicamente l'esecuzione di esperimenti di run-to-failure in laboratorio, in cui i componenti vengono fatti funzionare fino al guasto sotto monitoraggio continuo. Solo al termine dell'esperimento, conoscendo il tempo totale di vita, è possibile etichettare a ritroso ogni punto della storia del componente con la RUL corrispondente. Questa necessità di dati di vita completi è una delle ragioni per cui la prognosi supervisionata, sebbene potente, è più difficile da implementare in pratica rispetto alla diagnosi.

Le Support Vector Machines (SVM), introdotte da Vladimir Vapnik e i suoi collaboratori negli anni '90, rappresentano uno degli algoritmi più potenti e teoricamente eleganti del Machine Learning classico, radicati nella solida teoria dell'apprendimento statistico [116]. L'idea fondamentale di una SVM, nel caso più semplice di classificazione binaria, è geometricamente intuitiva: trovare l'iperpiano che non solo separa correttamente i punti delle due classi, ma lo fa nel modo più netto e robusto possibile. Il concetto di "separazione più netta" viene formalizzato matematicamente come la ricerca dell'iperpiano che massimizza il margine, ovvero la distanza tra l'iperpiano stesso e i punti più vicini di ciascuna classe. Questi punti, che giacciono sul bordo del margine, sono detti vettori di supporto (*support vectors*), poiché sono gli unici punti del dataset di addestramento che definiscono attivamente la posizione e l'orientamento del confine decisionale. Questa focalizzazione sui punti più

difficili e ambigui da classificare (quelli al confine) rende le SVM particolarmente robuste al rumore e agli outlier presenti nel resto dei dati.

La vera potenza e flessibilità delle SVM, tuttavia, risiede nel celebre "trucco del kernel" (*kernel trick*). Riconoscendo che molti problemi non sono linearmente separabili nello spazio delle feature originali, il trucco del kernel permette alle SVM di operare in uno spazio a dimensionalità molto più elevata (potenzialmente infinita) senza mai dover calcolare esplicitamente le coordinate dei punti in questo spazio. Attraverso l'uso di una funzione kernel, che calcola il prodotto scalare tra le immagini di due punti in questo spazio ad alta dimensionalità, l'algoritmo di ottimizzazione può trovare un iperpiano di separazione lineare in questo nuovo spazio, che corrisponde a un confine decisionale non lineare e altamente complesso nello spazio originale. Funzioni kernel comuni includono il kernel lineare (che riduce la SVM a un classificatore lineare), il kernel polinomiale, il kernel sigmoide e, il più diffuso e versatile, il kernel a base radiale (*Radial Basis Function* - RBF). Nella diagnosi dei guasti, le SVM, specialmente con kernel RBF, sono state ampiamente e con successo utilizzate per classificare diversi tipi di guasto in motori, cuscinetti e ingranaggi, mostrando un'elevata accuratezza e una buona capacità di generalizzazione [94]. L'adattamento delle SVM ai problemi di regressione, noto come Support Vector Regression (SVR), si basa su un'idea altrettanto ingegnosa, basata sulla definizione di una "zona di insensibilità" (*epsilon-insensitive tube*). La SVR cerca di trovare una funzione che si adatti ai dati in modo tale che il maggior numero possibile di punti cada all'interno di questo tubo, ignorando gli errori al suo interno e penalizzando solo quelli all'esterno. Questo la rende robusta a piccole fluttuazioni e, grazie al kernel trick, capace di modellare traiettorie di degrado altamente non lineari [37].

Un singolo modello, per quanto sofisticato, può essere soggetto a errori, a bias specifici o a overfitting. Per superare questo limite, il Machine Learning ha sviluppato i modelli di insieme (*ensemble learning*), che si basano su un principio tanto semplice quanto potente, spesso riassunto nell'adagio della "saggezza della folla". L'idea è quella di combinare le previsioni di un gran numero di modelli più semplici e diversificati (i "membri" dell'ensemble) per ottenere una previsione finale che sia più accurata, più stabile e più robusta di quella che ogni singolo membro potrebbe fornire.

L'algoritmo di insieme che ha probabilmente avuto il maggiore impatto pratico e che è diventato uno standard de facto per molti problemi di classificazione e regressione è il Random Forest, introdotto da Leo Breiman in un articolo seminale del 2001. Un Random Forest è un ensemble di alberi decisionali. Un singolo albero decisionale è un modello

"white-box" molto intuitivo, ma è anche notoriamente instabile e incline all'overfitting. Il Random Forest risolve brillantemente questo problema costruendo una "foresta" di centinaia o migliaia di alberi decisionali, ciascuno addestrato in modo da essere leggermente diverso e, idealmente, de-correlato dagli altri. La diversità, che è la chiave del successo di qualsiasi ensemble, è ottenuta attraverso due meccanismi di randomizzazione intelligenti:

1. *Bagging* (Bootstrap Aggregating): Ogni albero della foresta non viene addestrato sull'intero dataset, ma su un sotto-campione casuale con rimpiazzo (un campione di bootstrap) della stessa dimensione del dataset originale. Questo assicura che ogni albero veda una versione leggermente diversa dei dati, rendendoli meno sensibili a singoli punti influenti.

2. *Random Subspace Method* (Selezione Casuale delle Feature): Durante la costruzione di ogni albero, a ogni nodo, invece di cercare il miglior taglio tra tutte le feature disponibili, l'algoritmo ne considera solo un sottoinsieme casuale. Questo meccanismo, apparentemente semplice, è cruciale: impedisce a poche feature molto forti di dominare la struttura di tutti gli alberi, forzando la foresta a esplorare una varietà più ampia di relazioni nei dati e rendendo gli alberi più diversi e de-correlati.

Il risultato di questo doppio processo di randomizzazione è una collezione di alberi "esperti" di aspetti diversi del problema. La previsione finale viene fatta aggregando le loro "opinioni": tramite un voto di maggioranza per i problemi di classificazione, o calcolando la media delle previsioni per i problemi di regressione. Questo processo di aggregazione ha l'effetto statistico di ridurre drasticamente la varianza del modello complessivo, senza aumentarne significativamente il bias, rendendolo estremamente robusto all'overfitting e al rumore nei dati.
I Random Forest sono diventati uno strumento di lavoro prediletto nella comunità della PdM per una serie di ragioni convincenti. Offrono un'accuratezza predittiva che è spesso allo stato dell'arte, competendo e talvolta superando modelli più complessi. Sono relativamente facili da usare e da sintonizzare, con pochi parametri cruciali da impostare. Sono in grado di gestire in modo nativo un gran numero di feature e, un vantaggio notevole e quasi unico, forniscono una misura intrinseca dell'importanza delle feature (basata su quanto una feature contribuisce a ridurre l'impurità dei nodi o sull'aumento dell'errore quando i suoi valori vengono permutati casualmente), permettendo di capire quali variabili sono più predittive.

Questa combinazione di alta performance, robustezza e interpretabilità parziale li rende un modello di riferimento essenziale in qualsiasi progetto di diagnosi o prognosi supervisionata [105] [42].
In sintesi, l'apprendimento supervisionato fornisce un corredo di tecniche potenti, flessibili e matematicamente fondate per affrontare i compiti di diagnosi e prognosi, a condizione che si disponga di dati etichettati di alta qualità. La scelta tra algoritmi come SVM e Random Forest dipende dalle specifiche del problema, ma entrambi rappresentano un salto qualitativo significativo rispetto ai modelli statistici tradizionali. Tuttavia, la loro dipendenza dalle etichette e dal feature engineering manuale rimane un ostacolo, che altri paradigmi di apprendimento, come l'apprendimento non supervisionato e il deep learning, cercano di affrontare.

### 4.2.2. Apprendimento Non Supervisionato

L'Apprendimento Non Supervisionato rappresenta un cambio di prospettiva fondamentale, una transizione da un paradigma di apprendimento basato sull'imitazione (apprendere a mappare input a output noti) a uno basato sulla scoperta. In questo caso, l'algoritmo viene messo di fronte a un dataset "non etichettato" (*unlabeled data*). Non viene fornito alcun segnale di supervisione, nessuna "risposta corretta" da un insegnante esterno. L'obiettivo non è predire un output predefinito, ma agire come un esploratore, un "archeologo dei dati", per scoprire la struttura intrinseca, i pattern nascosti, i raggruppamenti naturali e le anomalie all'interno dei dati stessi. Questo paradigma è di enorme valore pratico e strategico in scenari di Manutenzione Predittiva, poiché affronta la sfida più grande e più comune dell'apprendimento supervisionato: la scarsità di dati etichettati di alta qualità. In contesti industriali reali, è relativamente facile e poco costoso raccogliere enormi quantità di dati da macchinari in funzione, ma è estremamente difficile, oneroso e spesso impossibile etichettare ogni singolo istante di questi dati con una precisa classe di guasto o un valore di RUL accurato. L'apprendimento non supervisionato fornisce gli strumenti per estrarre conoscenza e valore da questi vasti "oceani" di dati non etichettati, trasformando quello che altrimenti sarebbe un costo di archiviazione in una risorsa strategica.

I due compiti principali e più rilevanti dell'apprendimento non supervisionato, nel nostro contesto, sono il clustering, che cerca di trovare gruppi omogenei nei dati, e l'anomaly detection, che cerca di identificare osservazioni anomale e inaspettate.

Il clustering è il compito fondamentale di raggruppare un insieme di dati in sottoinsiemi, i cluster, in modo tale che gli oggetti all'interno di uno stesso cluster siano molto simili tra loro (alta intra-cluster similarity) e gli oggetti in cluster diversi siano molto dissimili (bassa inter-cluster similarity). La nozione di "similarità" o "dissimilarità" è definita formalmente da una metrica di distanza o di prossimità nello spazio delle feature. Nel contesto della PdM, il clustering può essere utilizzato per un compito di importanza cruciale: scoprire automaticamente e in modo agnostico i diversi stati operativi o di salute di un macchinario a partire dai dati dei sensori, senza che questi stati siano stati definiti a priori da un esperto.

Immaginiamo di raccogliere dati di vibrazione e temperatura da un macchinario durante il suo intero ciclo di vita. Applicando un algoritmo di clustering a questo dataset multidimensionale, potremmo scoprire che i dati non si distribuiscono in modo uniforme, ma si raggruppano naturalmente in un numero finito di "isole" o regioni dense nello spazio delle feature. Un'analisi successiva di questi cluster da parte di un ingegnere potrebbe rivelare che essi corrispondono a stati fisici ben precisi e interpretabili: un grande cluster centrale potrebbe rappresentare lo "stato sano" o normale; altri cluster distinti potrebbero corrispondere a diversi "regimi operativi" (es. lavorazione a basso carico vs. lavorazione ad alto carico); e altri cluster ancora, più piccoli e progressivamente più lontani dal cluster "sano", potrebbero rappresentare le diverse fasi del degrado ("usura lieve", "usura grave", "guasto incipiente").

Una volta identificati e validati, questi cluster diventano una potente fonte di conoscenza. In primo luogo, possono fornire una diagnosi non supervisionata: monitorando in quale cluster si trova il punto rappresentativo dello stato attuale del macchinario, si può avere un'idea qualitativa della sua salute e del suo regime operativo. In secondo luogo, e forse ancora più potentemente, questi cluster possono essere usati come "pseudo-etichette". Una volta che un esperto ha assegnato un significato a ciascun cluster (es. "sano", "guasto A"), queste etichette possono essere propagate a tutti i punti all'interno di quel cluster, trasformando un dataset non etichettato in uno parzialmente etichettato. Questo permette di addestrare un classificatore supervisionato, in un approccio noto come apprendimento semi-supervisionato, che sfrutta la grande quantità di dati non etichettati per migliorare

drasticamente le performance di un modello addestrato su un piccolo set di dati etichettati [20].
Tra i numerosi algoritmi di clustering, due sono particolarmente rappresentativi di filosofie opposte:

- k-Means Clustering: È l'algoritmo di clustering più famoso e utilizzato, un vero e proprio "cavallo di battaglia" grazie alla sua semplicità concettuale ed efficienza computazionale. È un algoritmo iterativo di partizionamento che suddivide il dataset in un numero *k* di cluster predefinito. Inizia scegliendo casualmente *k* punti come "centroidi" iniziali. Successivamente, ripete due passaggi fino a convergenza: (1) l'E-step (Expectation), in cui ogni punto del dataset viene assegnato al cluster il cui centroide gli è più vicino; (2) l'M-step (Maximization), in cui la posizione di ogni centroide viene ricalcolata come la media (il baricentro) di tutti i punti che gli sono stati assegnati. Sebbene sia veloce e semplice, k-Means ha dei limiti noti e ben documentati: richiede che l'utente specifichi a priori il numero di cluster, *k*, una scelta non banale e spesso difficile da giustificare; è sensibile all'inizializzazione casuale dei centroidi, potendo convergere a ottimi locali; e, soprattutto, la sua metrica di distanza Euclidea e la sua ricerca di un centroide basato sulla media lo portano a funzionare bene solo con cluster che sono di forma approssimativamente sferica (isotropica) e di dimensioni simili, faticando a identificare cluster di forme più complesse o allungate [47].

- DBSCAN (*Density-Based Spatial Clustering of Applications with Noise*): A differenza di k-Means, DBSCAN adotta una filosofia basata sulla densità, più flessibile e spesso più realistica. Invece di partizionare tutti i dati, esso definisce i cluster come regioni dense di punti, separate da regioni a bassa densità. L'algoritmo è governato da due parametri: una distanza di vicinato, *ε (epsilon)*, e un numero minimo di punti, *MinPts*. Un punto è considerato un "punto core" se nel suo intorno di raggio *ε* ci sono almeno *MinPts* punti. I cluster vengono quindi formati come insiemi di punti "densamente raggiungibili" a partire dai punti core. Questo approccio basato sulla densità conferisce a DBSCAN due vantaggi enormi: è in grado di trovare cluster di forma arbitraria e, cosa di importanza cruciale per la PdM, è in grado di identificare i punti che non appartengono a nessun cluster, classificandoli esplicitamente come rumore o outlier. Questa caratteristica lo rende non solo un

algoritmo di clustering, ma anche un efficace strumento di anomaly detection. I suoi principali punti di forza sono che non richiede di specificare il numero di cluster a priori e la sua robustezza al rumore. La sua principale sfida è la scelta sensibile dei parametri $\varepsilon$ e *MinPts*, che possono essere difficili da impostare per dataset con densità molto variabili [28].

L'Anomaly Detection (o *outlier detection*) è il compito specifico di identificare dati, eventi o osservazioni che si discostano in modo significativo dal comportamento "normale" o atteso. È la scienza della ricerca dell'inaspettato. In Manutenzione Predittiva, questo è un compito di importanza cruciale e di grande applicabilità pratica, poiché un'anomalia può essere il primo, debole ma fondamentale segnale di un guasto incipiente, di un degrado non ancora catastrofico, o di una condizione operativa imprevista e potenzialmente dannosa. Spesso, si opera in un contesto "one-class" o "semi-supervisionato", in cui si dispone di una grande quantità di dati raccolti durante il funzionamento sano e normale del macchinario, e l'obiettivo è costruire un modello della "normalità" per poter rilevare qualsiasi deviazione futura da essa [19].

- *Isolation Forest*: Questo algoritmo, introdotto da Liu, Ting e Zhou (2008) [64], si basa su un'idea tanto semplice quanto potente: le anomalie sono, per definizione, "poche e diverse", e quindi dovrebbero essere più facili da "isolare" rispetto ai punti normali, che sono densi e raggruppati. L'algoritmo costruisce un insieme di alberi decisionali casuali, detti *isolation trees*. A differenza dei Random Forest, questi alberi non vengono costruiti per predire un'etichetta, ma semplicemente per partizionare i dati. In ogni albero, i dati vengono separati ricorsivamente scegliendo a caso una feature e un punto di taglio casuale all'interno del range di quella feature. L'intuizione è che un punto anomalo, essendo isolato ai margini della distribuzione dei dati, richiederà in media molti meno tagli (un percorso molto più breve dalla radice alla foglia) per essere isolato rispetto a un punto normale, che si trova in una regione densa. La lunghezza media del percorso attraverso la foresta viene quindi usata come punteggio di anomalia. L'Isolation Forest è computazionalmente molto efficiente, funziona bene in spazi ad alta dimensionalità e non richiede una metrica di distanza, rendendolo uno strumento molto versatile e popolare.

- *One-Class SVM*: Questa è un'applicazione della potente filosofia delle Support Vector Machines al problema dell'anomaly detection, in un contesto in cui si dispone solo di esempi della classe "normale". Invece di trovare un iperpiano che separa due classi di dati, la One-Class SVM (OC-SVM) cerca di trovare un iperpiano che separa i dati "normali" dall'origine nello spazio delle feature, cercando di massimizzare la distanza dall'origine. Utilizzando il "kernel trick", questo equivale a trovare un "confine" o una "bolla" non lineare che racchiude la maggior parte dei dati normali. Qualsiasi nuovo punto che cade al di fuori di questo confine appreso viene classificato come un'anomalia. È una tecnica potente, capace di apprendere confini complessi e non lineari per definire la "normalità", ma la sua performance è sensibile alla scelta del kernel e dei suoi parametri (in particolare il parametro *nu*, che controlla la frazione di outlier attesi) [98].

In conclusione, l'apprendimento non supervisionato fornisce un corredo di tecniche indispensabili per estrarre valore e conoscenza dai vasti dataset non etichettati che caratterizzano il mondo industriale. Permette di scoprire automaticamente gli stati di un sistema e, soprattutto, di rilevare le prime, sottili deviazioni dalla normalità, agendo come un sistema di allarme precoce, un "canarino nella miniera di carbone" digitale. Spesso, questi approcci non sono un'alternativa, ma un complemento a quelli supervisionati, fornendo le basi (attraverso la scoperta di cluster o la rilevazione di anomalie) per un'analisi più mirata e approfondita. Tuttavia, anch'essi, nella loro forma classica, dipendono criticamente dalla qualità delle feature ingegnerizzate manualmente. La frontiera successiva, il Deep Learning, promette di superare anche questo ultimo ostacolo, apprendendo le feature direttamente dai dati.

### 4.2.3. Deep Learning

L'analisi condotta finora ha delineato un percorso evolutivo chiaro e coerente: dai modelli statistici tradizionali, con le loro rigide e spesso irrealistiche assunzioni, passando agli algoritmi di Machine Learning "classico" (come le Support Vector Machines e i Random Forest), che offrono una maggiore flessibilità e potenza nel modellare relazioni complesse e non lineari. Tuttavia, anche questi ultimi, pur rappresentando un significativo passo avanti,

condividono una limitazione fondamentale con i loro predecessori: la loro performance è intrinsecamente e indissolubilmente legata alla qualità e alla pertinenza delle feature ingegnerizzate manualmente. Questo processo di *feature engineering*, rappresenta un collo di bottiglia umano, soggettivo e laborioso, che limita la scalabilità e l'oggettività scientifica delle soluzioni prognostiche. È in questo contesto che emerge la più recente e, per certi versi, definitiva rivoluzione nel campo dell'Intelligenza Artificiale: il Deep Learning.

Il Deep Learning non è un paradigma di apprendimento completamente nuovo, ma piuttosto una re-immaginazione, una rinascita e una scalatura su vasta scala delle Reti Neurali Artificiali (*Artificial Neural Networks* - ANN), un'idea che affonda le sue radici negli studi pionieristici sulla cibernetica e sul connessionismo degli anni '40 e '50. Una rete neurale è un modello computazionale ispirato alla struttura e al funzionamento del cervello umano, composto da un gran numero di unità di elaborazione semplici (i "neuroni artificiali") organizzate in strati (*layers*). Ogni neurone riceve input ponderati dai neuroni dello strato precedente, calcola una somma di questi input, e applica una funzione di attivazione non lineare (come la sigmoide, la tangente iperbolica o, più comunemente oggi, la ReLU - Rectified Linear Unit) per produrre il suo output. Impilando più strati di neuroni tra lo strato di input e quello di output, si crea una rete neurale profonda (*Deep Neural Network* - DNN). È proprio questa "profondità", ovvero la presenza di molteplici strati nascosti (*hidden layers*), che conferisce al Deep Learning la sua caratteristica più potente e rivoluzionaria: la capacità di eseguire il representation learning end-to-end [59].

