



POLITECNICO DI TORINO

**Corso di Laurea Magistrale
in Ingegneria Gestionale**

Tesi di Laurea Magistrale

**Intelligenza artificiale e
gestione dell'incertezza
interpretativa attraverso il
Data Quality Management:
un'applicazione nel settore
assicurativo**

Relatore

Prof. Emilio Paolucci

Correlatori esterni

Prof.ssa Claudia Franzè

Ing. Davide Giuseppe

Bruno

Candidato

Riccardo Canali

ANNO ACCADEMICO 2023-2024

Ringraziamenti

Nel conseguimento di questo significativo traguardo del mio percorso accademico, desidero esprimere la mia più profonda gratitudine a tutti coloro che hanno contribuito alla realizzazione del presente lavoro.

Un ringraziamento sentito va al Professor Emilio Paolucci, la cui competenza e passione per l'innovazione hanno rappresentato per me una fonte inesauribile di ispirazione durante l'intero percorso di Laurea Magistrale. Conoscerla mi ha permesso di comprendere come il ruolo dell'Ingegnere non richieda solo preparazione e conoscenza, ma soprattutto una costante curiosità e apertura verso l'apprendimento.

Desidero, inoltre, esprimere un ringraziamento al gruppo che ha seguito l'evoluzione di questo lavoro, apportando suggerimenti, feedback preziosi e continuo incoraggiamento. In particolare, voglio ringraziare in modo speciale l'Ing. Simone Trincherò il cui supporto preciso e i preziosi consigli hanno arricchito notevolmente questo elaborato, aiutandomi a superare le sfide incontrate lungo il cammino.

Indice

Abstract	5
Capitolo 1 – Introduzione	6
1.1 Struttura della tesi	7
Capitolo 2 – Introduzione dell’Intelligenza Artificiale nelle organizzazioni burocratiche e riconfigurazione dei processi decisionali	8
2.1 Stato dell’arte dell’Intelligenza Artificiale nelle organizzazioni burocratiche	8
2.2 Processi decisionali e Intelligenza Artificiale.....	12
2.3 Introduzione dell’Intelligenza Artificiale nel settore assicurativo	16
Capitolo 3 – Data Quality Management	19
3.1 Definizioni e caratteristiche dei dati	19
3.2 Caratteristiche del processo di DQM in riferimento al settore assicurativo	22
3.3 DQM per il Machine Learning: importanza del Data Cleaning	24
3.4 Tempo e fattori chiave nella preparazione dei dati per il Machine Learning	28
3.5 Acquisizione e trasferimento del know how.....	30
3.6 Influenza del DQM sulle prestazioni attese di un modello di Machine Learning	33
Capitolo 4 – Progetto D.A.R.E	38
4.1 Descrizione progetto	38
4.2 Requisiti per lo sviluppo di un algoritmo di IA	40
4.3 Framework per il training e test dell’IA nel progetto	43
4.4 Metodologia per il controllo dell’incertezza interpretativa e per il data cleaning	43
Capitolo 5 – Analisi dei dati peritali	51
Analisi dei dati peritali.....	51
5.1 Risultati e statistiche del campione analizzato di 200 perizie (foto)	51
5.1.1 Analisi del campione analizzato di 200 perizie (foto).....	52
5.2 Confronto tra i due network peritali.....	53
5.2.1 Risultati analitici e statistiche del campione analizzato (foto)	53
5.2.2 Analisi analitica del campione analizzato (foto): confronto tra i due network.....	54
5.2.3 Osservazioni analitiche sul campione analizzato (testo perizia con importo liquidazione).....	55
5.3 Risultati e statistiche del dataset complessivo di 700 perizie (foto)	57
Capitolo 6 – Discussione dei risultati peritali	59
Capitolo 7 – Esperimento per valutaione comparativa	62
7.1 Introduzione	62

7.2 Risorse e strumenti operativi	64
7.3 Criteri e metriche di valutazione.....	65
7.4 Metodologia generale.....	68
7.5 Metodologia dell’esperimento pilota	70
7.6 Risultati dell’esperimento pilota	72
7.7 Osservazioni finali e sviluppi futuri.....	77
Capitolo 8 – Conclusioni e open points	78
Elenco delle figure	80
Elenco delle tabelle	81
Bibliografia	82

Abstract

La presente tesi si propone di illustrare gli aspetti più significativi relativi all'implementazione dell'Intelligenza Artificiale (IA) all'interno dei processi decisionali con uno specifico focus su un caso di studio riguardante la gestione dei sinistri assicurativi del segmento "non auto". L'obiettivo principale che ci si pone è quello di analizzare come l'incertezza interpretativa dei dati possa essere controllata attraverso una gestione efficace della qualità dei dati stessi, valutando altresì l'impatto di tale gestione sull'addestramento dei modelli di IA.

Il caso studio ha costituito, inoltre, la base per il design di un esperimento di valutazione comparativa volto a esplorare il potenziale miglioramento delle decisioni derivante dall'uso di strumenti IA e a esaminare la complementarità tra le capacità umane e tali strumenti.

Capitolo 1

Introduzione

L'avanzamento dell'Intelligenza Artificiale (IA) e la sua diffusione stanno trasformando in modo profondo le dinamiche aziendali, influenzando significativamente non solo le routine operative, ma anche le competenze professionali richieste e i processi decisionali. Sebbene l'IA offra un potenziale per migliorare l'efficienza e l'accuratezza, grazie alla capacità di processare volumi di dati notevolmente elevati con velocità e precisione rispetto agli esseri umani, la sua effettiva integrazione con persone, sistemi e processi consolidati incontra, oltre ai significativi investimenti finanziari e di tempo, numerose difficoltà che possono limitarne la piena realizzazione dei benefici attesi.

Le differenze tra l'elaborazione delle informazioni da parte dell'IA e quella umana rappresentano infatti un aspetto rilevante: l'IA, come detto può gestire e analizzare ingenti quantità di dati in modo rapido e preciso, ma gli esseri umani, sebbene possano produrre errori e distorsioni, mostrano maggiore adattabilità in contesti complessi o dove le informazioni sono scarse. Questa capacità umana di interpretare e reagire a situazioni impreviste conferisce un valore aggiunto che l'IA, per quanto avanzata, non è sempre in grado di replicare pienamente.

Per sfruttare al meglio le potenzialità offerte dall'IA, le organizzazioni devono implementare strategie mirate, che non solo rispondano ai requisiti tecnici e normativi imposti da questi sistemi, ma che considerino anche la centralità dei *dati*, risorsa fondamentale per alimentare e potenziare tale tecnologia. Infatti, per realizzare un'efficace integrazione nei processi decisionali è necessario trasformare i dati grezzi a disposizione in risorse strutturate, al fine di ridurre l'incertezza interpretativa e superare le limitazioni dei modelli IA nell'acquisizione e nella riproduzione della conoscenza tacita, spesso esclusiva delle persone e delle organizzazioni.

In questo contesto, la presente tesi si propone di andare a colmare un gap nella letteratura, relativo al modo in cui i dati possano essere raccolti, analizzati, selezionati, trattati e combinati tra loro per l'addestramento di un modello di Machine Learning finalizzato alla classificazione di immagini, ovvero un modello di Computer Vision.

Inoltre, a seguito dell'analisi dei dati relativi al caso studio nel settore assicurativo, è altresì emersa la necessità di strutturare un esperimento di valutazione comparativa. Tale esperimento ha l'obiettivo di verificare se strumenti basati su IA possano effettivamente contribuire a migliorare i processi decisionali in ambiti ad alta variabilità e complessità, offrendo un complemento valido al lavoro umano e ottimizzando l'interazione tra risorse umane e algoritmi.

1.1 Struttura della tesi

Il presente lavoro di tesi è articolato in tre parti principali, strettamente interconnesse tra loro.