Questo concetto merita una disamina approfondita. Invece di richiedere all'ingegnere di progettare e calcolare a mano le feature dai dati grezzi, una rete neurale profonda impara autonomamente una gerarchia di rappresentazioni (o feature) a diversi livelli di astrazione. Gli strati più vicini all'input imparano a riconoscere feature semplici, locali e di basso livello. Ad esempio, se l'input è un'immagine, il primo strato potrebbe imparare a riconoscere bordi, angoli e gradienti di colore. Se l'input è un segnale di vibrazione, potrebbe imparare a riconoscere componenti di frequenza elementari o impulsi semplici. Gli strati successivi, prendendo in input le rappresentazioni apprese dallo strato precedente, le combinano per creare rappresentazioni via via più complesse e astratte. Nell'esempio dell'immagine, gli strati intermedi potrebbero imparare a riconoscere texture, forme geometriche o parti di oggetti (come un occhio o una ruota), mentre gli strati finali potrebbero apprendere a riconoscere l'oggetto intero. L'intero processo di apprendimento, guidato da algoritmi di ottimizzazione come la discesa del gradiente stocastico e l'algoritmo di back propagation (che calcola in modo efficiente il gradiente della funzione di costo rispetto a

ogni peso della rete), regola simultaneamente i pesi di tutti gli strati per ottimizzare l'intera gerarchia di feature in funzione del compito finale (ad esempio, la classificazione di un guasto o la stima della RUL). Questa capacità di automatizzare il processo di estrazione delle feature non è solo un enorme vantaggio in termini di efficienza e scalabilità, ma rappresenta un vero e proprio cambio di paradigma: sposta il focus dalla progettazione di feature "intelligenti" alla progettazione di architetture di rete che possano scoprire autonomamente le feature più predittive, anche quelle che potrebbero essere troppo complesse o non intuitive per essere scoperte e formulate da un essere umano [35].

Nel contesto della Manutenzione Predittiva, due famiglie di architetture di Deep Learning si sono rivelate particolarmente efficaci e hanno dominato la letteratura scientifica recente, ciascuna specializzata in un diverso tipo di struttura dei dati: le Reti Neurali Convoluzionali e le Reti Neurali Ricorrenti.

Le Reti Neurali Convoluzionali (*Convolutional Neural Networks* - CNN) sono una classe di architetture neurali specificamente progettata per processare dati che hanno una topologia a griglia, come le immagini (dati 2D) o le serie temporali (dati 1D). La loro architettura, ispirata alla gerarchia della corteccia visiva primaria del cervello dei mammiferi, si basa su tre idee fondamentali che ne sfruttano le proprietà di stazionarietà locale e di composizionalità: i campi recettivi locali, la condivisione dei pesi e il sotto-campionamento (pooling). Invece di collegare ogni neurone di uno strato a tutti i neuroni dello strato precedente (come in una rete *fully connected* o densa), i neuroni di uno strato convoluzionale sono collegati solo a una piccola regione locale dell'input, il loro "campo recettivo". Questo neurone agisce come un filtro (o kernel), ovvero una piccola matrice di pesi, che scorre (o "convolve") sull'intero input, cercando un pattern specifico (es. un bordo verticale). Poiché lo stesso filtro viene applicato su tutta l'immagine, i pesi sono condivisi, il che riduce drasticamente il numero di parametri da apprendere (rispetto a una rete densa) e, cosa fondamentale, rende il modello invariante alla traslazione del pattern. Gli strati di pooling (es. Max Pooling), invece, riducono progressivamente la dimensionalità spaziale della rappresentazione, aggregando le risposte dei neuroni in una regione e rendendo la rappresentazione più robusta a piccole deformazioni, traslazioni e distorsioni.

Sebbene nate e rese celebri per le loro performance rivoluzionarie nell'analisi di immagini 2D [55], le CNN si sono rivelate straordinariamente efficaci anche per l'analisi di dati di serie temporali 1D. L'approccio più comune consiste nel trasformare il segnale 1D in una rappresentazione 2D tempo-frequenza, come uno spettrogramma (ottenuto tramite STFT) o uno scalogramma (ottenuto tramite Wavelet Transform). Questa "immagine" del segnale,

che visualizza come il contenuto di frequenza cambia nel tempo, viene quindi data in input a una CNN, che può apprendere automaticamente i pattern tempo-frequenza caratteristici dei diversi stati di salute o tipi di guasto [120]. Un'altra strategia, ancora più "end-to-end", consiste nell'applicare convoluzioni 1D direttamente al segnale grezzo. In questo modo, la rete impara da sola i filtri ottimali per l'analisi del segnale, superando la necessità di scegliere a priori una trasformata specifica e potenzialmente sub-ottimale [46]. Grazie a questa capacità di apprendere feature gerarchiche e invarianti, le CNN hanno raggiunto performance allo stato dell'arte nella diagnosi dei guasti, spesso superando significativamente gli approcci basati su feature ingegnerizzate manualmente.

Se le CNN sono specializzate nell'analisi di dati con una struttura spaziale, le Reti Neurali Ricorrenti (*Recurrent Neural Networks* - RNN) sono l'architettura d'elezione per modellare dati intrinsecamente sequenziali, dove l'ordine e il contesto temporale sono fondamentali. A differenza delle reti feedforward (come le CNN o le reti dense), le RNN possiedono delle connessioni ricorsive, o "cicli", che permettono all'informazione di persistere e di essere passata da un passo temporale al successivo. Un neurone ricorrente non riceve solo l'input attuale della sequenza, $x(t)$, ma anche l'output (o lo stato nascosto) che ha prodotto al passo temporale precedente. Il suo output attuale, $h(t)$, è quindi una funzione sia di $x(t)$ che di $h(t-1)$. Questo crea una forma di memoria interna, che permette alla rete di tenere conto del contesto passato per processare l'input attuale. Questa caratteristica le rende lo strumento naturale e teoricamente più adatto per modellare serie temporali e, di conseguenza, per il problema della stima della RUL, che dipende intrinsecamente dalla storia passata del degrado.

Tuttavia, le RNN semplici soffrono di un grave problema pratico, noto come “svanire” o “esplodere” del gradiente (*vanishing/exploding gradient*). Durante l'addestramento tramite l'algoritmo di back propagation-through-time (BPTT), i gradienti dell'errore vengono propagati all'indietro nel tempo. A causa delle molteplici moltiplicazioni matriciali coinvolte nel ciclo ricorrente, per sequenze lunghe, questi gradienti possono diventare esponenzialmente piccoli (svanire) o esponenzialmente grandi (esplodere). Lo svanire del gradiente rende di fatto impossibile per la rete apprendere dipendenze a lungo raggio tra elementi distanti nella sequenza, limitando la sua memoria a un passato molto recente [9].

Per risolvere questo problema cruciale, sono state introdotte architetture ricorrenti più sofisticate, che oggi rappresentano lo stato dell'arte e hanno reso le RNN uno strumento pratico ed efficace. Le due più importanti sono le Long Short-Term Memory (LSTM), introdotte da Hochreiter e Schmidhuber (1997) [44], e le Gated Recurrent Unit (GRU), una

loro variante leggermente più semplice proposta da Cho et al. (2014) [21]. L'idea chiave di queste architetture è quella di dotare la cella ricorrente di un meccanismo di gating esplicito, ovvero delle "porte" (gate) neurali che, tramite piccole reti con attivazione sigmoide, imparano a controllare dinamicamente il flusso di informazioni. Una cella LSTM possiede una memoria interna separata (il *cell state*) e tre porte: una porta di input (che decide quali nuove informazioni sono importanti e devono essere aggiunte alla memoria), una porta di oblio (*forget gate*) (che decide quali vecchie informazioni non sono più rilevanti e devono essere scartate dalla memoria) e una porta di output (che decide quali informazioni della memoria utilizzare per calcolare l'output attuale). Questa architettura permette alla cella di mantenere informazioni rilevanti per lunghi periodi di tempo, proteggendole dal flusso di input irrilevanti e superando il problema dello svanire del gradiente. Le GRU semplificano questo meccanismo utilizzando solo due porte, che combinano le funzionalità di quelle dell'LSTM, spesso raggiungendo performance simili con una minore complessità computazionale.

Nel contesto della prognostica, le reti basate su LSTM e GRU sono diventate il modello di riferimento per la stima della RUL a partire da dati di sensori di serie temporali. Esse possono essere addestrate in modo end-to-end per prendere in input una sequenza di dati grezzi o di feature e predire il valore della RUL alla fine della sequenza. La loro capacità di apprendere dipendenze temporali complesse e a lungo raggio le rende teoricamente e praticamente superiori ai modelli statistici tradizionali e anche a molti modelli di ML classico in questo specifico e difficile compito.

In conclusione, il Deep Learning rappresenta una rottura paradigmatica nella prognostica data-driven. La sua capacità di eseguire il feature learning end-to-end e di modellare complesse dipendenze spaziali e temporali sta portando a un nuovo livello di accuratezza e di automazione. Tuttavia, questa immensa potenza ha un costo: i modelli di Deep Learning sono "affamati di dati" e richiedono enormi quantità di dati di addestramento, necessitano di una notevole potenza computazionale (spesso basata su GPU) e, soprattutto, la loro natura di "scatola nera" con milioni di parametri solleva significative e pressanti sfide di interpretabilità, fiducia e certificazione, che la ricerca nel campo dell'Explainable AI (XAI) sta attivamente cercando di affrontare [1].

### 4.2.4. Tabella di Sintesi Comparativa

Per schematizzare la complessa gamma di tecniche di Machine Learning e Deep Learning discusse in questa sezione, la seguente tabella (*Tabella 3*) fornisce una sintesi comparativa. La tabella riassume, per ogni paradigma e algoritmo chiave, il suo obiettivo primario nel contesto della Manutenzione Predittiva, i requisiti fondamentali, i principali punti di forza e le debolezze. Questo schema mira a fornire al lettore una "mappa concettuale" per navigare nel panorama degli strumenti di ML e per comprendere i trade-off associati alla scelta di ciascuno di essi.

| **Paradigma** | **Compito** | **Algoritmo / Architettura** | **Obiettivo nella PdM** | **Requisiti Fondamentali** | **Punti di Forza** | **Debolezze / Sfide** |
|---|---|---|---|---|---|---|
| **ML Classico (Supervisionato)** | Classificazione / Regressione | SVM / Random Forest | Diagnosi dei guasti o stima della RUL. | Dati etichettati, feature engineering manuale di alta qualità. | Elevata accuratezza, robustezza, interpretabilità parziale (RF), gestisce bene feature strutturate. | Performance limitata dalla qualità delle feature, "scatola nera" (SVM), meno efficace su dati grezzi. |
| **ML Classico (Non Supervisionato)** | Clustering / Anomaly Detection | k-Means / DBSCAN / Isolation Forest | Scoperta di stati operativi, rilevamento di anomalie incipienti. | Dati non etichettati, feature engineering manuale. | Utile in assenza di etichette, scoperta di pattern, allarme precoce. | Meno preciso della supervisione, la qualità dipende totalmente dalle feature fornite. |
| **Deep Learning (Supervisionato)** | Diagnosi / Classificazione | Rete Neurale Convoluzionale (CNN) | Diagnosi dei guasti da segnali grezzi (1D) o da loro rappresentazioni tempo-frequenza (2D). | Grandi quantità di dati etichettati, notevole potenza computazionale (GPU). | Feature learning automatico (end-to-end), performance allo stato dell'arte, invariante a traslazioni. | Richiede molti dati, "scatola nera" (difficile da interpretare), costo computazionale elevato. |

| | | | | | | |
|---|---|---|---|---|---|---|
| **Deep Learning (Supervisionato)** | Prognosi / Stima RUL | Rete Neurale Ricorrente (LSTM / GRU) | Stima della RUL da sequenze di dati di sensori. | Lunghe sequenze di dati etichettati, potenza computazionale. | Modella dipendenze temporali a lungo raggio, feature learning da sequenze, stato dell'arte per la prognosi. | "Fame di dati", addestramento lento, "scatola nera" molto complessa, rischio di overfitting. |
| **Deep Learning (Non Supervisionato)** | Anomaly Detection / Feature Learning | Autoencoder | Rilevamento di anomalie (basato sull'errore di ricostruzione di dati "sani"), o apprendimento di feature non supervisionato da usare in modelli a valle. | Grandi quantità di dati non etichettati (prevalentemente sani). | Apprende rappresentazioni compatte e non lineari dei dati, eccellente per l'anomaly detection su dati complessi. | Difficile da addestrare correttamente, l'interpretabilità delle feature apprese può essere problematica. |

*Tabella 3 – Sintesi Comparativa delle tecniche di Machine Learning*

## 4.3. Il Paradigma del Digital Twin: il Gemello Digitale del Processo Produttivo

Anche i modelli di Machine Learning più avanzati, se usati da soli, hanno un limite: operano in un "vuoto di contesto". Un algoritmo di Deep Learning può analizzare un segnale di vibrazione con un'efficacia straordinaria e classificarne lo stato di salute, ma non "sa" nulla del cuscinetto che ha generato quel segnale. Non conosce il suo materiale, la sua geometria, la sua posizione o le leggi fisiche che ne governano l'usura. I modelli di ML sono quindi strumenti analitici molto potenti, ma altamente specializzati. La frontiera più avanzata della prognostica moderna cerca di superare questa disconnessione, muovendosi verso un approccio più integrato: il Digital Twin.

Il Digital Twin non è un algoritmo in competizione con il Machine Learning. Al contrario, è un'architettura che si propone di integrare in un unico sistema il Machine Learning, i modelli basati sulla fisica, i dati in tempo reale e la conoscenza del processo. Il suo obiettivo

è creare un ponte continuo tra il mondo fisico e quello digitale, dando vita a un sistema in cui la macchina reale e la sua copia virtuale evolvono insieme, in costante dialogo [54].
Possiamo pensare a questa relazione in modo semplice: se il Machine Learning è il "cervello" analitico che impara dai dati, il Digital Twin è il "corpo" virtuale e il "sistema nervoso" che connette questo cervello al mondo reale, dandogli il contesto necessario per rendere le sue previsioni non solo accurate, ma anche significative e utilizzabili. In pratica, il passaggio al Digital Twin segna una transizione importante: si passa da una Manutenzione Predittiva che risponde alla domanda "cosa succederà a questo segnale?", a una gestione della salute degli asset che risponde a una domanda molto più complessa e di valore: "cosa significa questa previsione per la performance, la sicurezza e la redditività del mio intero sistema produttivo?".

### 4.3.1. Oltre la Simulazione: La Definizione di una Rappresentazione Vivente

Per capire a fondo il potenziale del Digital Twin, è importante prima di tutto definirlo con precisione e distinguerlo da concetti simili ma diversi, come i modelli CAD o le simulazioni. Sebbene il termine sia diventato popolare solo di recente con l'Industria 4.0, l'idea di usare modelli speculari per supportare sistemi complessi non è nuova. Risale, infatti, al programma Apollo della NASA degli anni '60, dove a terra si usavano repliche e simulatori per rispecchiare le condizioni delle navicelle nello spazio e testare soluzioni a problemi imprevisti. Tuttavia, a questo concetto mancava un elemento chiave che definisce il Digital Twin moderno: una connessione dati automatizzata e continua.
La formalizzazione moderna del concetto è attribuita a Michael Grieves, che nel 2002 propose un modello basato su tre parti: un prodotto fisico, una sua rappresentazione virtuale e un flusso di dati che li collega [38].
È proprio la natura di questa connessione a fare la differenza. Un Digital Twin non è un modello statico, come una "fotografia" digitale, né una simulazione del comportamento atteso. Al contrario, è una rappresentazione dinamica e "vivente" della sua controparte fisica. La sua caratteristica fondamentale è un flusso di dati automatizzato che lo lega in tempo quasi reale all'asset fisico. Come definito dalla NASA, un Digital Twin è un sistema integrato che combina un modello complesso con i dati in tempo reale del veicolo fisico, per rispecchiare la sua intera vita [33]. È questa connessione a trasformare il modello da un'entità

passiva a una attiva, un vero e proprio "alter ego" digitale che evolve e si degrada in sincronia con il suo gemello fisico. Per chiarire questa distinzione, in letteratura si parla a volte di *Digital Model* (senza connessione dati automatica), *Digital Shadow* (con flusso dati in una sola direzione, dal fisico al virtuale) e *Digital Twin* (con flusso dati bidirezionale) [54]. Questa definizione implica che un Digital Twin non è un modello monolitico, ma un ecosistema complesso e stratificato di modelli, un "modello di modelli" o un *system-of-systems*, che deve integrare diverse forme di conoscenza e rappresentazione per raggiungere un'alta fedeltà:

- Rappresentazione Geometrica: È la base visiva del Digital Twin, tipicamente derivata da file CAD 3D, ma arricchita da informazioni sul Product Manufacturing Information (PMI) come tolleranze e finiture superficiali. Può essere ulteriormente dettagliata da dati di scansioni laser o fotogrammetria per catturare la geometria "as-built" dell'asset, con le sue imperfezioni, invece di quella "as-designed" idealizzata.

- Rappresentazione Comportamentale Basata sulla Fisica: Questa componente, che sfrutta i modelli Physics-of-Failure (PoF) discussi nel capitolo precedente, ne simula il comportamento secondo le leggi fondamentali della fisica. Modelli agli elementi finiti (FEM) per l'analisi strutturale e termica, o di fluidodinamica computazionale (CFD) per l'analisi dei flussi, permettono di prevedere come l'asset risponderà a determinate sollecitazioni, di calcolare la distribuzione degli sforzi interni e di simulare meccanismi di degrado noti.

- Rappresentazione Comportamentale Guidata dai Dati: Questa è la componente in cui risiedono gli algoritmi di Machine Learning e Deep Learning. Essi agiscono laddove la fisica è sconosciuta, troppo complessa da modellare o computazionalmente proibitiva. Apprendono pattern, anomalie e comportamenti complessi direttamente dai dati dei sensori, fornendo capacità diagnostiche e prognostiche che sono complementari e sinergiche a quelle dei modelli fisici.

- Rappresentazione Logica e Operativa: Questa componente, spesso trascurata ma cruciale, modella la logica di funzionamento del sistema (es. il programma PLC di una macchina utensile), le regole di business che governano un processo produttivo, o persino i modelli del comportamento umano degli operatori. La sua inclusione

permette di simulare non solo il comportamento fisico, ma anche quello logico e operativo del sistema, e di analizzare le interazioni tra di essi.

La capacità di integrare queste diverse forme di conoscenza — fisica, geometrica, basata sui dati e logica — in un unico quadro interattivo è ciò che conferisce al Digital Twin la sua straordinaria potenza rappresentativa e predittiva, ben al di là di qualsiasi modello di simulazione tradizionale [49].

Un esempio emblematico di questa transizione è visibile nel settore aerospaziale. Mentre in passato si eseguivano simulazioni FEM per calcolare la vita a fatica di un'ala di aereo in condizioni di carico standard, oggi aziende come Airbus o Boeing lavorano per creare un Digital Twin per ogni singolo velivolo. Questo gemello digitale non si basa su dati generici, ma viene costantemente aggiornato con i dati reali di volo di quell'aereo specifico: le ore di volo, i cicli di pressurizzazione, le turbolenze incontrate, le temperature a cui è stato esposto. Il Digital Twin, quindi, non calcola una vita residua teorica, ma la vita residua effettiva di quella specifica ala, riflettendo la sua storia unica. Questo permette di passare da ispezioni basate su un calendario statico a ispezioni ottimizzate sulla base della reale condizione di usura di ogni singolo aereo della flotta.

In sintesi, mentre una simulazione risponde alla domanda "Cosa succederebbe a un modello generico e idealizzato se applicassimo certe condizioni?", un Digital Twin mira a rispondere alla domanda molto più complessa e di valore: "Cosa sta succedendo adesso al mio asset fisico specifico, con la sua storia unica e le sue imperfezioni, e cosa gli succederà nel futuro, data la sua condizione attuale e le probabili condizioni operative?". Questa transizione dalla simulazione ipotetica e generica alla rappresentazione vivente, co-evolutiva e individualizzata è il vero cuore della rivoluzione concettuale e tecnologica del Digital Twin.

### 4.3.2. L'Architettura Fondamentale e la Dinamica del Flusso Circolare delle Informazioni

Per tradurre la visione concettuale del Digital Twin da un'idea affascinante a un sistema ingegneristico funzionante, è necessario definirne un'architettura robusta. Sebbene le implementazioni specifiche possano variare notevolmente a seconda del dominio applicativo, della scala e della maturità tecnologica, la letteratura scientifica e le best practice industriali convergono su un modello fondamentale basato su tre pilastri essenziali. La loro

interazione dinamica non crea semplicemente un modello digitale, ma un vero e proprio sistema ciber-fisico, un'entità ibrida che vive a cavallo tra il mondo degli atomi e quello dei bit. Questi pilastri, seguendo il modello concettuale originale di Grieves e le successive elaborazioni [39] [111], sono: l'entità fisica, l'entità virtuale e la connessione dati che li unisce in un legame indissolubile.

Il primo pilastro, il punto di partenza e il referente ultimo di tutto il sistema, è l'Entità Fisica (*Physical Twin*). Questo è l'asset reale che opera nel mondo fisico: il macchinario, il robot, la linea di produzione o persino l'intera fabbrica. Nell'era del Digital Twin, tuttavia, questo asset cessa di essere un oggetto passivo, "muto" e isolato, per trasformarsi in un'entità "aumentata", intelligente e, soprattutto, "loquace". Questa metamorfosi è resa possibile da una rete di sensori, che costituisce il livello fisico dell'Internet of Things (IoT). Questa rete agisce come un sistema nervoso sensoriale artificiale, avvolgendo l'asset fisico e catturandone i parametri operativi vitali e lo stato di salute con un'elevata frequenza.