La *prima parte* che comprende i capitoli 2 e 3 fornisce una panoramica della letteratura esistente e dei suoi principali contenuti e più specificatamente il capitolo 2 approfondisce, in un primo momento, le caratteristiche generali delle organizzazioni burocratiche, successivamente il focus si sposta sul mutamento dei processi decisionali in rapporto al crescente ruolo dell'Intelligenza Artificiale e sull'influenza che questa può esercitare nel settore assicurativo, caratterizzato, appunto, da struttura burocratica. Il capitolo 3, invece, si concentra sull'importanza del processo di gestione della qualità dei dati (Data Quality Management), esaminando sia l'aspetto "macro" (archiviazione, elaborazione e fruizione dei dati) sia l'aspetto "micro", relativo al caso studio che verrà trattato più avanti, ovvero lo sviluppo di un modello di Machine Learning e, in particolare, di Computer Vision.

La *seconda parte*, costituita dai capitoli 4, 5 e 6 presenta nel dettaglio il caso studio nel settore assicurativo (progetto D.A.R.E.) e affronta l'introduzione dell'Intelligenza Artificiale in tale ambito, con particolare attenzione ai processi sui dati per controllare l'incertezza interpretativa e funzionali allo sviluppo di algoritmi IA. Questa sezione include un'approfondita analisi dei dati e una discussione dei risultati ottenuti.

La *terza parte*, rappresentata dal solo capitolo 7, illustra il design e i risultati ottenuti di un esperimento di valutazione comparativa volto a verificare se l'uso di strumenti IA possa migliorare la qualità del processo decisionale, sia da parte di esperti che di non esperti del settore assicurativo, nella gestione di un sinistro attraverso un'analisi fotografica.

Infine, la tesi si conclude con un capitolo dedicato alle conclusioni.

Capitolo 2

Introduzione dell'Intelligenza Artificiale nelle organizzazioni burocratiche e riconfigurazione dei processi decisionali

2.1 Stato dell'arte dell'Intelligenza Artificiale nelle organizzazioni burocratiche

Molte aziende in diversi settori stanno vivendo un profondo cambiamento tecnologico, spesso definito “era degli algoritmi”, che sta trasformando i processi lavorativi e ridefinendo il futuro del lavoro stesso.

Questo cambiamento è trainato dall'avanzamento e dall'implementazione sempre più diffusa dell'Intelligenza Artificiale (IA) la quale, col tempo, sarà sempre più integrata all'interno delle organizzazioni modificando in modo significativo routine, competenze e processi decisionali.

L'IA avrà un impatto non indifferente, soprattutto, in quelle organizzazioni con struttura burocratica, sia pubbliche che private: in queste organizzazioni, infatti, le rigidità operative e procedurali rappresentano elementi di vulnerabilità ad esigenze derivanti dall'ambiente esterno che richiedono risposte repentine.

Caratteristiche delle organizzazioni burocratiche

Uno studio sistematico delle organizzazioni burocratiche venne formalizzato nel corso del Novecento dal sociologo Max Weber, il quale le descriveva come la forma più efficace e razionale per la gestione e l'amministrazione di organizzazioni di grandi dimensioni e complessità.

Weber, in relazione a tale tipologia di organizzazione, identificava le seguenti caratteristiche:

- Specializzazione e divisione del lavoro, in gran parte semplice e ripetitiva.
- Struttura gerarchica, in cui ogni soggetto ha un superiore e, al contempo, ne supervisiona altri.
- Procedure formalizzate e standardizzate con una chiara divisione dei compiti.
- Regole e procedure scritte e prestabilite.

- Personale altamente qualificato, fedele all'organizzazione.

Questo modello organizzativo permette di sostituire il lavoro generico con il lavoro specializzato e garantire precisione, rapidità, chiarezza e continuità nello svolgimento dei compiti al fine di gestire e ridurre l'incertezza dell'azione umana. Due elementi chiave sono, infatti, la razionalizzazione e la standardizzazione che garantiscono interconnessione e complementarità delle attività migliorando sensibilmente l'efficacia e l'efficienza dei processi per il raggiungimento degli obiettivi. Nonostante l'efficienza del modello weberiano, esso non è applicabile universalmente e dato che, in alcune circostanze, può rappresentare un ostacolo, alcune organizzazioni hanno via via sviluppato delle strutture ibride in cui i benefici derivanti dalla standardizzazione, tipici della burocrazia, vengono affiancati ad approcci più flessibili e adattabili ad imminenti esigenze.

Il contesto attuale, caratterizzato da un sempre più pervasivo ruolo dei computer, dei dati e dell'IA, spinge le organizzazioni burocratiche, seppur con una lenta capacità di adattamento dovuta a rigidità intrinseche proprie, a riconfigurare i propri *modi operandi* con l'obiettivo di semplificare i processi, ridurre costi, migliorare l'efficienza complessiva e ottenere risultati tangibili. A tal proposito, in letteratura, l'influenza del ruolo delle tecnologie IT in relazione alla burocrazia ha permesso di definirne tre tipologie: *di strada*, *di schermo* e *di sistema*.

La *burocrazia di strada* (Lipsky, 1980) fa riferimento a quella tipologia di burocrazia in cui il ruolo delle tecnologie risulta limitato e che è composta da tutti quegli individui che forniscono servizi direttamente al pubblico (letteralmente per strada, es. medici, poliziotti ecc.); tale tipologia di burocrazia è stata via via sostituita dalla *burocrazia di schermo* caratterizzata dall'introduzione del lavoro dietro lo schermo di un computer e da attività routinarie, guidate da strumenti digitali, come software e database, che quindi riducono la discrezionalità umana (Bovens & Zouridis, 2002). Anche le burocrazie a livello di schermo hanno iniziato a cedere il passo a usi ancora più ampi degli strumenti delle tecnologie dell'informazione e della comunicazione (ICT), i quali hanno incrementato la capacità di automatizzare le attività decisionali (Kurzweil, 2005, 2013): si parla, così, di *burocrazia di sistema*, individuabili in quelle burocrazie in cui gli strumenti tecnologici non si limitano soltanto a supportare i burocrati ma, in alcuni casi, sostituiscono il giudizio esperto dei funzionari umani. (Bovens & Zouridis, 2002). In questo quadro generale, il crescente impiego delle cosiddette *macchine intelligenti* è destinato a modificare in maniera profonda e diffusa le strutture occupazionali, gli assetti organizzativi, i processi lavorativi e l'insieme dei mestieri e delle professioni e, con riferimento alle organizzazioni burocratiche, si assisterà alla diffusione di un nuovo paradigma organizzativo, quello dell'*algocrazia* (Aneesh 2006, 2009).

Stato dell'arte dell'IA nelle organizzazioni burocratiche: benefici e limiti.

L'implementazione dell'IA nelle organizzazioni burocratiche, sia nel settore pubblico che in quello privato, è un fenomeno caratterizzato da una lenta diffusione poiché, oltre alla novità della tecnologia e agli onerosi investimenti, richiede anche necessari cambiamenti culturali e organizzativi graduali per superare le resistenze interne e garantire un uso etico e responsabile dell'IA stessa.

Allo stato attuale, da quello che si evince dalla letteratura, l'effettiva implementazione dell'IA nei processi risulta ancora poco diffusa in quanto, nonostante i presunti benefici che le macchine intelligenti possono portare, è emersa la forte difficoltà ad integrare l'AI con persone, processi consolidati e sistemi esistenti. Pertanto, i risultati di alcune ricerche evidenziano che spesso i cospicui investimenti in sistemi di IA non generano i benefici attesi e anzi possono dare luogo, non di rado, ad esiti fallimentari (Makarius, 2020)

Nonostante ciò, nel *settore pubblico* l'IA viene impiegata, da quello che emerge, per migliorare l'efficienza dei servizi erogati ai cittadini: ad esempio, in numerosi paesi sono stati introdotti *chatbot* per facilitare la navigazione all'interno dei portali governativi, fornire informazioni sui servizi pubblici e supportare i cittadini nel completamento delle procedure online.