La scelta e il posizionamento di questi sensori sono un'attività di progettazione critica. Si spazia da sensori tradizionali e consolidati, come accelerometri triassiali per misurare le vibrazioni, termocoppie o termocamere a infrarossi per monitorare le mappe termiche, sensori di pressione e di flusso per i circuiti idraulici e di lubrificazione, ed encoder e resolver per misurare con precisione posizione e velocità angolare, fino a sensori più esotici e specifici, come microfoni ad alta frequenza per l'analisi delle emissioni acustiche o sensori di corrente e tensione per il monitoraggio della firma elettrica dei motori (*Motor Current Signature Analysis* - MCSA). L'obiettivo è quello di creare un "flusso di coscienza" digitale, un flusso continuo di dati che racconti la "vita" dell'asset, momento per momento.

Oltre ai sensori, che rappresentano i "sensi" del sistema, l'entità fisica è dotata di attuatori (motori, valvole, pistoni, riscaldatori, bracci robotici) che ne costituiscono i "muscoli", permettendole di modificare il proprio comportamento e di agire sul mondo fisico. La caratteristica cruciale che abilita il Digital Twin è che sia i sensori (*input*) che gli attuatori (*output*) sono connessi in rete, capaci di inviare dati e di ricevere comandi tramite protocolli di comunicazione industriale standard. L'entità fisica, quindi, cessa di essere un sistema meccanico isolato per diventare un nodo attivo e interattivo di una rete ciber-fisica.

Il secondo e più complesso pilastro è l'Entità Virtuale (*Virtual Twin*). Questa è la controparte digitale dell'asset fisico, che risiede in un ambiente computazionale, sia esso un server on-premise ad alte prestazioni o, sempre più spesso, una piattaforma cloud scalabile che offre servizi di calcolo, storage e analytics. È qui che risiede la maggior parte dell'intelligenza, della logica e della capacità predittiva del sistema. È un errore concettuale grave ridurre

l'entità virtuale a un semplice modello 3D; essa è, piuttosto, una rappresentazione multi-dominio, multi-scala e, idealmente, probabilistica ad alta fedeltà.

"Multi-dominio" (o multi-fisica) si riferisce alla sua capacità di integrare e orchestrare modelli che descrivono diversi domini della fisica e della logica. Un Digital Twin di una macchina utensile, ad esempio, deve integrare un modello meccanico FEM per la dinamica strutturale e le deformazioni sotto carico, un modello termico per la dilatazione e la dissipazione del calore generato dal taglio, un modello fluidodinamico per il sistema di lubrificazione, un modello elettrico per il comportamento dei motori e dei drive, e un modello di controllo che emula la logica del PLC. La vera sfida, e il vero valore, risiede nel catturare le complesse interazioni e i coupling tra questi domini (ad esempio, come la deformazione termica influenzi la precisione geometrica).

"Multi-scala" significa che il Digital Twin può rappresentare il sistema a diversi livelli di granularità e astrazione, permettendo di "zoomare" dentro e fuori a seconda delle necessità dell'analisi. Può spaziare dal comportamento microstrutturale di un materiale in un punto critico (es. simulando la nucleazione di una cricca), al comportamento di un singolo componente (un cuscinetto), al comportamento dell'intera macchina (le sue vibrazioni modali e la sua cinematica), fino al suo inserimento in una linea di produzione (i flussi, i buffer e le interdipendenze con altre macchine) e persino all'interno dell'intera supply chain [31].

Infine, una rappresentazione matura del Digital Twin deve essere probabilistica. Riconoscendo l'incertezza dei modelli, delle misure e delle condizioni future, l'entità virtuale non dovrebbe produrre previsioni deterministiche, ma distribuzioni di probabilità, quantificando la confidenza associata a ogni stima. L'entità virtuale è quindi il "terreno di gioco" computazionale, la "sandbox" dove vengono addestrati i modelli di Machine Learning, dove vengono eseguite le complesse simulazioni agli elementi finiti, e dove vengono testate e validate le politiche di manutenzione in un ambiente virtuale privo di rischi e di costi.

Il terzo e più cruciale pilastro, quello che dà vita al Digital Twin e lo distingue da qualsiasi forma di modellazione offline, è la Connessione Dati. Questo è il "filo digitale" (*digital thread*), un'infrastruttura di comunicazione robusta e ad alte prestazioni che lega indissolubilmente il mondo fisico e quello virtuale, garantendo la loro coerenza e sincronia. Questo flusso di dati deve essere persistente, affidabile e, soprattutto, bidirezionale.

Il flusso dal Fisico al Virtuale è il flusso di dati afferente, il percorso sensoriale del sistema. I dati grezzi raccolti dai sensori vengono trasmessi, spesso tramite protocolli industriali

standardizzati e interoperabili come OPC-UA (*Open Platform Communications Unified Architecture*), MQTT (*Message Queuing Telemetry Transport*) o DDS (*Data Distribution Service*), all'entità virtuale. Questo flusso di dati ha un duplice e fondamentale scopo. In primo luogo, serve a sincronizzare costantemente lo stato del modello virtuale con lo stato istantaneo del modello reale. In secondo luogo, e in modo più sofisticato, serve a calibrare, correggere e validare il modello nel tempo. Nessun modello, per quanto complesso, è una rappresentazione perfetta della realtà. Esisterà sempre una discrepanza tra le previsioni del modello e le osservazioni reali. I dati reali vengono utilizzati per aggiornare i parametri incerti del modello (sia esso fisico o data-driven) e per correggerne le deviazioni, in un processo noto in letteratura come *model updating*, *data assimilation* o, in un contesto bayesiano, *belief updating*. Tecniche di filtraggio avanzate come il Filtro di Kalman Esteso (EKF) o il Filtro Particellare (*Particle Filter*) sono gli strumenti matematici ideali per realizzare questa fusione ricorsiva tra la previsione del modello (la conoscenza a priori) e l'evidenza dei dati (la verosimiglianza), producendo una stima dello stato più accurata e con un'incertezza ridotta [95].
Il flusso dal Virtuale al Fisico è il flusso di dati efferente, il percorso motorio del sistema, quello che chiude il cerchio e genera valore operativo. Le informazioni, le previsioni, le diagnosi e le decisioni generate nell'ambiente virtuale vengono ritrasmesse al mondo fisico per guidarne il comportamento. Questo può avvenire con diversi livelli di autonomia e maturità, come definito da alcuni framework di classificazione [54]:

- Livello Informativo (*Digital Model/Shadow*): Il Digital Twin fornisce *insight*, visualizzazioni e analisi agli operatori e ai manager attraverso dashboard, report o interfacce in Realtà Aumentata (AR) e Virtuale (VR), supportando le loro decisioni in modo più informato.

- Livello Consultivo (*Digital Twin parziale*): Il sistema genera allarmi predittivi e raccomandazioni di azione specifiche (es. "Si raccomanda la sostituzione del cuscinetto X entro le prossime 48 ore di funzionamento con una probabilità di guasto del 95%"). L'operatore umano mantiene la responsabilità finale della decisione.

- Livello Autonomo (*Digital Twin completo*): Nel suo stadio più avanzato e visionario, il Digital Twin invia comandi diretti agli attuatori dell'asset fisico per ottimizzare o correggere il processo in tempo reale, senza intervento umano. Ad esempio, può

regolare dinamicamente i parametri di taglio di una macchina utensile per compensare l'usura dell'utensile, o può ridurre il regime operativo di un motore per estenderne la vita residua fino al prossimo fermo macchina programmato, in un'ottica di controllo predittivo.

Nel settore automotive, aziende come Tesla utilizzano un'architettura simile a un Digital Twin per la gestione della flotta. Ogni veicolo (l'entità fisica) è equipaggiato con centinaia di sensori che trasmettono costantemente dati operativi a un server centrale (dove risiede l'entità virtuale). Questo flusso di dati dal fisico al virtuale non solo permette a Tesla di monitorare lo stato di salute della batteria di ogni singola auto, ma anche di raccogliere dati su come diversi stili di guida e condizioni climatiche influenzino il suo degrado. L'aspetto più innovativo è il flusso dal virtuale al fisico: sulla base delle analisi condotte su questi dati aggregati, Tesla può rilasciare aggiornamenti software "over-the-air" che modificano gli algoritmi di gestione della batteria, ottimizzandone la performance e la durata sull'intera flotta. In questo caso, il Digital Twin non solo prevede, ma agisce attivamente per migliorare il comportamento dell'asset fisico.

È questo ciclo continuo e ricorsivo di dati reali che arricchiscono e correggono il modello virtuale, e di informazioni e decisioni elaborate dal modello virtuale che guidano e ottimizzano il mondo fisico, che costituisce l'essenza operativa e il potere trasformativo del paradigma del Digital Twin.

Un esempio industriale emblematico, che traduce in pratica questa visione, è offerto da Siemens con la sua gestione della flotta di turbine a gas industriali attraverso la piattaforma MindSphere. Per ogni singola turbina fisica installata, Siemens mantiene un gemello digitale unico che non è una semplice replica 3D, ma un ecosistema di modelli ibridi ad alta fedeltà. Questo gemello virtuale integra la conoscenza fisica (modelli termodinamici e di fatica dei materiali) con modelli analitici addestrati su decenni di dati storici.

Il vero valore emerge dalla connessione dati continua: centinaia di sensori sulla turbina reale trasmettono un flusso di dati che aggiorna e calibra costantemente il gemello digitale, rendendolo un ritratto fedele e "vivente" della sua controparte fisica. Grazie a questa sincronia, Siemens può eseguire simulazioni accelerate per prevedere, con mesi di anticipo, la probabilità di guasto di un componente specifico su una turbina in qualsiasi parte del mondo, tenendo conto del suo profilo di carico unico.

Il ciclo si chiude con il passaggio dalla previsione alla prescrizione. L'operatore dell'impianto non riceve un semplice allarme, ma una raccomandazione ottimizzata e

contestualizzata, del tipo: "Prescrizione: Si prevede un aumento del rischio di guasto del componente X tra 400 ore. Si consiglia di ridurre il carico del 5% durante i picchi per estendere la vita residua di 300 ore e di pianificare la sostituzione durante il prossimo fermo programmato tra 6 settimane". In questo modo, Siemens non vende più solo turbine, ma offre ore di funzionamento affidabile, dimostrando come il Digital Twin trasformi non solo la manutenzione, ma l'intero modello di business.

### 4.3.3. Il Ruolo del Digital Twin nella Manutenzione Predittiva

L'architettura del Digital Twin, basata sui suoi tre pilastri (fisico, virtuale e connesso), non è statica, ma è la base su cui si svolge un processo dinamico e continuo: il flusso circolare delle informazioni. Questo ciclo è il cuore del sistema: è il processo attraverso cui i dati grezzi vengono trasformati in conoscenza, e la conoscenza viene a sua volta usata per guidare azioni nel mondo fisico.

Capire come funziona questo flusso è fondamentale per comprendere perché il Digital Twin è molto più di un semplice sistema di monitoraggio. Possiamo vederlo come un'evoluzione del classico ciclo "Percezione-Azione", tipico della robotica, a cui viene aggiunta una fase fondamentale di ragionamento e previsione del futuro. Questo ciclo si ripete costantemente durante la vita della macchina e può essere suddiviso in una serie di fasi.

Il ciclo ha inizio nel mondo fisico, il "Gemba" della terminologia Lean. L'entità fisica, attraverso la sua rete di sensori IoT, agisce come un organo di senso, acquisendo dati grezzi sul suo stato operativo e sulle condizioni ambientali. Questi dati, che costituiscono la percezione del sistema, vengono trasmessi attraverso il digital thread all'entità virtuale. Questo primo passaggio, tuttavia, va ben oltre una semplice registrazione o un data logging. I dati in arrivo innescano la fase cruciale della sincronizzazione, in cui il gemello virtuale, che fino a un istante prima era una rappresentazione "passata" del sistema, viene aggiornato e allineato con la realtà attuale. Questo processo di sincronizzazione non è un semplice aggiornamento di parametri, ma un'attività complessa che può richiedere l'aggiornamento simultaneo di modelli eterogenei: il modello geometrico potrebbe essere aggiornato per riflettere l'usura, il modello termico per riflettere la distribuzione attuale delle temperature, e i modelli di ML per registrare le ultime osservazioni.

Più profondamente, i dati reali vengono utilizzati per un processo continuo di calibrazione e correzione del modello, un'attività nota in letteratura con i termini di data

assimilation o model updating. Nessun modello, per quanto sofisticato, può catturare perfettamente la complessità stocastica e le incertezze del mondo reale. Esisterà sempre una discrepanza, un "errore residuo", tra le previsioni del modello e le osservazioni reali. Il flusso di dati dal mondo fisico viene utilizzato per ridurre questa discrepanza, aggiornando i parametri incerti del modello (siano essi parametri fisici come i coefficienti di attrito, o i pesi di una rete neurale) in modo che le sue previsioni si allineino meglio con l'evidenza empirica. Tecniche di filtraggio bayesiano, come il Filtro di Kalman e le sue estensioni non lineari (EKF, UKF), e in particolare il Filtro Particellare (*Particle Filter*), sono gli strumenti matematici d'elezione per realizzare questa fusione in modo rigoroso e probabilistico. Essi permettono di combinare la conoscenza a priorincon la verosimiglianza fornita dai dati osservati (la *likelihood*), per produrre una stima dello stato "posteriore" (*posterior*) che è statisticamente più accurata e ha un'incertezza ridotta rispetto alle due fonti prese singolarmente [95]. Questa fase assicura che il Digital Twin non sia un modello generico e idealizzato, ma un'entità individualizzata e contestualizzata, un ritratto fedele e costantemente aggiornato della sua specifica controparte fisica, con la sua storia unica, le sue imperfezioni e le sue peculiarità.

Una volta che il gemello virtuale è stato sincronizzato per rispecchiare fedelmente lo stato attuale del gemello fisico, inizia la fase più potente, distintiva e di maggior valore del paradigma: l'esplorazione del futuro. L'entità virtuale viene, in un certo senso, "sganciata" dal presente e proiettata nel futuro per rispondere non solo alla domanda "cosa succederà?", ma anche alla domanda più profonda e strategica: "What if?" (Cosa succederebbe se?). Utilizzando i modelli ad alta fedeltà che lo compongono, è possibile eseguire un gran numero di simulazioni di scenari futuri a una velocità molto superiore a quella del tempo reale, un processo che può essere visto come una forma di *accelerated life testing* virtuale o di esplorazione controfattuale.

Questa capacità di simulazione apre possibilità prognostiche che vanno ben oltre la semplice estrapolazione statistica dei trend passati. È possibile, ad esempio, eseguire analisi complesse e stratificate:

- Prognosi sotto condizioni nominali: Si simula l'evoluzione futura del degrado assumendo che il macchinario continui a operare con il profilo di carico e le condizioni ambientali medie osservate nel passato, per ottenere una stima della RUL di base.

- Prognosi sotto stress e scenari alternativi: Si simula l'impatto di condizioni operative future diverse e potenzialmente più severe. Cosa succede alla RUL se il carico di lavoro aumenta del 20% per il prossimo turno? E se la temperatura ambiente dovesse salire di 5 gradi? E se si utilizzasse un lotto di materiale con una durezza leggermente superiore? Il Digital Twin permette di rispondere a queste domande in modo quantitativo.

- Analisi di sensibilità e ottimizzazione: Si possono identificare quali parametri operativi (es. velocità di taglio, avanzamento) hanno l'impatto più significativo sulla vita utile del componente, aprendo la strada a strategie di controllo che ottimizzino questo trade-off tra produttività e affidabilità.

L'integrazione di modelli fisici e data-driven è particolarmente potente in questa fase. Un modello fisico può essere utilizzato per simulare la traiettoria di degrado di base, mentre un modello di Machine Learning, addestrato su dati storici, può essere utilizzato per correggere questa traiettoria, tenendo conto di effetti complessi, non modellati e specifici di quella macchina. Inoltre, sfruttando la natura probabilistica di molti di questi modelli, è possibile eseguire non una singola simulazione deterministica, ma migliaia di simulazioni Monte Carlo. In ogni simulazione, i parametri incerti del modello e le future condizioni operative vengono campionati dalle loro rispettive distribuzioni di probabilità. Il risultato non è una singola stima puntuale e fragile della RUL, ma un'intera distribuzione di probabilità della RUL, che fornisce una ricca e robusta valutazione del rischio di guasto nel tempo, completa di intervalli di confidenza [51]. Questa capacità di esplorare e quantificare l'incertezza futura è una delle caratteristiche più preziose e distintive del Digital Twin per una gestione proattiva e basata sul rischio.

Sulla base delle previsioni, delle simulazioni e delle analisi condotte nell'ambiente virtuale, il ciclo si chiude tornando al mondo fisico attraverso la fase di decisione e azione. Le informazioni elaborate dal Digital Twin devono essere tradotte in interventi concreti che generino valore economico e operativo. Questo "ritorno al reale" può avvenire con diversi livelli di autonomia, rappresentando un vero e proprio percorso di maturità del sistema.

Al livello più semplice, il Digital Twin agisce come un sistema di supporto alle decisioni (*Decision Support System* - DSS). Se le simulazioni prevedono l'insorgere di un guasto o di una deriva critica del processo, il sistema può generare un allarme predittivo e una raccomandazione di manutenzione per l'operatore o il responsabile dell'impianto. Queste

raccomandazioni non sono semplici alert, ma possono essere arricchite con informazioni contestuali di grande valore: la stima della RUL con il suo intervallo di confidenza, la probabilità di guasto nei prossimi turni, l'identificazione della causa radice più probabile, e persino le istruzioni operative per l'intervento, magari visualizzate tramite un'interfaccia in Realtà Aumentata (AR) che sovrappone le informazioni digitali direttamente sull'asset fisico, guidando l'operatore passo dopo passo [111].

A un livello più avanzato, il sistema passa da una logica predittiva a una prescrittiva. Oltre a prevedere cosa succederà, il Digital Twin può esplorare diverse possibili azioni correttive (es. "ridurre la velocità di taglio del 10%", "aumentare la frequenza di lubrificazione", "pianificare un fermo macchina tra 24 ore") e simularne l'impatto sulla RUL, sulla qualità del prodotto e sulla performance produttiva. Utilizzando algoritmi di ottimizzazione, può quindi raccomandare l'azione o la sequenza di azioni che offre il miglior compromesso tra costi, rischi e benefici, fornendo una vera e propria prescrizione operativa.

Infine, nello stadio più avanzato e visionario, quello del controllo ciber-fisico a ciclo chiuso, il Digital Twin esercita un'azione autonoma. Se la previsione indica un'imminente uscita dalle specifiche di qualità o un rischio di guasto inaccettabile, il Digital Twin può inviare un comando di correzione direttamente al controllore (PLC) della macchina reale, senza intervento umano. L'esempio citato nei colloqui iniziali è paradigmatico di questo livello di integrazione: il Digital Twin di una macchina di tornitura, prevedendo la rottura imminente dell'utensile, invia autonomamente un comando alla macchina reale per fermare il processo e richiedere un cambio utensile. Questo approccio, noto come controllo predittivo, rappresenta la massima espressione dell'autonomia industriale [90].

È questo ciclo continuo, ricorsivo e virtuoso di dati reali che arricchiscono e correggono il modello virtuale, e di informazioni elaborate, previsioni e comandi che guidano e ottimizzano il mondo fisico, che costituisce l'essenza operativa e il potere trasformativo del paradigma del Digital Twin. È questo ciclo che trasforma un insieme di modelli e dati in un sistema cognitivo completo, capace di percepire, ragionare, prevedere e agire, realizzando pienamente la promessa di una fabbrica veramente intelligente.

### 4.3.4. Sinergia con il Machine Learning e Sfide Future

L'adozione del Digital Twin nella Manutenzione Predittiva (PdM) non è un semplice miglioramento, ma un vero e proprio salto di qualità. La sua architettura non è solo un

"contenitore" per i modelli di Machine Learning, ma agisce come un catalizzatore che ne amplifica il valore. Il Digital Twin fornisce il contesto e il ciclo di feedback necessari per trasformare la prognostica da un'analisi isolata a una gestione integrata e, in ultima analisi, prescrittiva. Vediamo in dettaglio i ruoli che svolge.