Nel *settore privato* le grandi aziende industriali, ad esempio, usano l'IA sottoforma di chatbot e assistenti virtuali per il dialogo con la clientela, per automatizzare il backoffice o per la manutenzione predittiva, permettendo di prevenire guasti ai macchinari e riducendo così i tempi di inattività e i costi di manutenzione. Inoltre, le funzioni di gestione delle risorse umane stanno progressivamente adottando strumenti di IA per automatizzare la selezione del personale e valutare le prestazioni dei dipendenti (es. sistema Watson di IBM).

Tuttavia, nonostante i benefici evidenti, l'adozione dell'IA nelle organizzazioni burocratiche, è altresì rallentata da una serie di vincoli normativi e legali: nel settore pubblico, ad esempio, le organizzazioni operano in contesti caratterizzati da requisiti di trasparenza, responsabilità e conformità legali e la presenza di normative stringenti sulla protezione dei dati personali, unite alle questioni etiche sollevate dall'introduzione dell'IA, limita l'implementazione di queste tecnologie; tali ostacoli sollevano interrogativi su come integrare l'IA in modo ottimale e responsabile all'interno dei processi.

Un problema cruciale legato all'uso di algoritmi IA è rappresentato dal fatto che essi operano come *black box*, risultando complesso comprendere le logiche dietro l'apprendimento e l'elaborazione dei dati. Ciò può comprometterne la fiducia nei risultati e quindi l'integrazione nelle pratiche lavorative, specialmente se determinate scelte si rendono automatizzate, in ambiti burocratici, dove la giustificazione e la legittimità delle decisioni sono requisiti essenziali.

In aggiunta, l'apprendimento basato su dati storici potrebbe riflettere, ad esempio, pregiudizi sociali o escludere informazioni rilevanti provenienti da gruppi minoritari, compromettendo così l'equità delle decisioni. In contesti burocratici, dove l'imparzialità è un principio fondamentale, la manifestazione di tali *bias algoritmici* potrebbe, non solo minare la giustizia delle decisioni, ma anche esporre le organizzazioni a responsabilità legali e a danni reputazionali. Risultano inoltre da attenzionare le questioni relative alla privacy e quindi alla gestione dei dati personali, finanziari e comportamentali con cui l'IA viene addestrata.

Per ovviare a queste criticità che, come precisato, limitano l'adozione di questa tecnologia, sono state definite dall'organo HLAG AI (High-Level Expert Group on Artificial Intelligence, 2019) della Commissione Europea *sette linee guida etiche* che i sistemi di IA dovrebbero soddisfare per essere considerati affidabili quando utilizzati in sistemi decisionali che hanno un impatto sui diritti degli individui (come ricevere un risarcimento equo e non discriminatorio dall'assicurazione o l'approvazione di un mutuo, ecc.) che vengono riportate di seguito:

- *Azione umana e supervisione*: i sistemi di IA devono responsabilizzare gli esseri umani, garantendo loro di prendere decisioni informate. Devono inoltre essere previsti adeguati meccanismi di supervisione tramite approcci human-in-the-loop, human-on-the-loop e human-in-command.
- *Robustezza tecnica e sicurezza*: i sistemi di IA devono essere sicuri, cioè garantire un piano di ripiego nel caso in cui si verifichi una criticità, accurati, affidabili e riproducibili. Questo è l'unico modo per garantire che anche i danni non intenzionali possano essere ridotti al minimo e prevenuti.
- *Privacy e governance dei dati*: devono essere rispettate la privacy e la protezione dei dati personali con meccanismi di governance dei dati che tengano conto della qualità e dell'integrità dei dati e ne garantiscano un accesso legittimato.
- *Trasparenza*: i sistemi di IA e i relativi processi decisionali devono essere trasparenti e spiegabili. Le persone devono essere informate quando interagiscono con un sistema di IA e avere consapevolezza delle sue capacità e limitazioni.
- *Diversità, non discriminazione ed equità*: i sistemi di IA devono prevenire pregiudizi e discriminazioni e garantire accessibilità a tutti, indipendentemente da eventuali disabilità, coinvolgendo tutte le parti interessate lungo l'intero ciclo di vita del sistema.
- *Benessere sociale e ambientale*: i sistemi di IA dovrebbero andare a vantaggio di tutti gli esseri umani, comprese le generazioni future. Occorre quindi garantire che siano sostenibili, rispettosi dell'ambiente e degli altri esseri viventi.
- *Responsabilità*: dovrebbero essere messi in atto meccanismi per garantire la responsabilità e la verificabilità delle decisioni prese dai sistemi di IA, specialmente nelle applicazioni critiche e dei loro risultati.