Una delle sfide principali del Machine Learning puro è che opera in un "vuoto semantico". Un algoritmo può analizzare un segnale di vibrazione e classificarlo, ma non "sa" nulla del sistema fisico che lo ha generato. Il Digital Twin risolve questo problema fornendo un contesto fisico e operativo ai dati. Quando un segnale viene analizzato all'interno di un Digital Twin, il sistema sa che quella vibrazione proviene da un cuscinetto specifico, con un certo materiale, montato in una certa posizione, che sta lavorando a una certa velocità. Questo ricco contesto, che integra dati dal CAD, dai modelli di simulazione e dal sistema di gestione della produzione (MES), permette di arricchire l'analisi. Si può così passare da una semplice correlazione statistica ("questo pattern è correlato a un guasto") a una comprensione della causalità fisica ("questo pattern è causato da un difetto sulla pista esterna, che peggiora quando la macchina lavora ad alta velocità"). Questa capacità di contestualizzare i risultati è fondamentale per migliorare l'interpretabilità dei modelli e la fiducia degli utenti [17].

Inoltre, il Digital Twin è l'ambiente ideale per implementare gli approcci ibridi, che fondono modelli basati sulla fisica (PoF) e modelli guidati dai dati. Questi approcci rappresentano la frontiera più promettente della prognostica. Ad esempio, un modello fisico (come un FEM che simula la fatica) può essere usato per fornire la traiettoria di degrado di base, mentre un modello di ML, analizzando i dati dei sensori in tempo reale, corregge e calibra continuamente questa previsione, tenendo conto delle specificità del singolo asset e delle reali condizioni operative.

La capacità di simulazione del Digital Twin è un altro strumento prognostico di per sé, che supera uno dei limiti principali dei modelli puramente basati sui dati. Un modello di ML può solo estrapolare dai pattern che ha già visto; non può prevedere in modo affidabile il comportamento del sistema in condizioni operative nuove. Il Digital Twin, grazie ai suoi modelli fisici, può invece rispondere a domande "what-if". Permette di esplorare l'impatto di diverse strategie o di futuri profili di missione sulla RUL, una capacità che i modelli data-driven da soli non possono fornire, consentendo di valutare la resilienza dell'asset a scenari alternativi [3].

Un caso d'uso potente si trova nel settore dell'energia eolica. Il gestore di un parco eolico può utilizzare il Digital Twin di una turbina per simulare l'impatto di diverse strategie di

controllo. Ad esempio, può porsi la domanda: "Se prevediamo venti molto forti per la prossima settimana, è più redditizio far funzionare la turbina al massimo della sua capacità, accettando un maggiore stress strutturale e una riduzione della sua vita utile, oppure è meglio limitarne leggermente la potenza (*derating*) per preservare i componenti critici come il riduttore e i cuscinetti?". Il Digital Twin, integrando un modello aerodinamico, un modello strutturale e un modello di degrado, può quantificare il guadagno economico e la "spesa" in termini di vita residua per ogni scenario, permettendo al gestore di prendere una decisione informata e basata sui dati.

Una delle sfide più grandi e persistenti nell'applicazione del Machine Learning supervisionato, e in particolare del Deep Learning, è la sua "fame" di dati. Esso richiede grandi quantità di dati di addestramento etichettati, che includano numerosi esempi di tutte le classi di interesse. In molti contesti industriali, i dati relativi ai guasti sono, fortunatamente, estremamente rari. Questo problema, noto come problema delle classi sbilanciate o della scarsità di dati di guasto, limita fortemente la capacità dei modelli di apprendere a riconoscere le condizioni di anomalia, portando a modelli che sono eccellenti nel riconoscere lo stato "sano" ma pessimi nel diagnosticare i guasti.

Il Digital Twin offre una soluzione elegante a questo problema attraverso la generazione di dati sintetici (*synthetic data generation*). Un Digital Twin ad alta fedeltà, validato su dati reali, può essere utilizzato come un "generatore di realtà virtuale" per simulare una vasta gamma di scenari di guasto, anche quelli più rari, più estremi o più pericolosi, in un ambiente virtuale sicuro e a basso costo.

Nel campo della robotica industriale, i guasti catastrofici di un braccio robotico sono eventi estremamente rari, rendendo quasi impossibile addestrare un modello analitico a riconoscerli. Un'azienda come KUKA o ABB può utilizzare un Digital Twin ad alta fedeltà del proprio robot, che include un modello dinamico e cinematico accurato, per generare dati sintetici. Si possono simulare migliaia di scenari di guasto (es. un guasto a un giunto, un problema al motore, un'usura eccessiva del riduttore) e registrare la "risposta" virtuale dei sensori di coppia e di posizione.

È possibile iniettare virtualmente guasti di diversa natura e gravità (es. diversi livelli di usura, diverse dimensioni di cricca) e simulare la risposta dei sensori, generando così grandi quantità di dati sintetici, ma fisicamente realistici. Questi dati sintetici possono poi essere utilizzati per aumentare il dataset reale, arricchendolo con esempi di guasto e bilanciando la distribuzione delle classi. Questo processo di *data augmentation* basato sulla simulazione si è dimostrato estremamente efficace per migliorare la robustezza, l'accuratezza e la capacità

di generalizzazione dei modelli di ML, specialmente per le reti neurali profonde, che beneficiano enormemente di dataset più grandi e diversificati [51] [65].

Forse il contributo più trasformativo del Digital Twin è la sua capacità di spostare l'orizzonte della manutenzione da una logica puramente predittiva (che risponde alla domanda "cosa succederà e quando?") a una logica prescrittiva (che risponde alla domanda "dato ciò che succederà, qual è la migliore azione da intraprendere ora?"). Avendo a disposizione un modello olistico del sistema e la capacità di simulare scenari futuri, il Digital Twin diventa uno strumento per l'ottimizzazione globale e multi-obiettivo.

Un'applicazione pratica di manutenzione prescrittiva si trova nel settore dei beni di consumo ad alta velocità, ad esempio in un impianto di imbottigliamento di Coca-Cola o PepsiCo. Il Digital Twin di una macchina etichettatrice potrebbe prevedere, sulla base del degrado di un attuatore, che la probabilità di applicare etichette storte aumenterà del 15% nelle prossime 24 ore. Invece di limitarsi a segnalare un allarme, il sistema prescrittivo potrebbe analizzare il piano di produzione e raccomandare un'azione ottimizzata: "Prescrizione: Continuare la produzione del lotto attuale a velocità ridotta del 5% per mantenere la qualità, e pianificare un intervento di manutenzione di 15 minuti durante il cambio formato previsto tra 4 ore, che è il momento a minor impatto produttivo". In questo modo, la decisione non è solo basata sulla condizione della macchina, ma è ottimizzata rispetto agli obiettivi di produzione e ai costi complessivi.

Invece di generare un semplice allarme, il sistema può esplorare lo spazio delle decisioni possibili. Prima di eseguire un intervento di manutenzione costoso, che implica il fermo della produzione e l'impiego di risorse, è possibile utilizzare il Digital Twin per valutarne l'impatto complessivo. Attraverso la simulazione, si possono confrontare diverse strategie, valutandole non solo in termini di affidabilità, ma anche di costo, qualità e produttività. Si possono formalizzare problemi di ottimizzazione complessi, come: "Trovare la politica di manutenzione che minimizza il costo totale del ciclo di vita, dato un vincolo sulla probabilità di guasto".

Questa ottimizzazione può estendersi oltre la manutenzione, per includere le operazioni in tempo reale. Se il Digital Twin prevede una RUL ridotta per un utensile, invece di fermare semplicemente la macchina, potrebbe prescrivere una modifica dei parametri di taglio (es. ridurre la velocità di avanzamento) per estendere la vita dell'utensile fino al prossimo fermo macchina già programmato, massimizzando così l'uptime complessivo. Questa capacità di prendere decisioni che bilanciano dinamicamente e in tempo reale i trade-off tra affidabilità, performance, qualità e costi, basandosi su una comprensione predittiva del sistema, è il cuore

della manutenzione prescrittiva, considerata il livello più alto di maturità nella gestione degli asset [106] [99].

In sintesi, il Digital Twin agisce come un ecosistema ciber-fisico che eleva la Manutenzione Predittiva a un nuovo e superiore livello di intelligenza e integrazione. Fornisce contesto e significato ai dati, abilita una prognostica ibrida, robusta e probabilistica, risolve il problema cronico della scarsità dei dati attraverso la simulazione, e, soprattutto, trasforma la previsione in prescrizione ottimizzata. È il framework che permette di realizzare pienamente la promessa di una gestione degli asset industriali che non sia solo proattiva, ma veramente olistica, adattiva e basata sul valore.

# Capitolo 5: Tassonomia e Analisi Comparativa della Letteratura

## 5.1. Metodologia di Classificazione (Tassonomia)

Una volta completata la selezione degli articoli, il risultato è un insieme di studi pertinenti ma ancora non organizzati (*Figura 3*). Per analizzarli in modo sistematico, è necessario classificarli. A questo scopo, ho sviluppato una tassonomia, che è molto più di una semplice tabella: è uno schema di classificazione che permette di "smontare" ogni articolo nei suoi elementi chiave. Questo processo non è fine a se stesso, ma è uno strumento fondamentale per dare ordine alla complessità, confrontare i diversi approcci e identificare i trend e le lacune nella ricerca [6]. La costruzione di una tassonomia è una pratica standard nelle revisioni sistematiche, perché fornisce la struttura necessaria per una sintesi chiara e replicabile [118] [52].

Per sviluppare la tassonomia, ho seguito un approccio ibrido, che combina un metodo "dall'alto" (deduttivo) e uno "dal basso" (induttivo), come raccomandato da Nickerson, Varshney, e Muntermann [76].

L'approccio deduttivo (*top-down*) mi ha guidato nella definizione delle dimensioni principali della classificazione. Queste macrocategorie, come la caratterizzazione metodologica e quella applicativa, non sono state scelte a caso, ma derivano direttamente dalle domande di ricerca (RQ e SQ) definite nel Capitolo 1. Questo assicura che la classificazione sia finalizzata a raccogliere le informazioni necessarie per rispondere a quelle domande.

L'approccio induttivo (*bottom-up*), invece, l'ho usato per definire le categorie più specifiche all'interno di ogni dimensione. Ho letto attentamente un gruppo di articoli rappresentativi per capire quali terminologie e tecniche venivano usate più di frequente, e ho costruito un insieme di categorie che fosse completo e senza sovrapposizioni.

Le dimensioni principali sono state raggruppate in tre macrocategorie: la caratterizzazione metodologica, quella del contesto applicativo e quella del contributo scientifico.

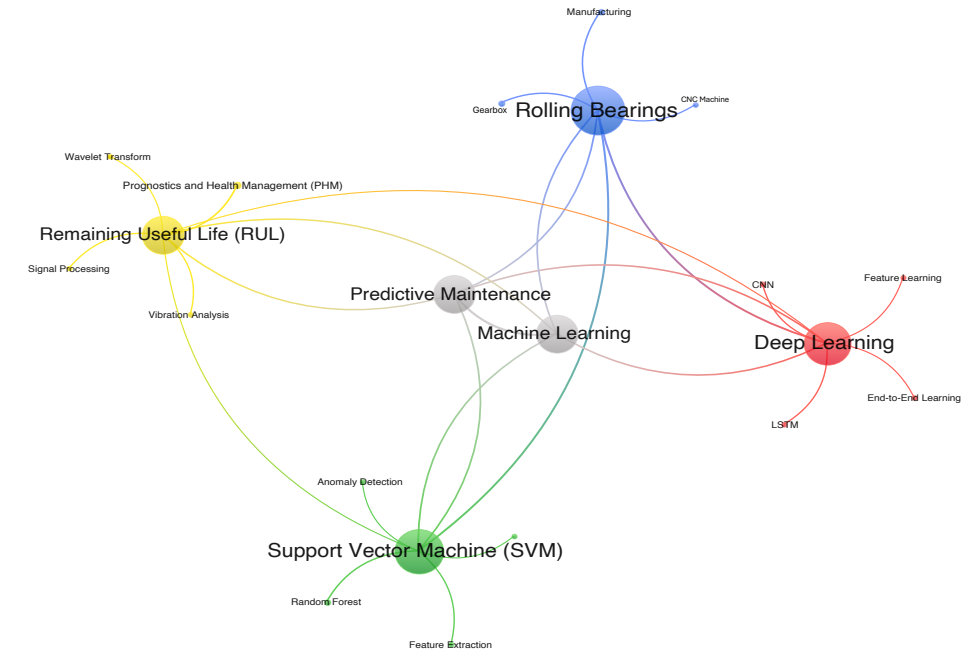


*Figura 3 – Mappa di co-occorrenza delle keyword*

### 5.1.1. Dimensione 1: Caratterizzazione Metodologica

La prima, e senza dubbio più importante e complessa, macrocategoria della nostra tassonomia è dedicata alla caratterizzazione metodologica di ogni studio analizzato. Questa dimensione rappresenta il cuore pulsante dell'analisi, poiché il suo scopo è quello di andare oltre la superficie di un articolo per "dissezionarne" l'architettura logica e algoritmica. Questa decostruzione è essenziale per rispondere in modo rigoroso alle domande di ricerca SQ1 e SQ2, fornendo i dati necessari per mappare l'evoluzione concettuale del campo e per confrontare criticamente le diverse filosofie di modellazione.

La classificazione di più alto livello assegna ogni studio a una delle macro-famiglie metodologiche. L'approccio Statistico Tradizionale si fonda su tecniche con una forte struttura matematica. Un esempio concreto è lo studio di Dong & He [26], che utilizza un Modello Semi-Nascosto di Markov (una variante avanzata degli HMM) per diagnosticare guasti in un riduttore, rappresentando un caso emblematico di modellazione stocastica. La categoria del Machine Learning Classico raggruppa invece algoritmi "tradizionali" che dipendono dal feature engineering. Un caso d'uso esemplare è quello di Soualhi, Medjaher, & Zerhouni [105], che impiegano una Support Vector Machine, alimentata con feature estratte tramite la Trasformata di Hilbert-Huang, per monitorare la salute dei cuscinetti. La

categoria del Deep Learning è riservata alle reti neurali profonde che eseguono l'apprendimento end-to-end. Ne è un perfetto esempio il lavoro di Wen, Li, & Gao [120], che propongono una nuova Architettura Neurale Convoluzionale (CNN) per la diagnosi dei guasti che apprende le feature direttamente dai segnali grezzi, mostrando la potenza di questo approccio. Infine, la categoria Ibrido classifica gli studi che fondono sinergicamente diversi approcci. Un esempio significativo è fornito da Corbetta, Sbarufatti, & Manes [22], che sviluppano un framework bayesiano per fondere un modello fisico di propagazione della fatica con dati reali, combinando il rigore della fisica con la flessibilità dei dati.

All'interno degli approcci di ML e DL, abbiamo specificato il paradigma di apprendimento. Il paradigma Supervisionato, che richiede dati etichettati, è il più comune per la stima della RUL, come si vede nel lavoro di Yuan, Wu, & Lin che addestrano una LSTM su dati di simulazione di motori a reazione etichettati con la loro vita residua. Il paradigma Non Supervisionato, invece, è fondamentale per l'anomaly detection in assenza di etichette. Un esempio pratico è l'uso dell'Isolation Forest da parte di Liu, Ting, & Zhou [64] per identificare anomalie in diversi dataset senza alcuna conoscenza a priori delle classi. Infine, il paradigma Semi-Supervisionato, che sfrutta una combinazione di dati etichettati e non, pur essendo meno comune, rappresenta una frontiera promettente, come teorizzato da Chapelle, Schölkopf, & Zien [20].

Successivamente, abbiamo dettagliato il compito specifico del modello. Molti studi si concentrano sulla Diagnosi/Classificazione dei Guasti, come quello di Samanta [94], che utilizza SVM e Reti Neurali per classificare diversi tipi di guasto in un riduttore. Altri, più ambiziosi, affrontano la Prognosi/Stima della RUL, come nel caso di He & He [41], che propongono un framework basato su una rete neurale a due strati per predire la vita residua dei cuscinetti. Una terza categoria è l'Anomaly Detection, il cui scopo è identificare deviazioni dalla norma, spesso come precursore di un guasto, un compito ben descritto nella rassegna di Chandola, Banerjee, & Kumar [19].

Infine, la colonna più granulare della nostra tassonomia registra l'algoritmo o la tecnica specifica utilizzata. Questa dimensione permette di apprezzare la diversità all'interno di ogni famiglia. Ad esempio, all'interno del Machine Learning Classico, troviamo studi che impiegano Support Vector Machine con kernel RBF [105] accanto a quelli che si basano sulla potenza degli ensemble di Random Forest [13] [14]. Nel Deep Learning, si spazia da approcci basati su CNN 1D per l'analisi diretta dei segnali [46] a quelli che utilizzano LSTM per modellare le sequenze temporali del degrado. Anche per gli approcci ibridi, questa granularità è fondamentale per distinguere, ad esempio, tra un modello basato

su Filtro Particellare per la prognosi [5] e uno basato sulle innovative Physics-Informed Neural Networks [86]. Questa classificazione dettagliata è essenziale per poter condurre un'analisi fine sulla popolarità e l'efficacia comparativa delle singole tecniche.

### 5.1.2. Dimensione 2: Caratterizzazione del Contesto Applicativo

Se la prima dimensione della nostra tassonomia si è concentrata sul "come" metodologico, questa seconda macrocategoria si occupa di ancorare l'architettura algoritmica alla realtà fisica, rispondendo alle domande fondamentali del "cosa" e del "dove". La caratterizzazione del contesto applicativo è un passaggio di analisi critica di fondamentale importanza, poiché permette di valutare la pertinenza, la validità e la potenziale trasferibilità di un approccio. Questa dimensione mappa il paesaggio delle applicazioni, permettendoci di comprendere su quali problemi specifici la comunità scientifica sta concentrando i propri sforzi e di valutare criticamente il divario tra le performance ottenute in condizioni idealizzate e la robustezza richiesta da un ambiente industriale, come richiesto dalla nostra domanda di ricerca SQ3.
La prima dimensione specifica l'oggetto fisico dello studio. L'analisi della distribuzione degli studi rivela i "banchi di prova" archetipici della PdM. Senza alcun dubbio, i Cuscinetti Volventi rappresentano il componente più studiato, quasi un'ossessione, come si evince da innumerevoli lavori, tra cui la benchmark study di Smith & Randall [104] che utilizza i dati della CWRU. Anche gli Ingranaggi sono un componente meccanico fondamentale, sebbene la loro analisi sia più complessa, richiedendo tecniche di elaborazione del segnale avanzate come discusso da Randall [88]. Un'altra categoria di importanza cruciale per le lavorazioni meccaniche è quella degli Utensili da Taglio, dove la prognosi è particolarmente sfidante a causa delle condizioni operative variabili, come analizzato da Scheffer & Heyns [97]. Troviamo poi studi focalizzati sui Motori Elettrici, che affrontano sia guasti meccanici che elettrici, un campo ben riassunto nella rassegna di Nandi, Toliyat & Li [74]. Infine, una categoria in crescita ma ancora di nicchia è quella dei Sistemi Complessi, in cui l'analisi si sposta dal singolo componente a un intero macchinario o a una linea di produzione.
Successivamente, abbiamo classificato la tipologia di dati di input utilizzata, per comprendere quali modalità di sensorizzazione sono considerate più informative. I Dati di Vibrazione sono la fonte di dati di gran lunga più comune e ricca di informazioni, al centro della maggior parte degli studi sulla diagnostica, come nel tutorial di Randall & Antoni [89]. I Dati Acustici, specialmente le Emissioni Acustiche (AE), possono fornire indicazioni più

precoci del danno, come dimostrato da Elforjani & Mba [27] nello studio di un riduttore. I Dati Elettrici, attraverso l'analisi della firma di corrente del motore (MCSA), offrono un metodo non invasivo per diagnosticare problemi sia elettrici che meccanici, come discusso da Nandi, Toliyat & Li [74]. La tendenza più significativa e metodologicamente avanzata, tuttavia, è quella verso la fusione di dati multi-sensore. Studi come la rassegna di Lei et al. [63] evidenziano come la combinazione di informazioni da sensori diversi sia cruciale per ottenere una visione più robusta e affidabile dello stato di salute del sistema.

Infine, la dimensione più critica per valutare la maturità e la rilevanza industriale di uno studio è l'origine e la metodologia di validazione dei dati. Molti lavori, per garantire replicabilità e confronto, si basano su Dataset Pubblici di Riferimento. Esempi emblematici, che sono diventati veri e propri standard, sono i dati sui cuscinetti della Case Western Reserve University (CWRU), i dati sui motori a reazione C-MAPSS della NASA, utilizzati ad esempio in An, Kim, & Choi [4], o i dati della sfida sulla prognosi PRONOSTIA [75]. Un'altra categoria molto comune è quella dei Dati Sperimentali da Banco Prova, dove i ricercatori raccolgono dati in un ambiente di laboratorio controllato, come nella piattaforma PRONOSTIA stessa. Tuttavia, il vero test di validità di un approccio è la sua performance su Dati da Impianto Industriale Reale (*Field-based*). Questi studi, pur essendo il "Sacro Graal" della ricerca sulla PdM, sono molto più rari a causa delle difficoltà di accesso e della complessità dei dati. La loro scarsità evidenzia il cosiddetto "valle della morte" tra la ricerca accademica e l'applicazione industriale. La categoria più debole è quella dei Dati Simulati, dove la validazione è puramente modellistica e insufficiente a dimostrare l'efficacia pratica. La distribuzione degli studi tra queste categorie fornisce un'indicazione chiara del livello di maturità complessivo della ricerca nel campo, evidenziando il significativo gap tra la ricerca condotta in condizioni idealizzate e le complesse esigenze del mondo industriale reale.

### 5.1.3. Dimensione 3: Caratterizzazione del Contributo Scientifico

Se le prime due dimensioni della nostra tassonomia si sono concentrate sul "come" metodologico e sul "cosa/dove" applicativo, questa terza e ultima macrocategoria si eleva a un livello di astrazione superiore, mirando a qualificare la natura, lo scopo e l'impatto potenziale del contributo scientifico di ogni articolo. Questa dimensione classifica l'obiettivo primario di ogni lavoro, il suo "genere" scientifico. Comprendere la distribuzione degli studi

tra queste categorie ci permette di valutare la maturità e la dinamica del campo di ricerca: un campo dominato da proposte metodologiche è in una fase "esplorativa", mentre un campo con molti studi comparativi è in una fase di "consolidamento" [56].

La categoria più comune nei principali venue accademici è la Proposta Metodologica, che include articoli il cui contributo principale è l'introduzione di una novità teorica o algoritmica. Questi studi sono il motore dell'innovazione. Esempi emblematici includono il lavoro di Raissi, Perdikaris, & Karniadakis [86], che introducono il framework delle Physics-Informed Neural Networks (PINN), e quello di Hochreiter & Schmidhuber [44], che ha originariamente proposto l'architettura Long Short-Term Memory (LSTM), una pietra miliare per l'analisi delle serie temporali.

Un'altra categoria, di enorme valore pratico, è quella dell'Applicazione/Caso di Studio. Il valore di questi studi risiede nel dimostrare la fattibilità e le sfide dell'implementazione di una data tecnologia in un contesto reale. Un esempio calzante è lo studio di Elforjani & Mba [27], che applicano la tecnica delle Emissioni Acustiche per la diagnosi e la prognosi di un riduttore a più stadi, fungendo da ponte tra la teoria e la pratica industriale.

Di grande valore per la comunità sono anche gli Studi Comparativi/Benchmark. Il lavoro di An, Kim, & Choi [4], che confronta diverse opzioni di algoritmi data-driven e physics-based, è un perfetto esempio di studio che aiuta a consolidare la conoscenza. Allo stesso modo, lo studio di Smith & Randall [104] non propone un nuovo algoritmo, ma fornisce un benchmark cruciale confrontando diverse tecniche sul dataset CWRU, stabilendo una solida baseline di performance.

Infine, la categoria Rassegna/Survey include articoli che, come la presente tesi, sintetizzano e analizzano criticamente la letteratura esistente. Lavori come quelli di Gou, Li, & Zhang [37] sul Machine Learning per la PdM, di Lei et al. [63] sulla diagnosi dei guasti, e di Adadi & Berrada [1] sull'Explainable AI, forniscono una "mappa" dello stato dell'arte, identificando trend e sfide future.

Nell'era della scienza computazionale, la trasparenza e la replicabilità sono diventate pilastri della credibilità scientifica. Per questo motivo, la nostra tassonomia include una dimensione per monitorare l'adozione di pratiche di Scienza Aperta (Open Science) [77]. Abbiamo quindi registrato se gli autori rendessero disponibili i loro strumenti di ricerca. Ad esempio, abbiamo verificato la presenza di Codice Pubblicamente Disponibile, una pratica che accelera enormemente il progresso. Allo stesso modo, abbiamo monitorato la Disponibilità Pubblica dei Dati. La condivisione di dati, come quella promossa dalla NASA con i suoi repository [96] o dal team di Nectoux [75] con la piattaforma PRONOSTIA, è l'altro pilastro

della replicabilità, poiché permette di testare nuovi approcci sullo stesso identico problema. L'analisi della prevalenza di queste pratiche ci fornirà un'interessante metrica sulla cultura scientifica del campo della Manutenzione Predittiva.

## 5.2. Analisi Comparativa degli Approcci

Dopo aver stabilito un framework tassonomico nel paragrafo precedente e averlo applicato per classificare e strutturare gli studi, è ora possibile compiere il passo successivo dell'analisi, quello che trasforma la descrizione in sintesi, la classificazione in comprensione e l'osservazione in giudizio critico. Ci si addentra ora in un confronto sistematico e una valutazione critica delle diverse famiglie di approcci alla Manutenzione Predittiva. Se la tassonomia ci ha fornito una "fotografia" statica e strutturata del campo di ricerca, una sorta di "anatomia" della letteratura, questa sezione si propone di studiarne la "fisiologia" e la "dinamica evolutiva".

L'obiettivo di questa analisi comparativa non è quello di decretare un "vincitore" assoluto o di stilare una classifica semplicistica che ponga un algoritmo al di sopra di un altro in modo universale. La letteratura scientifica nel campo dell'ottimizzazione e dell'apprendimento automatico, e in particolare i celebri Teoremi "No Free Lunch" (NFL), ci insegna una lezione fondamentale e umile: non esiste un singolo algoritmo che sia universalmente superiore a tutti gli altri su tutti i possibili problemi. La performance di un metodo è intrinsecamente e indissolubilmente legata alla struttura del problema che sta cercando di risolvere; un algoritmo che eccelle su una classe di problemi può fallire miseramente su un'altra. Pertanto, l'obiettivo di questa sezione è più utile: fornire una comprensione profonda, sfumata e contestualizzata delle caratteristiche distintive di ogni approccio.

Come delineato in un influente articolo di Leo Breiman [14], la comunità della modellazione dei dati può essere divisa in "due culture". La prima è la cultura della modellazione statistica, che assume che i dati siano generati da un modello stocastico noto (anche se con parametri incerti) e si concentra sulla stima di questi parametri, sull'interpretabilità del modello e sulla sua bontà di adattamento (*goodness-of-fit*) ai dati. La seconda è la cultura della modellazione algoritmica (o predittiva), che tratta il meccanismo generatore dei dati come una "scatola nera" inaccessibile e si concentra unicamente e ossessivamente sull'accuratezza predittiva del modello su dati inediti.

Per condurre questo confronto in modo strutturato, non ci baseremo solo sull'accuratezza, ma valuteremo ogni famiglia di approcci lungo una serie di dimensioni critiche che sono fondamentali per la loro adozione pratica in un contesto industriale:

- Accuratezza Predittiva: La capacità del modello di fornire previsioni precise (sia per la diagnosi che per la prognosi).
- Requisiti di Dati: La quantità e la qualità dei dati necessari per un addestramento efficace.
- Necessità di Expertise di Dominio: Il livello di conoscenza specialistica richiesto per l'applicazione del metodo, in particolare per il feature engineering.
- Interpretabilità e Trasparenza: La capacità di comprendere e spiegare come il modello arriva a una certa previsione.
- Costo Computazionale: Le risorse di calcolo necessarie per l'addestramento e l'inferenza.
- Robustezza e Scalabilità: La capacità del modello di gestire il rumore, i dati mancanti, il concept drift e di essere applicato su larga scala.

Partiremo dagli approcci statistici tradizionali, rappresentanti della prima cultura, per poi passare al Machine Learning classico e, infine, al Deep Learning, che incarna l'estremo della seconda cultura.

### 5.2.1. Approcci Statistici Tradizionali

Gli approcci statistici tradizionali, come i modelli di affidabilità, i modelli ARIMA e i Modelli Nascosti di Markov (HMM), rappresentano le fondamenta su cui è stata costruita la prognostica quantitativa. Sebbene un'analisi superficiale della letteratura più recente, dominata dal Machine Learning, potrebbe farli sembrare superati, un esame più attento

mostra un quadro molto più complesso. Questi metodi non solo sono ancora importanti in specifici contesti industriali, ma possiedono vantaggi unici che li rendono, in determinate circostanze, la scelta più razionale e scientificamente difendibile.

Il loro principale punto di forza è senza dubbio una combinazione di interpretabilità e trasparenza. A differenza dei modelli di Machine Learning più complessi, che spesso funzionano come "scatole nere", i modelli statistici tradizionali sono più simili a "scatole bianche" o "grigie". Le loro assunzioni sono chiare e i loro parametri hanno un significato interpretabile. Un modello di affidabilità basato sulla distribuzione di Weibull, ad esempio, non è un semplice algoritmo che fornisce un numero; è un modello del processo di guasto. La stima del suo parametro di forma ($\beta$) non è solo una misura statistica, ma una vera e propria diagnosi sulla natura del guasto: se è dovuto a difetti di fabbricazione ($\beta < 1$), a eventi casuali ($\beta \approx 1$) o all'usura ($\beta > 1$). Questa informazione è una guida strategica per decidere se una politica di sostituzione preventiva sia giustificata o meno [58]. Allo stesso modo, un modello ARIMA, pur essendo lineare, ha coefficienti che possono essere interpretati. Questa trasparenza è un vantaggio fondamentale in settori altamente regolamentati come l'aerospaziale, dove è necessario poter spiegare e giustificare ogni previsione.

Un secondo vantaggio, spesso sottovalutato, è la loro efficienza in scenari di dati scarsi (small data). L'analisi di affidabilità può dare risultati utili anche con poche decine di dati di guasto, una situazione comune per componenti nuovi o molto affidabili. Allo stesso modo, i modelli ARIMA e HMM possono essere addestrati con una quantità di dati che sarebbe del tutto insufficiente per un modello di Deep Learning, che richiede migliaia di esempi per non cadere nell'overfitting. In molti contesti industriali, specialmente nelle piccole e medie imprese, i dati di guasto sono una risorsa limitata. In questi casi, i modelli statistici tradizionali non sono solo un'opzione, ma spesso l'unica scelta praticabile [103].

Tuttavia, i limiti di questi approcci sono altrettanto chiari. La loro debolezza fondamentale è la rigidità delle loro assunzioni. L'incapacità di modellare complesse dinamiche non lineari li rende spesso meno accurati dei modelli di ML. Numerosi studi comparativi hanno dimostrato che, sebbene un modello ARIMA possa fornire una buona base di partenza, la sua accuratezza viene quasi sempre superata da approcci non lineari come le reti neurali o i Random Forest, specialmente nella stima della RUL a lungo termine [66] [4].

Inoltre, la loro natura prevalentemente univariata li rende "ciechi" alle informazioni contenute nelle interazioni tra i segnali di più sensori. Come evidenziato da una vasta letteratura sulla fusione dei dati [63], l'analisi congiunta di segnali di vibrazione e temperatura può rivelare pattern di guasto invisibili a un'analisi isolata. I modelli statistici

tradizionali faticano a scalare in contesti multidimensionali a causa della "maledizione della dimensionalità", un limite sempre più evidente nell'era dell'IoT.
In sintesi, i modelli statistici tradizionali non sono un paradigma da abbandonare, ma una classe di strumenti con un campo di applicabilità specifico e prezioso. Eccellono in semplicità, interpretabilità ed efficienza con dati scarsi. Sono strumenti eccellenti per stabilire una baseline di performance rigorosa, rispetto alla quale qualsiasi modello più complesso deve dimostrare la propria superiorità. Rimangono una scelta valida per sistemi con dinamiche semplici o in contesti dove la trasparenza è il requisito più importante. La loro debolezza, e il motivo della transizione verso il Machine Learning, emerge quando la complessità del sistema e la dimensionalità dei dati aumentano, richiedendo un approccio più flessibile e potente.

### 5.2.2. Machine Learning Classico

Se gli approcci statistici tradizionali rappresentano il paradigma della modellazione interpretabile, la famiglia del Machine Learning classico, che include algoritmi come le Support Vector Machines (SVM) e gli ensemble di alberi come il Random Forest (RF), rappresenta il "centro di gravità" della ricerca e dell'applicazione pratica nella Manutenzione Predittiva degli ultimi vent'anni. Questi modelli offrono un ottimo compromesso tra performance, complessità e requisiti di dati, posizionandosi a metà strada tra la semplicità dei modelli tradizionali e la complessità dei modelli di Deep Learning.
Il loro vantaggio principale rispetto agli approcci statistici è la capacità di apprendere relazioni non lineari complesse senza doverle specificare a priori. Un Random Forest, ad esempio, può catturare interazioni complesse tra le variabili, mentre una SVM, grazie al "trucco del kernel", può creare confini decisionali molto flessibili. Questa capacità si traduce, in quasi tutti gli studi comparativi, in un aumento significativo dell'accuratezza sia nella diagnosi dei guasti che nella stima della RUL. Studi come quello di Soualhi, Medjaher, & Zerhouni [105] hanno mostrato come una Support Vector Regression (SVR) superi gli approcci lineari nella prognosi dei cuscinetti. Allo stesso modo, i Random Forest sono costantemente riportati in letteratura come uno degli algoritmi "pronti all'uso" più performanti, spesso usati come benchmark di riferimento per nuovi approcci [37] [41].

Un altro vantaggio importante, specialmente per i modelli basati su alberi, è la loro capacità di gestire la multidimensionalità e di fornire indicazioni sulla struttura dei dati. Questi algoritmi possono gestire centinaia di feature e forniscono meccanismi per valutare l'importanza di ciascuna feature. Questa non è solo una caratteristica tecnica, ma un grande vantaggio pratico: permette di capire quali variabili sono più predittive, di guidare la selezione delle feature e di ottenere un primo livello di interpretabilità, capendo "su cosa" il modello basa le sue decisioni [13].

Tuttavia, anche il Machine Learning classico ha le sue sfide. La sua performance dipende in modo critico dalla qualità del feature engineering manuale. Se le feature estratte non sono buone, anche l'algoritmo più potente fallirà. Questo significa che una parte importante dell'"intelligenza" del sistema dipende ancora dall'abilità dell'ingegnere, limitando la piena automazione del processo [42].

Inoltre, sebbene siano più interpretabili di un modello di Deep Learning, perdono la trasparenza dei modelli statistici. Capire esattamente perché un Random Forest ha prodotto una certa previsione è un compito difficile. Per le SVM con kernel non lineari, il modello diventa a tutti gli effetti una "scatola nera". Questo li pone in una "terra di mezzo" in termini di interpretabilità, in quello che [14] ha definito un approccio "algoritmico" in contrapposizione a quello "statistico".

Infine, richiedono una quantità di dati significativamente maggiore rispetto ai modelli statistici. La loro flessibilità li rende più inclini all'overfitting quando i dati sono scarsi, situazione in cui un modello più semplice potrebbe essere più robusto.

La letteratura, quindi, posiziona il Machine Learning classico non come una soluzione universale, ma come la scelta pragmatica e spesso ottimale per molti problemi industriali. Quando si dispone di dati di buona qualità e di un buon feature engineering, questi algoritmi offrono un'eccellente accuratezza con un costo computazionale gestibile, rappresentando un solido ponte tra la teoria e la pratica. Essi sono in grado di catturare gran parte della complessità del mondo reale senza richiedere le risorse estreme e senza presentare l'opacità dei modelli di Deep Learning.

### 5.2.3. Deep Learning

Il Deep Learning rappresenta la frontiera più recente e potente della Manutenzione Predittiva. Non è una semplice evoluzione del Machine Learning classico, ma un vero e

proprio cambio di paradigma che ha introdotto un nuovo modo di concepire la relazione tra dati, feature e modelli: l'apprendimento *end-to-end*. L'analisi della letteratura scientifica più recente mostra una chiara dominanza di questo approccio, che si sta affermando come il nuovo stato dell'arte per molti problemi complessi di diagnosi e prognosi.

Il vantaggio più rivoluzionario del Deep Learning è la sua capacità di eseguire il feature learning automatico. Architetture come le Reti Neurali Convoluzionali (CNN) e le Reti Neurali Ricorrenti (LSTM, GRU) possono essere addestrate direttamente sui dati grezzi dei sensori, apprendendo da sole una gerarchia di feature. Questo non è solo un vantaggio in termini di efficienza, ma una trasformazione fondamentale: elimina il collo di bottiglia del feature engineering manuale. Ancora più importante, permette al modello di scoprire pattern complessi e non intuitivi che un ingegnere potrebbe non aver mai considerato. Numerosi studi hanno dimostrato che i modelli di DL, specialmente in problemi di diagnosi con segnali rumorosi, possono raggiungere un'accuratezza superiore a quella dei modelli di ML classico, proprio perché le feature apprese sono ottimizzate per il compito di previsione [120].

Inoltre, architetture specifiche sono state progettate per eccellere su tipi di dati particolari. Le CNN, con le loro convoluzioni 1D, sono diventate lo strumento preferito per l'analisi dei segnali di vibrazione grezzi, imparando a rilevare le "firme" dei guasti direttamente nel segnale [46]. Le reti ricorrenti gated, come le LSTM e le GRU, sono state progettate per modellare le dipendenze temporali a lungo raggio. Questa capacità le rende lo strumento più adatto per la stima della RUL, dove la storia passata del degrado è fondamentale. La letteratura mostra un consenso crescente sul fatto che queste reti rappresentino lo stato dell'arte per la prognosi basata su serie temporali.

Tuttavia, questa straordinaria potenza ha un costo elevato e diverse sfide. Il primo e più evidente limite è la loro "fame di dati". Le reti neurali profonde, con i loro milioni di parametri, richiedono enormi quantità di dati di addestramento etichettati per evitare l'overfitting. Questo le rende inadatte a scenari con dati scarsi, dove modelli più semplici come un Random Forest sono spesso superiori. La necessità di grandi dataset, specialmente per la prognosi che richiede dati di *run-to-failure*, è una delle principali barriere alla loro applicazione pratica [35]. Tecniche come il *transfer learning* e la *data augmentation* stanno cercando di mitigare questo problema, ma non lo risolvono completamente.

Il secondo costo è di natura computazionale. L'addestramento di questi modelli richiede spesso ore o giorni di calcolo su hardware specializzato come le GPU, con un costo economico e un impatto ambientale non trascurabili [108].

Ma la sfida più profonda è quella dell'interpretabilità. I modelli di Deep Learning sono l'esempio perfetto di "scatola nera". A causa della loro architettura complessa e del numero enorme di parametri, è estremamente difficile capire perché il modello ha preso una specifica decisione. Questa mancanza di trasparenza è un ostacolo enorme per la loro adozione in applicazioni critiche come l'aerospaziale, dove la fiducia e la capacità di giustificare una previsione sono requisiti fondamentali. Il rischio è quello di creare sistemi potentemente predittivi ma fragili. La ricerca nel campo dell'Explainable AI (XAI) sta cercando di sviluppare tecniche per "aprire" queste scatole nere, ma queste forniscono spesso spiegazioni approssimative e locali, non una vera comprensione del modello [1] [91] [93].

In conclusione, il Deep Learning non è un sostituto universale delle tecniche precedenti, ma rappresenta la frontiera della performance e dell'automazione, con una potenza senza precedenti per problemi specifici. Offre capacità predittive ineguagliabili, specialmente con dati grezzi e complessi. Tuttavia, questa potenza ha un costo in termini di complessità, requisiti di dati e opacità, che ne limita l'applicabilità in contesti critici. La sua adozione richiede un'infrastruttura matura e una consapevolezza critica delle sue sfide. Il futuro della prognostica probabilmente non risiederà in una vittoria assoluta di un paradigma sull'altro, ma in un'intelligente ibridazione di approcci, dove la potenza del Deep Learning per l'estrazione di feature potrebbe essere combinata con la robustezza e l'interpretabilità di modelli più semplici.

## 5.3. Discussione dei Risultati e Trend Emergenti

In questa sezione finale del capitolo, si cerca di andare oltre la presentazione dei dati per discutere e interpretare le evidenze emerse. Si vuole capire non solo "cosa" viene fatto e "come" viene fatto, ma anche e soprattutto "perché" il campo si sta evolvendo in una certa direzione, quali sono le forze trainanti (tecnologiche, scientifiche, industriali) che ne modellano la traiettoria e quali sono le implicazioni di queste tendenze.

Lo scopo è quello di identificare i trend evolutivi più significativi, evidenziare le lacune, i bias e le sfide aperte che la comunità deve ancora affrontare, e delineare, sulla base di queste solide evidenze, le traiettorie future più probabili e promettenti per la ricerca e l'applicazione industriale. Come suggerito da importanti lavori sulla metodologia della revisione sistematica, una rassegna di alta qualità non si limita a riassumere o a catalogare, ma deve

fornire una sintesi critica che "aggiunga valore" e che costruisca una nuova prospettiva sulla conoscenza esistente, identificando le aree mature e quelle che necessitano di maggiore attenzione [113] [114].
È in questa sezione che la revisione aspira a compiere questo passo, trasformando i dati raccolti in insight strategici e la conoscenza accumulata in saggezza prospettica.