In aggiunta, le normative sulla protezione dei dati, come il Regolamento Generale sulla Protezione dei Dati (GDPR) in Europa, impongono rigorosi vincoli su come i dati debbano essere raccolti, conservati e utilizzati al fine di proteggere la privacy e prevenire usi impropri dei dati stessi. Numerose iniziative, infatti, mirano a promuovere una visione etica dell'IA e centrata sull'essere umano: un progresso significativo in questa direzione è rappresentato dall'Artificial Intelligence Act, la prima legge quadro al mondo sull'Intelligenza Artificiale, che rende i Paesi dell'Unione Europea pionieri nella regolamentazione dello sviluppo, della fornitura e dell'uso dei sistemi di IA in Europa.

2.2 Processi decisionali e Intelligenza Artificiale

L'interesse per l'IA e per il potenziamento delle sue capacità e possibilità di applicazione è in costante crescita da tempo. Già nel 2016 il Consiglio Economico Nazionale e il Chief Technology Officer degli Stati Uniti scrivevano: "Accelerare le capacità dell'intelligenza artificiale permetterà l'automazione di numerose attività che hanno a lungo richiesto il lavoro umano. Piuttosto che fare affidamento su regole strettamente personalizzate ed esplicitamente create dai programmatori, i moderni programmi di intelligenza artificiale possono imparare qualsiasi dato che incontrano e sviluppare le proprie regole su come interpretare nuove informazioni. Ciò significa che l'intelligenza artificiale può risolvere problemi e imparare con pochissimo input umano. Questo (la combinazione di intelligenza artificiale e robotica) consentirà l'automazione di molte attività ora svolte dai lavoratori umani e potrebbe cambiare la forma del mercato del lavoro e dell'attività umana. (Ufficio esecutivo del Presidente, 2016, p. 8)

Contrariamente a tecnologie precedenti, il forte interesse per l'implementazione dell'IA (confermato anche dal grafico sottostante) dipende dalle caratteristiche uniche che la caratterizzano: l'IA, infatti, permette alle macchine di apprendere e agire in modo autonomo (Balasubramanian et al., 2021) e ad interagire con gli esseri umani nel processo decisionale e nella risoluzione di problemi (Murray et al., 2021). Fornendo in input dati ed istruzioni, l'IA "apprende come apprendere", cioè crea il proprio algoritmo senza una programmazione umana esplicita e migliora il proprio apprendimento attraverso un meccanismo di "learning by doing". Così, compiti tradizionalmente svolti dall'intelligenza umana tendono ad essere affidati ai sistemi di IA poiché essi possono garantire maggiore precisione (ad esempio, come nella diagnostica medica) e perché ormai l'enorme quantità di dati da gestire e le capacità computazionali necessarie richiedono l'automazione in molti ambiti.

Un tema dunque rilevante e centrale in letteratura risulta essere quello che approfondisce l'evoluzione della *discrezionalità decisionale*, tipicamente detenuta dagli esseri umani, rispetto alle decisioni automatizzate che verranno sempre più implementate proprio con l'avvento e la diffusione dell'IA.