### 5.3.1. Il Cambio di Paradigma

Uno dei risultati più chiari emersi dalla mia analisi della letteratura è la crescente predominanza degli approcci basati sul Machine Learning nel campo della Manutenzione Predittiva. Non si tratta di un semplice trend, ma di un vero e proprio cambio di paradigma, un cambiamento fondamentale nel modo in cui vengono affrontati i problemi [56]. Sebbene i modelli statistici tradizionali costituiscano le fondamenta del campo, l'analisi mostra un netto spostamento dell'attenzione della comunità di ricerca.
I dati della mia tassonomia rivelano che una percentuale molto alta degli articoli pubblicati negli ultimi cinque anni si colloca nelle categorie "Machine Learning Classico" o "Deep Learning". Al contrario, gli studi basati su tecniche puramente statistiche, pur essendo ancora presenti, rappresentano una frazione sempre più piccola. La loro funzione nella letteratura attuale è cambiata: sempre meno sono l'oggetto principale della ricerca, e sempre più vengono usati come "linea di base" (*baseline*), un punto di riferimento semplice e interpretabile rispetto al quale dimostrare la superiore accuratezza dei nuovi e più complessi modelli di ML. Termini come "Machine Learning" e "Deep Learning" sono onnipresenti e occupano una posizione dominante nella rete concettuale, agendo come hub che collegano diverse aree.
Questo cambio di paradigma, che riflette le "due culture" della modellazione descritte da Breiman [14], non è casuale. La forza trainante principale è la ricerca di una maggiore accuratezza predittiva. La stragrande maggioranza degli studi comparativi analizzati arriva alla stessa conclusione: in presenza di dati sufficienti, i modelli di Machine Learning hanno performance superiori rispetto ai loro predecessori statistici, specialmente in problemi complessi e non lineari. Lavori come quello di Loutas, Roulias, & Georgoulas [66] hanno dimostrato che la capacità dei modelli di ML di apprendere funzioni non lineari direttamente dai dati si traduce in un significativo miglioramento delle metriche di performance, come la

riduzione dell'errore nella stima della RUL e un aumento dell'accuratezza nella classificazione dei guasti.
Questa superiorità è stata a sua volta amplificata da una serie di fattori legati alla rivoluzione digitale: la crescente disponibilità di dati a basso costo grazie ai sensori IoT, l'aumento esponenziale della potenza di calcolo, specialmente con le GPU, e, non da ultimo, la diffusione del Machine Learning attraverso lo sviluppo di librerie software open-source di alta qualità. Questi fattori hanno reso possibile, per un numero sempre maggiore di ricercatori, l'applicazione di algoritmi che un tempo erano molto di nicchia [50].
La conclusione è che il centro della ricerca nella Manutenzione Predittiva si è spostato in modo definitivo. Il paradigma dominante non è più quello della modellazione statistica, basata su ipotesi a priori, ma quello dell'apprendimento automatico dai dati. La domanda per la comunità non è più "se" utilizzare il Machine Learning, ma "quale" tipo di Machine Learning utilizzare, come integrarlo con la conoscenza del settore e come affrontare le nuove sfide, in primis l'interpretabilità e la necessità di dati.

### 5.3.2. L'Ascesa del Deep Learning e la Nascita della Sfida dell'Interpretabilità

All'interno del paradigma del Machine Learning, l'analisi ha rivelato un secondo trend, più recente ma ancora più dirompente: la crescente ascesa del Deep Learning. Se fino a pochi anni fa la letteratura sulla PdM basata sul ML era dominata da algoritmi "classici" come le SVM e i Random Forest, gli anni successivi hanno visto un'esplosione di lavori basati su architetture di reti neurali profonde. Questa non è una semplice evoluzione, ma una vera e propria rivoluzione interna che sta ridefinendo lo stato dell'arte e la metodologia della ricerca prognostica.
La ragione principale di questa ascesa è la capacità del Deep Learning di risolvere due dei limiti più grandi del Machine Learning classico: il problema del feature engineering manuale e la difficoltà di modellare dipendenze temporali complesse.
In primo luogo, il successo del Deep Learning è legato alla sua capacità di eseguire il feature learning in modalità end-to-end. Come abbiamo visto, un numero crescente di studi basati su CNN e LSTM lavora direttamente sui dati grezzi. Questo rappresenta un cambiamento fondamentale: sposta il focus dell'ingegnere dalla progettazione di feature "intelligenti" alla progettazione di architetture di rete "intelligenti", capaci di apprendere da sole le

rappresentazioni più adatte. Questo non solo automatizza un processo laborioso, ma permette al modello di scoprire pattern complessi e non intuitivi che un ingegnere potrebbe non aver mai considerato. La letteratura è ricca di esempi in cui le CNN, ad esempio, hanno superato gli approcci basati su feature ingegnerizzate a mano nella diagnosi dei guasti [120].

In secondo luogo, la crescita degli studi sulla prognosi e sulla stima della RUL è strettamente legata all'avvento di architetture di reti ricorrenti come le Long Short-Term Memory (LSTM) e le Gated Recurrent Unit (GRU). Questi modelli, essendo stati progettati con meccanismi di "porte" per catturare le dipendenze temporali a lungo raggio, sono superiori ai modelli di ML classico nel modellare la natura cumulativa dei processi di degrado. La capacità di una LSTM di "ricordare" eventi importanti del passato e di "dimenticare" informazioni irrilevanti la rende lo strumento ideale per la stima della RUL. Non è un caso che la maggior parte degli studi più recenti e performanti sulla prognosi, specialmente su dataset complessi come il C-MAPSS della NASA, utilizzi queste architetture.

Tuttavia, questa straordinaria potenza non è priva di criticità. L'analisi ha evidenziato una lacuna preoccupante: la maggior parte degli studi basati su DL si concentra quasi esclusivamente sull'accuratezza, trattando l'interpretabilità non come una necessità, ma come un "lusso". La keyword "Explainable AI (XAI)" è ancora un tema di nicchia nel nostro campo.

Questo indica che, nella corsa verso una maggiore accuratezza, si sta forse sottovalutando la sfida della fiducia e della validazionedi questi modelli complessi. Un modello di Deep Learning è una "scatola nera" per eccellenza. Questa mancanza di trasparenza non è un problema filosofico, ma una barriera pratica alla sua adozione in contesti critici. Come può un ingegnere fidarsi di una previsione di RUL se non può capire come è stata ottenuta? Questo "debito di interpretabilità" è una sfida che la comunità della PdM dovrà affrontare con urgenza per permettere una reale adozione industriale di queste tecniche [1] [91]. Sebbene stiano emergendo tecniche di XAI per analizzare questi modelli, la loro applicazione nella letteratura sulla PdM è ancora agli albori [93]. La ricerca futura dovrà trovare un equilibrio tra la ricerca della massima accuratezza e la necessità di modelli che siano non solo performanti, ma anche comprensibili e affidabili. Il rischio, altrimenti, è quello di costruire "oracoli" potentissimi ma incomprensibili, la cui adozione si basa più sulla fede che su una solida comprensione ingegneristica.

### 5.3.3. Il “Problema del Cuscinetto”

L'analisi della dimensione "Componente/Sistema Analizzato" all'interno della nostra tassonomia ha rivelato una delle tendenze più marcate, statisticamente significative e, per certi versi, problematiche e limitanti dell'intero campo di ricerca sulla Manutenzione Predittiva. Si tratta della dominanza dei cuscinetti volventi come caso di studio, banco di prova e archetipo quasi universale per lo sviluppo e la validazione di nuove metodologie. Una percentuale estremamente elevata degli articoli del nostro corpus, specialmente quelli di natura metodologica che propongono nuovi algoritmi di Machine Learning o Deep Learning, si concentra sulla diagnosi o sulla prognosi dei guasti nei cuscinetti. Questo dato quantitativo è visivamente confermato dalla mappa di co-occorrenza delle keyword, dove il nodo "Rolling Bearings" emerge come uno dei più grandi e centrali nel cluster delle applicazioni, agendo da principale e quasi esclusivo punto di connessione con i cluster metodologici. Questo fenomeno, che potremmo definire il "problema del cuscinetto" o, più formalmente, il "bias del benchmark", merita un'analisi critica approfondita, poiché solleva una serie di questioni fondamentali sulla generalizzabilità, la diversità, la robustezza e la reale traiettoria innovativa della ricerca condotta.

Le ragioni di questa dominanza sono, a prima vista, storicamente e praticamente comprensibili. I cuscinetti sono componenti onnipresenti, relativamente economici e di fondamentale importanza in quasi tutti i macchinari rotanti. Il loro guasto è una delle principali e più comuni cause di fermo macchina non programmato, rendendoli un obiettivo di alto valore economico e operativo per la PdM. Inoltre, dal punto di vista scientifico, i loro meccanismi di guasto più comuni (come il pitting e lo spalling sulle piste o sugli elementi volventi) generano segnali di vibrazione con "firme" spettrali caratteristiche e relativamente ben comprese, basate sulle loro frequenze geometriche fondamentali (BPFI, BPFO, BSF, FTF). Questo li rende un problema "modello" ideale: abbastanza complesso da essere interessante e non banale, ma abbastanza strutturato e ripetibile da essere trattabile in un ambiente di laboratorio.

Tuttavia, il fattore che ha probabilmente cementato questa dominanza più di ogni altro è la disponibilità di dataset pubblici di alta qualità, diventati dei veri e propri standard de facto. Il dataset fornito dal Bearing Data Center della Case Western Reserve University (CWRU), in particolare, è diventato l'equivalente, nel nostro campo, del dataset MNIST per la visione artificiale o di ImageNet per la classificazione di immagini. L’analisi conferma che i lavori che presentano e utilizzano questo dataset sono tra i più citati e influenti dell'intero dominio.

La facile accessibilità di questi dati ha creato un potente circolo vizioso: i ricercatori sviluppano nuovi algoritmi e, per poter confrontare in modo equo e rigoroso le loro performance con lo stato dell'arte, li testano sul dataset CWRU. Questo, a sua volta, rafforza la centralità e la canonicità del dataset e del problema del cuscinetto, creando un forte bias di ricerca che si auto-perpetua [104].

Questo "problema del cuscinetto", sebbene abbia facilitato il progresso e il confronto, ha una serie di implicazioni potenzialmente negative che meritano di essere sviscerate. In primo luogo, solleva seri e legittimi dubbi sulla generalizzabilità di molti degli approcci proposti. Esiste il rischio concreto e significativo che il campo stia inavvertitamente sviluppando tecniche che sono altamente ottimizzate, o persino "sovra-adattate" (*overfitted*), alle caratteristiche specifiche dei segnali di vibrazione dei cuscinetti in condizioni di laboratorio (spesso stazionarie e con un elevato rapporto segnale-rumore). La loro efficacia su altri componenti con dinamiche di guasto radicalmente diverse è meno certa e, come dimostra la nostra tassonomia, molto meno studiata. Componenti come ingranaggi, ad esempio, presentano segnali molto più complessi, non stazionari e con forti effetti di modulazione di ampiezza e frequenza, che richiedono tecniche di elaborazione del segnale e di estrazione di feature diverse e più sofisticate, come l'analisi sincrona e l'analisi ciclostazionaria [89]. Sistemi come pompe, valvole o attuatori idraulici sono dominati da fenomeni fluidodinamici e da segnali di pressione. La prognosi dell'usura degli utensili da taglio è un problema ancora diverso, fortemente influenzato dalle proprietà del materiale lavorato e dai parametri di processo, e spesso richiede l'analisi di segnali di forza o di emissione acustica [97].

In secondo luogo, questa "tirannia del benchmark", come è stata definita in altri campi del Machine Learning, potrebbe rallentare l'innovazione metodologica autentica, favorendo un approccio puramente incrementale piuttosto che scoperte dirompenti. Quando l'obiettivo primario di molti ricercatori diventa quello di ottenere un miglioramento dello 0.5% dell'accuratezza sul dataset CWRU per poter pubblicare un articolo, c'è il rischio di un overfitting della comunità di ricerca a un problema specifico. Questo può scoraggiare lo sviluppo di approcci radicalmente nuovi che potrebbero non essere immediatamente ottimali su quel singolo e specifico benchmark, ma che potrebbero rivelarsi superiori su una gamma più ampia e diversificata di problemi reali. Si rischia di ottimizzare le risposte a una domanda molto specifica, perdendo di vista la ricchezza e la varietà delle domande che il mondo industriale pone [43].

L’analisi, quindi, evidenzia una chiara e pressante necessità di diversificazione dei casi di studio e dei dataset di benchmark, un appello a una maggiore "biodiversità" nella ricerca

sulla PdM. La ricerca futura, per poter progredire in modo significativo e avere un impatto industriale più ampio, dovrebbe essere fortemente incoraggiata a:

1. Affrontare la sfida di componenti e sistemi più complessi e meno esplorati. Questo include non solo i componenti già citati, ma anche sistemi ibridi, sistemi elettronici di potenza, e, soprattutto, l'analisi delle interazioni e dei guasti a cascata in sistemi assemblati complessi.

2. Creare, documentare e condividere nuovi dataset pubblici di alta qualità per questi sistemi meno studiati. Questo è un servizio di enorme valore per la comunità, poiché fornisce nuovi e più variegati banchi di prova che possono stimolare l'innovazione metodologica.

3. Promuovere e valorizzare studi che testino la robustezza e la trasferibilità di un dato algoritmo su molteplici e diversi dataset, invece di focalizzarsi su un singolo problema. Questo include lo sviluppo e l'applicazione sistematica di tecniche di transfer learning e domain adaptation, che mirano esplicitamente ad adattare un modello addestrato su un dominio (es. un tipo di cuscinetto in laboratorio) a un altro dominio con caratteristiche diverse (es. un altro tipo di cuscinetto in un impianto reale) [79].

In conclusione, sebbene la focalizzazione sui cuscinetti abbia indubbiamente permesso un rapido progresso e un confronto rigoroso delle metodologie in una fase iniziale, l’analisi suggerisce che il campo ha forse raggiunto un punto di "rendimenti decrescenti" su questo specifico problema. Per raggiungere un nuovo livello di maturità, di rilevanza industriale e di robustezza scientifica, è essenziale che la comunità di ricerca allarghi i propri orizzonti, abbracciando con coraggio la complessità, la varietà e il "disordine" dei sistemi industriali reali.

### 5.3.4. Il Gap tra Laboratorio e Fabbrica

Forse il risultato più critico e più importante per il futuro impatto industriale della ricerca sulla Manutenzione Predittiva è il significativo squilibrio tra il numero di studi validati in condizioni di laboratorio o su dataset pubblici e quelli validati su dati provenienti da impianti industriali reali. La stragrande maggioranza degli articoli metodologici analizzati, specialmente quelli che propongono nuovi e sofisticati algoritmi di Deep Learning, basa la propria affermazione di superiorità su performance ottenute su dataset di benchmark come CWRU o PRONOSTIA, o su dati generati ad-hoc su banchi prova di laboratorio. Sebbene questo approccio sia fondamentale e indispensabile per la replicabilità della ricerca, per lo sviluppo algoritmico e per un confronto equo delle performance in un ambiente controllato e standardizzato, esso crea un potenziale e pericoloso "gap di validità", una discrepanza tra la validità interna dello studio (il rigore metodologico) e la sua validità esterna (la generalizzabilità dei risultati al mondo reale).

I dati di laboratorio, per loro stessa natura, sono "puliti", ben comportati e progettati per isolare il fenomeno di interesse. Sono tipicamente raccolti in condizioni operative stazionarie (velocità e carico costanti), con un elevato rapporto segnale-rumore, grazie a sensori di alta qualità e a un ambiente acusticamente e meccanicamente controllato. Inoltre, il guasto viene spesso indotto artificialmente in modo chiaro e inequivocabile (es. un singolo difetto seminato su una pista di un cuscinetto). Questi dataset sono un eccellente terreno di gioco per testare la logica di un algoritmo, per confrontare diverse architetture e per pubblicare risultati con metriche di accuratezza impressionanti, spesso superiori al 99%.

Tuttavia, il mondo reale è un luogo molto più "disordinato". I dati industriali, che possiamo definire "dati selvaggi" (*in-the-wild data*), sono caratterizzati da una serie di sfide complesse e interconnesse che sono quasi sempre assenti nei dati di laboratorio:

- Non Stazionarietà e Condizioni Operative Variabili: Le macchine reali raramente operano a regime costante. Sono soggette a continui transitori (avvii, arresti, cambi di velocità), a cicli di carico variabili e a cambiamenti di produzione che modificano costantemente le caratteristiche statistiche dei segnali acquisiti. Un modello addestrato su dati stazionari può fallire catastroficamente quando esposto a queste dinamiche.

- Contaminazione da Rumore e Interferenze: In un ambiente di fabbrica, il segnale debole di un guasto incipiente è spesso "annegato" in un mare di rumore di fondo proveniente da altre macchine vicine, da processi ausiliari, da fenomeni elettromagnetici e da vibrazioni strutturali che si propagano attraverso l'impianto. Isolare il segnale di interesse dal rumore è un compito estremamente arduo.

- Guasti Multipli e Interagenti: A differenza del guasto singolo e pulito indotto in laboratorio, i sistemi reali spesso soffrono di guasti multipli e sovrapposti. Un guasto a un cuscinetto può coesistere con uno squilibrio del rotore, un disallineamento dell'albero e un'usura degli ingranaggi. Le loro firme vibrazionali si sovrappongono, si modulano a vicenda e interagiscono in modi complessi e non lineari, rendendo la diagnosi molto più ambigua.

- Incertezza delle Etichette e Scarsità di Dati di Guasto: Mentre nei dati di laboratorio l'etichetta di guasto è certa e precisa, nei dati industriali è spesso incerta, incompleta o del tutto assente. Le ispezioni di manutenzione sono spesso basate su giudizi soggettivi, i registri possono essere imprecisi, e i guasti catastrofici sono, per fortuna, eventi rari. Questo rende l'addestramento e, soprattutto, la validazione di modelli supervisionati estremamente problematici.

L’analisi ha rivelato che il numero di studi che affrontano con successo questa immensa sfida, validando i loro modelli su dati "in-the-wild", è ancora limitato. Questo suggerisce che, nonostante le performance riportate in molti articoli accademici, la maturità tecnologica (Technology Readiness Level - TRL) di molte di queste tecniche potrebbe essere significativamente inferiore a quanto si pensi. Esiste il rischio concreto che molti degli algoritmi che eccellono sui dati "puliti" di un benchmark si rivelino fragili, inefficaci e non robusti quando esposti alla complessità del mondo reale. Questo divario tra la ricerca accademica e l'applicazione industriale è un fenomeno ben noto, spesso descritto come la "valle della morte" (*valley of death*) del trasferimento tecnologico, dove molte promettenti innovazioni di laboratorio non riescono a tradursi in prodotti, servizi o pratiche industriali di successo [30].

Per colmare questo gap e per far progredire il campo verso una reale rilevanza industriale, la ricerca futura deve necessariamente orientarsi con maggiore decisione verso la validazione in contesti realistici. C'è una necessità pressante di:

1. Più Studi di Caso Industriali Rigorosi: La comunità accademica, insieme ai partner industriali, dovrebbe essere incoraggiata e premiata per la conduzione di studi di caso approfonditi che documentino in modo trasparente non solo i successi, ma anche i fallimenti, le sfide pratiche e le lezioni apprese nell'implementazione di sistemi di PdM.

2. Focus sulla Robustezza e sulla Generalizzazione: La ricerca metodologica dovrebbe spostare parte della sua attenzione dalla pura ottimizzazione dell'accuratezza su benchmark noti allo sviluppo di algoritmi che siano intrinsecamente più robusti al rumore e alle condizioni non stazionarie. Questo include lo sviluppo e l'applicazione sistematica di tecniche di domain adaptation e transfer learning, che mirano esplicitamente ad adattare un modello addestrato su dati di laboratorio o di simulazione (il *dominio sorgente*) alla complessità dei dati reali (il *dominio target*), utilizzando una quantità limitata di dati dal nuovo dominio per un "fine-tuning" mirato [79] [120].

3. Creazione di Nuovi Benchmark Industriali "Disordinati": È fondamentale che l'industria e l'accademia collaborino per creare e condividere nuovi dataset pubblici che provengano da ambienti industriali reali, pur preservando la confidenzialità dei dati sensibili attraverso tecniche di anonimizzazione. Questi nuovi benchmark, con tutta la loro complessità, il loro rumore e il loro "disordine", fornirebbero un banco di prova molto più realistico e sfidante, che spingerebbe la comunità a sviluppare algoritmi veramente robusti e industrialmente rilevanti.