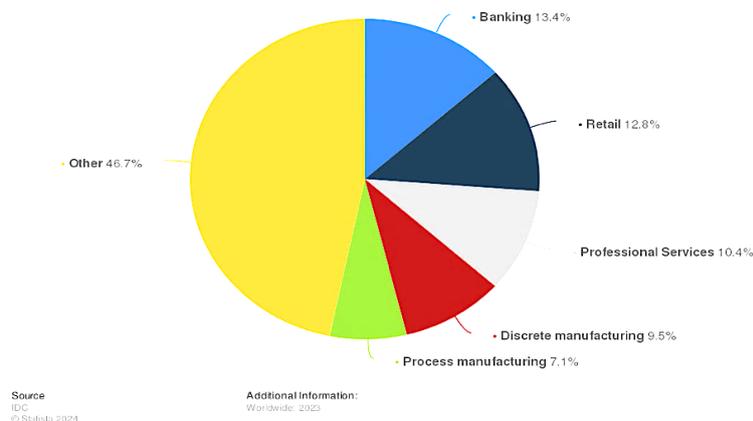


Figura 2.1 Investimenti nell'IA per settore nel 2023

Fattori che influenzano la sostituzione delle decisioni umane con quelle automatizzate.

Un *processo decisionale* razionale consiste in un confronto di alternative e nella successiva selezione dell'alternativa che, in termini di accuratezza ed efficienza, si prevede porti al risultato migliore; tradizionalmente affidato alla *discrezione umana*, ossia alla capacità di prendere decisioni senza regole prestabilite, questa attività è stata via via sempre più supportata dalle tecnologie dell'informazione e della comunicazione (ICT), andando quindi a definire un fenomeno noto come *discrezionalità digitale*. In questo contesto, l'evoluzione dell'IA e delle tecnologie di apprendimento automatico tenderanno ad automatizzare sempre più le attività cognitive. Anche Bush ed Henriksen (2018), in riferimento alle organizzazioni burocratiche menzionate nel capitolo precedente, sostengono che la discrezionalità umana stia lasciando sempre più spazio alla tecnologia e che i burocrati si affidano sempre di più a sistemi digitali e, per cui a tal proposito, si parla di *burocrazia digitale e discrezione digitale*.

Dunque, come si evolverà la discrezionalità con l'introduzione dell'IA nelle attività e nei processi? Essa può ampliare la discrezionalità umana in alcune attività o ridurla in altre? Per rispondere a queste domande è bene fare alcune precisazioni. Perrow (1967) classifica le attività lavorative secondo due dimensioni: la frequenza con cui l'attività richiede deviazioni dalle normali procedure standard e il grado in cui l'attività è analizzabile o può essere suddivisa in un processo di ricerca razionale e sistematica; Buffat (2015), Busch ed Henriksen (2018) nell'ambito dell'automazione delle attività lavorative sostengono che il *contesto* del compito ha un'importanza considerevole.

Secondo queste premesse, la possibilità che un compito venga automatizzato o supportato dall'IA aumenta con il livello di analizzabilità: i compiti analizzabili, cioè che possono essere esaminati o scomposti per capirne i dettagli, risultano meno complessi e più ripetitivi in quanto caratterizzati da bassa incertezza e pertanto maggiormente suscettibili ad essere automatizzati o assistiti dall'IA; al contrario, compiti poco analizzabili che si discostano frequentemente dalle norme sono caratterizzati da alta incertezza e quindi soggetti a minor predisposizione all'automazione.