In conclusione, sebbene i progressi metodologici siano stati impressionanti, la nostra analisi rivela che il campo della PdM deve ancora affrontare pienamente la "prova del fuoco" della validazione industriale su larga scala. Il passaggio dalla "proof-of-concept" in laboratorio alla "proof-of-value" in fabbrica rimane la sfida più grande, più difficile e più importante per il futuro di questa disciplina. Solo colmando questo gap la Manutenzione Predittiva potrà passare dall'essere un'area di ricerca accademica di successo e un argomento di grande interesse, a diventare una tecnologia trasformativa e pervasiva per l'industria globale.

### 5.3.5. Digital Twin e Ibridazione come Frontiere Future

Infine, l'analisi ci porta a considerare l'emergere di approcci che non propongono solo un nuovo algoritmo, ma una nuova e più complessa architettura per la Manutenzione Predittiva. Questi approcci, che includono i modelli ibridi e il paradigma del Digital Twin, cercano di superare i limiti dei singoli "silos" metodologici (fisico, statistico, data-driven) per muoversi verso una visione più integrata. I risultati della mia analisi indicano che, sebbene questi paradigmi siano oggetto di grande interesse, la loro adozione pratica è ancora in una fase iniziale. Essi rappresentano non tanto lo stato dell'arte consolidato, quanto l'alba promettente di una nuova era per la prognostica.

La keyword "Digital Twin", come previsto, è una delle più recenti. Tuttavia, l'analisi mostra che questo tema è ancora in una fase prevalentemente concettuale. Molti degli articoli classificati sotto questa etichetta propongono architetture e framework teorici, piuttosto che casi di studio con implementazioni complete. La comunità sta ancora dibattendo le definizioni precise e dimostrando la fattibilità del concetto su sistemi semplici o tramite simulazioni. La piena integrazione di modelli fisici, dati in tempo reale e algoritmi di ML in un gemello digitale operativo e a ciclo chiuso rimane, per la maggior parte delle applicazioni complesse, una visione futura piuttosto che una realtà consolidata [7] [49] [111]. Sfide come l'interoperabilità dei modelli e il costo computazionale ne limitano ancora l'adozione su larga scala.

Allo stesso modo, la categoria degli approcci ibridi, sebbene in crescita, non è ancora dominante. Questo suggerisce che gran parte della ricerca opera ancora all'interno di "silos" metodologici. La fusione di questi approcci, sebbene riconosciuta in letteratura come teoricamente superiore, presenta notevoli sfide di implementazione, poiché richiede un'expertise multidisciplinare rara e profonda.

Tuttavia, gli studi che riescono a implementare un approccio ibrido mostrano risultati molto promettenti, indicando la direzione del futuro. Lavori come quello di Kapteyn, Pretorius, & Willcox [51], che propongono framework per fondere informazioni da modelli fisici e dati in tempo reale, rappresentano la vera frontiera della ricerca. Questi approcci forniscono previsioni non solo più accurate, ma anche più robuste e con una quantificazione dell'incertezza più rigorosa rispetto ai modelli puramente basati sui dati. Essi riescono a mitigare i punti deboli di entrambi i mondi: la conoscenza fisica fornisce un "guard-rail" che impedisce al modello di ML di fare previsioni irrealistiche, mentre i dati reali permettono di

correggere e personalizzare le previsioni del modello fisico, che è spesso troppo idealizzato [22] [102].

In sintesi, la discussione dei risultati dipinge il ritratto di un campo di ricerca in una transizione vibrante ed eccitante. Un paradigma maturo, quello statistico, sta cedendo il passo al Machine Learning, che a sua volta sta vivendo una rivoluzione interna guidata dal Deep Learning. Emergono trend chiari verso una maggiore automazione e accuratezza, ma anche sfide cruciali legate all'interpretabilità e alla generalizzabilità. Le visioni più olistiche, come il Digital Twin e l'ibridazione, sono indicate come la destinazione finale di questo percorso, il punto di convergenza in cui la conoscenza fisica e quella basata sui dati collaboreranno in una sintesi ciber-fisica. La strada per raggiungere questa destinazione è ancora lunga, ma la direzione è tracciata.

# Capitolo 6: Conclusioni e Sviluppi Futuri

Il percorso intrapreso in questa tesi ci ha condotti attraverso un'esplorazione sistematica, approfondita e critica del vasto e dinamico campo della Manutenzione Predittiva. Partendo dalla definizione del problema e dalla sua crescente importanza strategica nel contesto della Quarta Rivoluzione Industriale, tracciando un'evoluzione metodologica che muove dalle solide ma rigide fondamenta della statistica classica, attraversa la rivoluzione pragmatica e potente del Machine Learning, e giunge infine alle frontiere più avanzate del Deep Learning e del paradigma del Digital Twin. Attraverso una revisione sistematica della letteratura, guidata dal protocollo PRISMA, si ha non solo descritto le singole tecniche in isolamento, ma son state anche classificate, confrontate e analizzate criticamente.

L'obiettivo di questo capitolo non è quello di presentare nuove informazioni o ulteriori analisi granulari, ma di compiere l'ultimo e più importante passo del processo di ricerca: quello della sintesi, della riflessione e della visione prospettica. Se i capitoli precedenti hanno fornito i dati, le analisi e le discussioni, questo capitolo si propone di distillarne l'essenza, di tirare le somme del percorso compiuto, di valutarne la portata e, infine, di guardare avanti, verso l'orizzonte delle possibilità future.

Lo scopo qui è triplice. In primo luogo, si intende sintetizzare i risultati chiave della ricerca. Questo va oltre un semplice riassunto. L'obiettivo è fornire risposte concise e basate sull'evidenza alle domande di ricerca che hanno guidato questa tesi fin dal suo inizio. Verranno riassunti i principali trend evolutivi, le metodologie dominanti, le aree di applicazione più studiate e le sfide più significative che sono emerse in modo ricorrente e robusto dall’analisi della letteratura.

Infine, si cercherà di delineare una roadmap per gli sviluppi futuri. Sulla base delle lacune, dei bias e delle sfide che son stati identificati nell’analisi. Si cercherà di rispondere alla domanda strategica: "Dove sta andando il campo della Manutenzione Predittiva e quali sono le sfide più urgenti, interessanti e di valore da affrontare per ricercatori e professionisti nei prossimi anni?".

## 6.1. Sintesi dei Risultati della Ricerca

Il percorso di ricerca intrapreso in questa tesi è stato guidato da una domanda principale (RQ) e tre sotto-domande (SQ) sequenziali. In questo paragrafo, si intende dare una risposta sintetica ma completa, basata sull'evidenza emersa dall'analisi della letteratura, consolidando così i risultati principali di questo lavoro.

La prima sotto-domanda (SQ1) chiedeva quali fossero le metodologie fondative e gli approcci statistici che hanno storicamente costituito le basi della Manutenzione Predittiva. L'analisi, presentata nel Capitolo 3, ha confermato che le fondamenta della disciplina poggiano su tre pilastri teorici. I primi sono i modelli di affidabilità, come la distribuzione di Weibull, che operano a livello di popolazione analizzando i tempi al guasto (*event data*) per definire strategie di sostituzione preventive. I secondi sono i modelli di analisi delle serie storiche, come i modelli ARIMA, che rappresentano il primo tentativo di prognosi individualizzata, estrapolando il trend di un indicatore di salute nel tempo. I terzi sono i processi stocastici, come i Modelli Nascosti di Markov, che hanno introdotto il concetto elegante di "stati di salute" discreti. Sebbene questi approcci siano ancora oggi utili in contesti di dati scarsi per la loro interpretabilità, l'analisi critica ha evidenziato i loro limiti strutturali: la rigidità, basata su assunzioni di linearità e stazionarietà, e la difficoltà nel gestire dati provenienti da più sensori contemporaneamente, si sono rivelate inadeguate a catturare la complessità dei moderni sistemi industriali. È proprio da questi limiti che è emersa la necessità scientifica di un nuovo paradigma.

La seconda sotto-domanda (SQ2) si interrogava su come il Machine Learning e il Digital Twin abbiano rivoluzionato l'approccio alla PdM. Il Capitolo 4 e l'analisi comparativa del Capitolo 5 hanno dimostrato che la rivoluzione è stata profonda e su più fronti. Il Machine Learning "classico", con algoritmi come le Support Vector Machines e i Random Forest, ha introdotto la capacità fondamentale di modellare relazioni non lineari complesse direttamente dai dati, portando a un significativo aumento dell'accuratezza predittiva. Tuttavia, la vera rottura paradigmatica è avvenuta con il Deep Learning. Architetture come le CNN e le LSTM hanno introdotto l'apprendimento *end-to-end*, una delle scoperte più significative dell'analisi. Questa capacità di apprendere le feature direttamente dai dati grezzi non solo automatizza un processo prima manuale e soggettivo, ma permette ai modelli di scoprire pattern che altrimenti rimarrebbero nascosti. Le LSTM, in particolare, sono emerse come lo strumento di riferimento per la stima della RUL, grazie alla loro abilità unica

di modellare le dipendenze temporali a lungo raggio. Il paradigma del Digital Twin, infine, ha rivoluzionato non tanto il singolo algoritmo, quanto l'architettura in cui questi algoritmi operano. Come discusso, il Digital Twin propone un framework olistico che integra modelli fisici e data-driven, permettendo non solo la previsione, ma anche la simulazione di scenari "what-if" e, in prospettiva, una manutenzione che non è più solo predittiva, ma prescrittiva [111].

La terza sotto-domanda (SQ3) mirava a identificare i principali trend applicativi, le sfide ricorrenti e le direzioni future. L'analisi tassonomica e bibliometrica del Capitolo 5 ha fatto emergere diversi trend chiari e critici. Dal punto di vista applicativo, ho identificato un forte e problematico bias di ricerca verso i cuscinetti volventi come caso di studio, a discapito di altri componenti altrettanto importanti come ingranaggi o utensili. Questo "problema del cuscinetto" è alimentato dalla disponibilità di dataset pubblici e rischia di limitare la generalizzabilità della ricerca. Per quanto riguarda le sfide ricorrenti, ne sono emerse principalmente due. La prima è il "gap di validità" tra i risultati eccellenti ottenuti su dati di laboratorio e la difficile performance su dati industriali reali, che sono rumorosi e non stazionari. Questo indica che la robustezza dei modelli nel mondo reale è una sfida ancora aperta [79]. La seconda, direttamente legata all'ascesa del Deep Learning, è la sfida dell'interpretabilità: la natura "black-box" di questi modelli rappresenta una barriera significativa alla loro adozione in contesti critici, dove la fiducia è fondamentale [1]. Le direzioni future più promettenti, come confermato dall'analisi, risiedono negli approcci ibridi (che fondono fisica e dati) e nell'implementazione di architetture Digital Twin complete, che rappresentano la frontiera più avanzata della ricerca [51].

In sintesi, rispondendo alla domanda di ricerca principale (RQ), si può affermare che lo stato dell'arte della Manutenzione Predittiva per le lavorazioni meccaniche è in uno stato di profonda transizione. La traiettoria evolutiva ha visto un chiaro e inesorabile spostamento dagli approcci statistici fondativi verso il paradigma del Machine Learning, guidato dalla ricerca di una maggiore accuratezza. All'interno di questo paradigma, una seconda ondata guidata dal Deep Learning sta spingendo verso l'automazione e l'apprendimento end-to-end. Tuttavia, questa evoluzione non è una semplice marcia trionfale, ma ha introdotto nuove e complesse sfide legate all'interpretabilità dei modelli, alla loro generalizzabilità al di fuori dei laboratori e alla loro integrazione in sistemi olistici come il Digital Twin. Il futuro della disciplina non risiede tanto nello sviluppo di algoritmi ancora più accurati, quanto nella capacità di risolvere queste nuove sfide, creando sistemi predittivi che siano non solo accurati, ma anche robusti, trasparenti e affidabili.

## 6.2. Implicazioni e Prospettive Future

L'analisi condotta in questa tesi ha dipinto il ritratto di un campo di ricerca in rapida trasformazione, guidato dalla potenza del Machine Learning ma ancora alle prese con le sfide della sua applicazione pratica. Sulla base delle evidenze raccolte, è ora possibile delineare le implicazioni di questi trend e le prospettive future più promettenti per la ricerca e per l'industria.

Una delle implicazioni più profonde è che il futuro della Manutenzione Predittiva non risiederà nella vittoria di un singolo paradigma, ma in una sintesi intelligente tra approcci diversi. La direzione più promettente è quella degli approcci ibridi, che combinano il rigore dei modelli fisici con la flessibilità dei modelli basati sui dati. Tecniche innovative come le Physics-Informed Neural Networks (PINN), che costringono i modelli di Deep Learning a rispettare le leggi fisiche, rappresentano una frontiera affascinante [86]. Dal punto di vista pratico, questo significa che le soluzioni di maggior valore saranno quelle capaci di integrare la conoscenza ingegneristica con l'analisi automatica dei dati.

Per quanto riguarda le applicazioni industriali, il potenziale di queste tecnologie si manifesta in scenari sempre più complessi. Nelle lavorazioni meccaniche ad alta precisione, come la fresatura CNC a 5 assi per componenti aerospaziali in titanio, un sistema di PdM avanzato non si limiterebbe a prevedere l'usura dell'utensile. Potrebbe, all'interno di un'architettura Digital Twin, analizzare le forze di taglio e le vibrazioni in tempo reale e regolare dinamicamente i parametri del processo (come velocità di avanzamento e profondità di passata) per compensare il degrado, garantendo il rispetto delle tolleranze micrometriche sul pezzo e massimizzando al contempo la vita utile dell'utensile, evitando rotture impreviste. Nelle linee di assemblaggio automotive, caratterizzate da centinaia di robot e stazioni interconnesse in una logica JIT, un guasto a un singolo robot di saldatura può fermare l'intera linea, con costi enormi. Un sistema di PdM olistico non solo monitorerebbe la salute del singolo robot prevedendo il degrado dei suoi giunti, ma analizzerebbe le interdipendenze dell'intera linea. Potrebbe, ad esempio, suggerire interventi di manutenzione prescrittivi da eseguire durante i brevi cambi turno, ri-allocando temporaneamente i compiti su robot adiacenti per non interrompere mai il flusso produttivo.

Questo potenziale si estende oltre la manifattura. Nel settore energetico, il gestore di un parco eolico potrebbe utilizzare la prognosi per decidere se far funzionare le turbine a pieno regime durante una tempesta (massimizzando i profitti) o limitarne la potenza per

preservarne la vita utile. Persino nel settore dei beni di consumo, il produttore di un elettrodomestico smart potrebbe utilizzare i dati raccolti sul campo per prevedere il guasto di un componente e inviare una notifica all'utente per prenotare un intervento tecnico prima che il guasto si manifesti, trasformando il servizio di assistenza da reattivo a proattivo e migliorando la soddisfazione del cliente.

Per realizzare pienamente queste applicazioni, tuttavia, la ricerca futura dovrà concentrarsi su alcune sfide cruciali emerse dalla nostra analisi. La prima direzione di ricerca fondamentale è quella dell'Explainable AI (XAI) per la PdM. La natura "black-box" dei modelli di Deep Learning è una barriera alla fiducia. La ricerca dovrà andare oltre le semplici tecniche post-hoc come LIME o SHAP, per sviluppare modelli intrinsecamente più trasparenti. Un filone molto promettente è quello dei modelli basati su meccanismi di attenzione, che possono non solo fornire una previsione, ma anche "evidenziare" visivamente quali parti del segnale di input o quali istanti temporali sono stati più importanti per quella decisione, fornendo un primo livello di "razionale" comprensibile per l'operatore [117]. Parallelamente, sarà cruciale sviluppare tecniche per una quantificazione robusta dell'incertezza, come le Reti Neurali Bayesiane, per fornire non solo una previsione, ma anche un livello di confidenza associato, indispensabile per la gestione del [57].

La seconda direzione di ricerca, altrettanto cruciale, riguarda la validazione e la robustezza dei modelli nel mondo reale, per colmare il "gap" tra laboratorio e fabbrica. Questo richiede un forte impulso nello sviluppo di tecniche di Transfer Learninge Domain Adaptation, per adattare modelli pre-addestrati su grandi quantità di dati a nuove macchine o a condizioni operative mutevoli con un minimo sforzo di ri-addestramento [79]. Una frontiera particolarmente interessante è il Federated Learning, un paradigma di apprendimento distribuito dove i modelli vengono addestrati in modo decentralizzato direttamente sui macchinari (*edge computing*), condividendo solo gli aggiornamenti dei modelli e non i dati grezzi. Questo approccio non solo migliora la robustezza, ma risolve anche cruciali problemi di privacy e sicurezza industriale, poiché i dati sensibili non lasciano mai l'impianto [53]. Ma, soprattutto, la sfida più grande sarà lo sviluppo di Digital Twin industriali validati, che possano agire come banchi di prova virtuali ad alta fedeltà. La sfida non è più solo costruire un modello analitico, ma creare un gemello digitale che sia una rappresentazione così accurata della realtà da poter essere usato per testare, validare e certificare gli algoritmi di PdM prima della loro implementazione sull'asset fisico, riducendo drasticamente i rischi e i costi del "go-live".

In conclusione, il futuro della Manutenzione Predittiva si preannuncia come un campo di straordinaria innovazione. Le sfide da affrontare sono significative, ma le potenzialità sono ancora più grandi. La transizione da una manutenzione reattiva a una prognostica prescrittiva e autonoma, orchestrata all'interno di ecosistemi Digital Twin, promette di trasformare radicalmente l'efficienza e la resilienza dei sistemi industriali, e questa tesi, attraverso la sua analisi, ha cercato di fornire una mappa per navigare in questo percorso.

# Bibliografia


[1] Adadi and M. Berrada, “Peeking Inside the Black-Box: A Survey on Explainable Artificial Intelligence (XAI)”, *IEEE Access*, vol. 6, pp. 52138–52160, 2018.

[2] P. S. Ahuja and J. S. Khamba, “Total productive maintenance: literature review and directions”, *International Journal of Quality & Reliability Management*, vol. 25, no. 7, pp. 709–739, 2008.

[3] P. Aivaliotis, K. Georgoulias, and G. Chryssolouris, “A framework for prognostics and health management of manufacturing systems based on the digital twin concept”, *Procedia CIRP*, vol. 81, pp. 951–956, 2019.

[4] D. An, J. H. Choi, and N. H. Kim, “Practical options for selecting data-driven or physics-based prognostics algorithms with reviews”, *Reliability Engineering & System Safety*, vol. 133, pp. 223–236, Jan. 2015.

[5] D. An, J. H. Choi, and N. H. Kim, “Prognostics 101: A tutorial for particle filter-based prognostics”, *Reliability Engineering & System Safety*, vol. 174, pp. 1–13, Jun. 2018.

[6] K. D. Bailey, *Typologies and Taxonomies: An Introduction to Classification Techniques*. Thousand Oaks, CA, USA: Sage Publications, 1994.

[7] B. R. Barricelli, E. Casiraghi, and D. Fogli, “A survey on digital twin: Definitions, characteristics, applications, and design implications”, *IEEE Access*, vol. 7, pp. 167653–167671, 2019.

[8] R. E. Bellman, *Adaptive Control Processes: A Guided Tour*. Princeton, NJ, USA: Princeton University Press, 1961.

[9] Y. Bengio, P. Simard, and P. Frasconi, “Learning long-term dependencies with gradient descent is difficult”, *IEEE Transactions on Neural Networks*, vol. 5, no. 2, pp. 157–166, Mar. 1994.

[10] C. M. Bishop, *Pattern Recognition and Machine Learning*. New York, NY, USA: Springer, 2006.

[11] T. Boutros and M. Liang, “A multi-sensor fusion based prognostic approach for rolling element bearings”, *Mechanical Systems and Signal Processing*, vol. 25, no. 6, pp. 2095–2108, Aug. 2011.

[12] G. E. P. Box, G. M. Jenkins, and G. C. Reinsel, *Time Series Analysis: Forecasting and Control*, 4th ed. Hoboken, NJ, USA: John Wiley & Sons, 2008.

[13] L. Breiman, “Random forests”, *Machine Learning*, vol. 45, no. 1, pp. 5–32, Oct. 2001.

[14] L. Breiman, “Statistical modeling: The two cultures”, *Statistical Science*, vol. 16, no. 3, pp. 199–215, Aug. 2001.

[15] J. B. Buckheit and D. L. Donoho, “Wavelab and reproducible research”, in *Wavelets and Statistics*, A. Antoniadis and G. Oppenheim, Eds. New York, NY, USA: Springer, 1995, pp. 55–81.

[16] J. Bulla, *Hidden Markov Models for Time Series: An Introduction Using R*. Boca Raton, FL, USA: CRC Press, 2011.

[17] M. Canizo, E. Onieva, and A. Conde, “A context-aware machine learning framework for the prognostics of a turbofan engine”, *Sensors*, vol. 19, no. 9, p. 2110, May 2019.

[18] T. P. Carvalho *et al.*, “A systematic literature review of machine learning methods applied to predictive maintenance”, *Computers & Industrial Engineering*, vol. 137, p. 106024, Nov. 2019.

[19] V. Chandola, A. Banerjee, and V. Kumar, “Anomaly detection: A survey”, *ACM Computing Surveys*, vol. 41, no. 3, art. 15, pp. 1–58, Jul. 2009.

[20] O. Chapelle, B. Schölkopf, and A. Zien, Eds., *Semi-Supervised Learning*. Cambridge, MA, USA: MIT Press, 2006.

[21] K. Cho *et al.*, “Learning phrase representations using RNN encoder-decoder for statistical machine translation”, in *Proc. EMNLP*, 2014, pp. 1724–1734.

[22] M. Corbetta, C. Sbarufatti, and A. Manes, “A Bayesian framework for the hybridization of physics-based and data-driven models in prognostics”, *Journal of Manufacturing Systems*, vol. 52, pp. 14–27, Jul. 2019.

[23] D. R. Cox, “Regression Models and Life-Tables”, *Journal of the Royal Statistical Society: Series B (Methodological)*, vol. 34, no. 2, pp. 187–220, 1972.

[24] G. Cybenko, “Approximation by superpositions of a sigmoidal function”, *Mathematics of Control, Signals and Systems*, vol. 2, no. 4, pp. 303–314, Dec. 1989.

[25] W. E. Deming, *Out of the Crisis*. Cambridge, MA, USA: MIT Press, 1986.

[26] M. Dong and D. He, “A segmental hidden semi-Markov model (HSMM)-based diagnostics and prognostics framework and methodology”, *Mechanical Systems and Signal Processing*, vol. 21, no. 5, pp. 2248–2266, Jul. 2007.

[27] M. Elforjani and D. Mba, “Diagnosing and prognosticating the health of a multi-stage gearbox using acoustic emission”, *Engineering Fracture Mechanics*, vol. 78, no. 1, pp. 88–102, Jan. 2011.

[28] M. Ester, H.-P. Kriegel, J. Sander, and X. Xu, “A density-based algorithm for discovering clusters in large spatial databases with noise”, in *Proc. 2nd Int. Conf. on Knowledge Discovery and Data Mining (KDD-96)*, 1996, pp. 226–231.

[29] A. V. Feigenbaum, “Total Quality Control”, *Harvard Business Review*, vol. 34, no. 6, pp. 93–101, Nov. 1956.

[30] S. J. Ford, “The nature of innovation”, Imperial College London, Tanaka Business School, London, UK, Tech. Rep., 2011.

[31] A. Fuller, M. Helu, and S. Rachuri, “A proposed framework for a Digital Twin in manufacturing”, *Procedia Manufacturing*, vol. 45, pp. 29–34, 2020.

[32] J. Gama, I. Žliobaitė, A. Bifet, M. Pechenizkiy, and A. Bouchachia, “A survey on concept drift adaptation”, *ACM Computing Surveys*, vol. 46, no. 4, art. 44, pp. 1–37, Mar. 2014.

[33] E. H. Glaessgen and D. S. Stargel, “The digital twin paradigm for future NASA and U.S. Air Force vehicles”, in *53rd AIAA/ASME/ASCE/AHS/ASC Structures, Structural Dynamics and Materials Conference*, Honolulu, HI, USA, 2012, p. 1818.

[34] K. Goebel, B. Saha, and A. Saxena, “A comparison of three data-driven-based-prognostics techniques”, in *Proc. Int. Conf. on Prognostics and Health Management*, Denver, CO, USA, 2008, pp. 1–12.

[35] I. Goodfellow, Y. Bengio, and A. Courville, *Deep Learning*. Cambridge, MA, USA: MIT Press, 2016.

[36] I. Goodfellow *et al.*, “Generative adversarial nets”, in *Advances in Neural Information Processing Systems 27*, 2014, pp. 2672–2680.

[37] B. Gou, Y. Li, and T. Zhang, “A comprehensive review of the application of machine learning in the field of predictive maintenance”, *Reliability Engineering & System Safety*, vol. 203, p. 107085, Nov. 2020.

[38] M. Grieves, “Digital Twin: Manufacturing Excellence through Virtual Factory Replication”, Digital Twin Consortium, White Paper, 2014.

[39] M. Grieves and J. Vickers, “Digital Twin: Mitigating Unpredictable, Undesirable Emergent Behavior in Complex Systems”, in *Transdisciplinary Perspectives on Complex Systems*, F. J. Kahlen, S. Flumerfelt, and A. Alves, Eds. Cham, Switzerland: Springer, 2017, pp. 85–113.

[40] T. Hastie, R. Tibshirani, and J. Friedman, *The Elements of Statistical Learning: Data Mining, Inference, and Prediction*, 2nd ed. New York, NY, USA: Springer, 2009.

[41] D. He and W. He, "A novel framework for bearing remaining useful life prediction based on an adaptive filtering and a two-layer neural network", *IEEE Transactions on Instrumentation and Measurement*, vol. 66, no. 11, pp. 2848–2859, Nov. 2017.

[42] D. He, R. Li, and Y. Li, "A review on prognostics methods for lithium-ion battery", *Journal of Power Sources*, vol. 482, p. 228941, Jan. 2021.

[43] M. Henrion, "The tyranny of the benchmark: An essay on the rise of machine learning and the decline of the statistical worldview", *The American Statistician*, vol. 72, no. 2, pp. 115–120, Apr. 2018.

[44] S. Hochreiter and J. Schmidhuber, "Long short-term memory", *Neural Computation*, vol. 9, no. 8, pp. 1735–1780, Nov. 1997.

[45] K. Hornik, M. Stinchcombe, and H. White, "Multilayer feedforward networks are universal approximators", *Neural Networks*, vol. 2, no. 5, pp. 359–366, 1989.

[46] T. Ince, S. Kiranyaz, L. Eren, M. Askar, and M. Gabbouj, "Real-time motor fault detection by 1-D convolutional neural networks", *IEEE Transactions on Industrial Electronics*, vol. 63, no. 11, pp. 7067–7075, Nov. 2016.

[47] A. K. Jain, "Data clustering: 50 years beyond K-means", *Pattern Recognition Letters*, vol. 31, no. 8, pp. 651–666, Jun. 2010.

[48] A. K. S. Jardine, D. Lin, and D. Banjevic, "A review on machinery diagnostics and prognostics implementing condition-based maintenance", *Mechanical Systems and Signal Processing*, vol. 20, no. 7, pp. 1483–1510, Oct. 2006.

[49] D. Jones, C. Snider, A. Nassehi, J. Yon, and B. Hicks, "Characterising the Digital Twin: A systematic literature review", *CIRP Journal of Manufacturing Science and Technology*, vol. 29, pp. 36–52, May 2020.

[50] M. I. Jordan and T. M. Mitchell, "Machine learning: Trends, perspectives, and prospects", *Science*, vol. 349, no. 6245, pp. 255–260, Jul. 2015.

[51] M. G. Kapteyn, D. J. Pretorius, and K. E. Willcox, "A probabilistic graphical model for data-driven, physics-based, and hybrid prognostics", *Mechanical Systems and Signal Processing*, vol. 159, p. 107779, Oct. 2021.

[52] B. Kitchenham and S. Charters, "Guidelines for performing Systematic Literature Reviews in Software Engineering", Keele University and University of Durham, UK, Tech. Rep. EBSE-2007-01, 2007.

[53] J. Konečný, H. B. McMahan, F. X. Yu, P. Richtárik, A. T. Suresh, and D. Bacon, "Federated learning: Strategies for improving communication efficiency", *arXiv preprint arXiv:1610.05492*, 2016.

[54] W. Kritzinger, M. Karner, G. Traar, J. Henjes, and W. Sihn, "Digital Twin in manufacturing: A categorical literature review and classification", *IFAC-PapersOnLine*, vol. 51, no. 11, pp. 1016–1022, 2018.

[55] A. Krizhevsky, I. Sutskever, and G. E. Hinton, "ImageNet classification with deep convolutional neural networks", in *Advances in Neural Information Processing Systems 25*, 2012, pp. 1097–1105.

[56] T. S. Kuhn, *The Structure of Scientific Revolutions*. Chicago, IL, USA: University of Chicago Press, 1962.

[57] B. Lakshminarayanan, A. Pritzel, and C. Blundell, "Simple and scalable predictive uncertainty estimation using deep ensembles", in *Advances in Neural Information Processing Systems 30*, 2017, pp. 6402–6413.

[58] J. F. Lawless, *Statistical Models and Methods for Lifetime Data*, 2nd ed. Hoboken, NJ, USA: John Wiley & Sons, 2011.

[59] Y. LeCun, Y. Bengio, and G. Hinton, "Deep learning", *Nature*, vol. 521, no. 7553, pp. 436–444, May 2015.

[60] E. A. Lee, H. A. Kao, and S. Yang, "Cyber-physical systems", in *The Industrial Information Technology Handbook*, R. Zurawski, Ed. Boca Raton, FL, USA: CRC Press, 2014.

[61] J. Lee, B. Bagheri, and H. A. Kao, "A Cyber-Physical Systems architecture for Industry 4.0-based manufacturing systems", *Manufacturing Letters*, vol. 3, pp. 18–23, Jan. 2015.

[62] J. Lee *et al.*, "Prognostics and health management design for rotary machinery systems—Reviews, methodology and applications", *Mechanical Systems and Signal Processing*, vol. 42, no. 1–2, pp. 314–334, Jan. 2014.

[63] Y. Lei *et al.*, "Applications of machine learning to machine fault diagnosis: A review and roadmap", *Mechanical Systems and Signal Processing*, vol. 138, p. 106587, Apr. 2020.

[64] F. T. Liu, K. M. Ting, and Z. H. Zhou, "Isolation forest", in *Proc. 8th IEEE Int. Conf. on Data Mining*, 2008, pp. 413–422.

[65] Z. Liu, N. Meyendorf, and N. Mrad, "The role of digital twin in structural health management: A survey", *Sensors*, vol. 21, no. 3, p. 859, Jan. 2021.

[66] T. Loutas, D. Roulias, and G. Georgoulas, “Remaining useful life estimation in rolling bearings utilizing data-driven and hybrid approaches”, *Mechanical Systems and Signal Processing*, vol. 40, no. 1, pp. 193–212, Oct. 2013.

[67] A. McAfee and E. Brynjolfsson, “Big data: The management revolution”, *Harvard Business Review*, vol. 90, no. 10, pp. 60–68, Oct. 2012.

[68] W. Q. Meeker and L. A. Escobar, *Statistical Methods for Reliability Data*. New York, NY, USA: John Wiley & Sons, 1998.

[69] T. M. Mitchell, *Machine Learning*. New York, NY, USA: McGraw-Hill, 1997.

[70] R. K. Mobley, *An Introduction to Predictive Maintenance*, 2nd ed. Woburn, MA, USA: Butterworth-Heinemann, 2002.

[71] D. Moher, A. Liberati, J. Tetzlaff, D. G. Altman, and The PRISMA Group, “Preferred Reporting Items for Systematic Reviews and Meta-Analyses: The PRISMA Statement”, *PLoS Medicine*, vol. 6, no. 7, p. e1000097, Jul. 2009.

[72] D. C. Montgomery, *Introduction to Statistical Quality Control*, 6th ed. Hoboken, NJ, USA: John Wiley & Sons, 2009.

[73] P. Muchiri, L. Pintelon, H. Martin, and A. De Meyer, “A review of maintenance performance measurement: A conceptual framework and research agenda”, *Journal of Quality in Maintenance Engineering*, vol. 17, no. 2, pp. 116–137, 2011.

[74] S. Nandi, H. A. Toliyat, and X. Li, “Condition monitoring and fault diagnosis of electrical motors—a review”, *IEEE Transactions on Energy Conversion*, vol. 20, no. 4, pp. 719–729, Dec. 2005.

[75] P. Nectoux *et al.*, “PRONOSTIA: An experimental platform for bearings accelerated degradation tests”, in *IEEE Int. Conf. on Prognostics and Health Management*, Denver, CO, USA, 2012, pp. 1–8.

[76] R. C. Nickerson, U. Varshney, and J. Muntermann, “A method for taxonomy development and its application in information systems”, *European Journal of Information Systems*, vol. 22, no. 3, pp. 336–359, May 2013.

[77] B. A. Nosek *et al.*, “Promoting an open research culture”, *Science*, vol. 348, no. 6242, pp. 1422–1425, Jun. 2015.

[78] M. E. Orchard and G. J. Vachtsevanos, “A particle-filtering approach for on-line fault diagnosis and failure prognosis”, *Transactions of the Institute of Measurement and Control*, vol. 31, no. 3–4, pp. 221–246, Jun. 2009.

[79] S. J. Pan and Q. Yang, “A survey on transfer learning”, *IEEE Transactions on Knowledge and Data Engineering*, vol. 22, no. 10, pp. 1345–1359, Oct. 2010.

[80] M. Pecht, *Prognostics and Health Management of Electronics*. Hoboken, NJ, USA: John Wiley & Sons, 2008.

[81] D. Peña, G. C. Tiao, and R. S. Tsay, *A Course in Time Series Analysis*. Hoboken, NJ, USA: John Wiley & Sons, 2011.

[82] Z. K. Peng, P. W. Tse, and F. L. Chu, "A comparison study of improved Hilbert–Huang transform and wavelet transform: Application to fault diagnosis for rolling bearing", *Mechanical Systems and Signal Processing*, vol. 19, no. 5, pp. 974–988, Sep. 2005.

[83] M. Petticrew and H. Roberts, *Systematic Reviews in the Social Sciences: A Practical Guide*. Malden, MA, USA: John Wiley & Sons, 2008.

[84] L. Pintelon and L. N. Van Wassenhove, "A maintenance management tool", *Omega*, vol. 18, no. 1, pp. 59–69, 1990.

[85] L. R. Rabiner, "A tutorial on hidden Markov models and selected applications in speech recognition", *Proceedings of the IEEE*, vol. 77, no. 2, pp. 257–286, Feb. 1989.

[86] M. Raissi, P. Perdikaris, and G. E. Karniadakis, "Physics-informed neural networks: A deep learning framework for solving forward and inverse problems involving nonlinear partial differential equations", *Journal of Computational Physics*, vol. 378, pp. 686–707, Feb. 2019.

[87] Y. Ran, X. Zhou, and P. Lin, "A Survey of Predictive Maintenance: Systems, Purposes and Approaches", *IEEE Communications Surveys & Tutorials*, vol. 21, no. 1, pp. 75–94, 1st Quart., 2019.

[88] R. B. Randall, *Vibration-based Condition Monitoring: Industrial, Aerospace and Automotive Applications*. Chichester, UK: John Wiley & Sons, 2011.

[89] R. B. Randall and J. Antoni, "Rolling element bearing diagnostics—A tutorial", *Mechanical Systems and Signal Processing*, vol. 25, no. 2, pp. 485–520, Feb. 2011.

[90] R. Rosen, G. von Wichert, G. Lo, and K. D. Bettenhausen, "About the importance of autonomy and digital twins for the future of manufacturing", *IFAC-PapersOnLine*, vol. 48, no. 3, pp. 567–572, 2015.

[91] C. Rudin, "Stop explaining black box machine learning models for high stakes decisions and use interpretable models instead", *Nature Machine Intelligence*, vol. 1, no. 5, pp. 206–215, May 2019.

[92] S. J. Russell and P. Norvig, *Artificial Intelligence: A Modern Approach*, 3rd ed. Upper Saddle River, NJ, USA: Pearson, 2016.

[93] W. Samek, T. Wiegand, and K. R. Müller, “Explainable artificial intelligence: understanding, visualizing and interpreting deep learning models”, *ITU Journal: ICT Discoveries*, vol. 1, no. 1, pp. 39–48, 2017.

[94] B. Samanta, “Gear fault detection using artificial neural networks and support vector machines with genetic algorithms”, *Mechanical Systems and Signal Processing*, vol. 18, no. 3, pp. 625–644, May 2004.

[95] S. Särkkä, *Bayesian Filtering and Smoothing*. Cambridge, UK: Cambridge University Press, 2013.

[96] A. Saxena and K. Goebel, Eds., “PHM08 data challenge competition”, NASA Ames Prognostics Data Repository, Moffett Field, CA, USA, 2008.

[97] C. Scheffer and P. S. Heyns, *Wear Debris Analysis Handbook*. Binsted, UK: Coxmoor Publishing, 2008.

[98] B. Schölkopf, J. C. Platt, J. Shawe-Taylor, A. J. Smola, and R. C. Williamson, “Estimating the support of a high-dimensional distribution”, *Neural Computation*, vol. 13, no. 7, pp. 1443–1471, Jul. 2001.

[99] S. Schork and M. Kucera, “Prescriptive Maintenance in Industry 4.0: A literature review”, *Procedia CIRP*, vol. 104, pp. 933–938, 2021.

[100] A. Sharma, G. S. Yadava, and S. G. Deshmukh, “A literature review and future perspectives on maintenance optimization”, *Journal of Quality in Maintenance Engineering*, vol. 17, no. 1, pp. 5–25, 2011.

[101] W. A. Shewhart, *Economic Control of Quality of Manufactured Product*. New York, NY, USA: D. Van Nostrand Company, 1931.

[102] S. Siahpour, A. Bertin, and B. Kamsu-Foguem, “A systematic review of hybrid physics-informed neural networks for industrial prognostics”, *Computers in Industry*, vol. 133, p. 103554, Dec. 2021.

[103] X. S. Si, W. Wang, C. H. Hu, and D. H. Zhou, “Remaining useful life estimation–A review on the statistical data-driven approaches”, *European Journal of Operational Research*, vol. 213, no. 1, pp. 1–14, Aug. 2011.

[104] W. A. Smith and R. B. Randall, “Rolling element bearing diagnostics using the Case Western Reserve University data: A benchmark study”, *Mechanical Systems and Signal Processing*, vol. 64–65, pp. 100–121, Dec. 2015.

[105] A. Soualhi, K. Medjaher, and N. Zerhouni, “Bearing health monitoring based on Hilbert–Huang transform, support vector machine, and regression”, *IEEE*

*Transactions on Instrumentation and Measurement*, vol. 64, no. 1, pp. 52–62, Jan. 2015.

[106] R. Stark, T. Damerau, and S. Kind, "Industrial-scale digital twins", in *The Digital Twin*, R. Stark, T. Damerau, and S. Kind, Eds. Cham, Switzerland: Springer, 2019, pp. 95–125.

[107] V. Stodden, S. Miguez, and J. Witzel, Eds., *The Practice of Reproducible Research: Case Studies and Lessons from the Data-Intensive Sciences*. Cham, Switzerland: Springer Nature, 2020.

[108] E. Strubell, A. Ganesh, and A. McCallum, "Energy and policy considerations for deep learning in NLP", in *Proc. 57th Annual Meeting of the Association for Computational Linguistics*, 2019, pp. 3645–3650.

[109] R. S. Sutton and A. G. Barto, *Reinforcement Learning: An Introduction*, 2nd ed. Cambridge, MA, USA: MIT Press, 2018.

[110] L. Swanson, "Linking maintenance strategies to performance", *International Journal of Production Economics*, vol. 70, no. 3, pp. 237–244, Apr. 2001.

[111] F. Tao *et al.*, "Digital Twin in Industry: State-of-the-Art", *IEEE Transactions on Industrial Informatics*, vol. 15, no. 4, pp. 2405–2415, Apr. 2019.

[112] J. Tidd, "A review of innovation models", Imperial College London, Tanaka Business School, London, UK, Tech. Rep., 2006.

[113] R. J. Torraco, "Writing integrative literature reviews: Guidelines and examples", *Human Resource Development Review*, vol. 4, no. 3, pp. 356–367, Sep. 2005.

[114] D. Tranfield, D. Denyer, and P. Smart, "Towards a methodology for developing evidence-informed management knowledge by means of systematic review", *British Journal of Management*, vol. 14, no. 3, pp. 207–222, Sep. 2003.

[115] N. J. van Eck and L. Waltman, "Software survey: VOSviewer, a computer program for bibliometric mapping", *Scientometrics*, vol. 84, no. 2, pp. 523–538, Aug. 2010.

[116] V. N. Vapnik, *The Nature of Statistical Learning Theory*. New York, NY, USA: Springer-Verlag, 1995.

[117] A. Vaswani *et al.*, "Attention is all you need", in *Advances in Neural Information Processing Systems 30*, 2017, pp. 5998–6008.

[118] J. Webster and R. T. Watson, "Analyzing the Past to Prepare for the Future: Writing a Literature Review", *MIS Quarterly*, vol. 26, no. 2, pp. xiii–xxiii, Jun. 2002.

[119] W. Weibull, “A statistical distribution function of wide applicability”, *Journal of Applied Mechanics*, vol. 18, no. 3, pp. 293–297, Sep. 1951.

[120] L. Wen, X. Li, L. Gao, and Y. Zhang, “A new convolutional neural network-based data-driven fault diagnosis method”, *IEEE Transactions on Industrial Electronics*, vol. 65, no. 7, pp. 5990–5998, Jul. 2018.