In sintesi, come riportato sotto in figura 2.2, i compiti che sono ad alta complessità (più deviazioni dalla norma) e ad alta incertezza (meno analizzabili) rimarranno probabilmente sotto il controllo discrezionale degli esseri umani, mentre le attività che sono meno complesse (con minori deviazioni dalla norma) e con minori incertezze (più analizzabili) saranno, invece, più facilmente affidate alle macchine. (Bullock, 2019)

		Uncertainty	
		Low	High
Complexity	Low	Low Complexity, Low Uncertainty High Analyzability, Few Deviations	Low Complexity, High Uncertainty Low Analyzability, Few Deviations
	High	High Complexity, Low Uncertainty High Analyzability, Many Deviations	High Complexity, High Uncertainty Low Analyzability, Many Deviations
		AI Dominates	Leaning AI
		Leaning Human	Human Dominates

Figura 2.2 Complessità del compito, incertezza del compito e Intelligenza Artificiale

Collaborazione Uomo-IA

Gli esseri umani e l'IA differiscono nel modo in cui elaborano le informazioni per creare competenze di dominio: l'IA può elaborare quantità molto maggiori di informazioni a una velocità e un'accuratezza maggiori, mentre gli esseri umani si affidano a scorciatoie per l'elaborazione delle informazioni, le quali potrebbero causare potenziali errori o distorsioni; tuttavia, risultano versatili in ambienti informativi scarsi o complessi.

Piuttosto che sostituire completamente i lavoratori, i sistemi di IA mirano ad integrare e supportare il lavoro umano: infatti, un tema particolarmente approfondito in letteratura riguarda la *collaborazione uomo-IA* (Bankins et al., 2024).

Una possibile complementarità tra gli individui ed IA può avvenire, ad esempio, quando una generica attività può essere suddivisa in sotto-attività che in parte vengono svolte da macchine basate su IA e in parte rimangono sotto la supervisione degli esseri umani: ad esempio, negli ultimi anni e nei paesi più tecnologicamente avanzati, i medici si concentrano esclusivamente sul trattamento del paziente, mentre la diagnosi medica è affidata a macchine con alla base l'IA (Talby, 2019); tale complementarità può anche verificarsi negli stessi compiti: ad esempio, i medici possono integrare la diagnosi medica delle macchine utilizzando la loro comprensione contestuale per individuare eventuali pregiudizi nei risultati delle macchine (Talby, 2019). Entrambe le visioni sottolineano le capacità distintive degli esseri umani e delle macchine e il loro stretto legame potrebbe rappresentare, in determinate realtà, una nuova fonte di vantaggio competitivo (Davenport e Kirby 2016).

Tuttavia, l'introduzione dell'IA non è sempre accolta positivamente e la sua adozione spesso è frenata in quanto emerge un'avversione algoritmica da parte degli individui che prende il nome di *resistenza all'uso della consulenza algoritmica* (Dietvorst et al., 2015): i lavoratori più esperti tendono a manifestare una maggiore avversione al cambiamento poiché sono convinti di avere capacità superiori all'IA, dovuto alla consapevolezza di avere un'esperienza e convinzione di "dominare" processi consolidati in cui partecipano o sovrintendono, e percepiscono maggiori responsabilità per i risultati dall'IA (Allen & Choudhury, 2022), quindi vedono pochi benefici nel suo utilizzo (Kim et al., 2022). Gli effetti di questa avversione, tuttavia, possono essere marginalmente attenuati consentendo ai lavoratori di incorporare le loro conoscenze insieme all'output dell'IA (Kawaguchi, 2021) o avendo un maggiore input decisionale umano in generale (Haesevoets et al., 2021).

In aggiunta, mentre i lavoratori meno esperti mostrano anch'essi avversione agli algoritmi in quanto hanno difficoltà a comprendere e utilizzare efficacemente l'output dell'IA (Allen & Choudhury, 2022), coloro che hanno un'esperienza di dominio moderata sono solitamente più propensi a utilizzare l'IA e a trarne benefici (Allen & Choudhury, 2022).

Se dunque i lavoratori percepiscono nell'IA *fiducia, facilità d'uso* e compatibilità con le tecnologie e pratiche organizzative esistenti (Chatterjee, Rana, et al., 2021) sono più propensi al suo utilizzo.

Limiti dell'IA nei processi decisionali

Nonostante i numerosi vantaggi dell'IA, è essenziale comprenderne i limiti in quanto la sua applicazione nei processi decisionali deve essere accompagnata da un'adeguata consapevolezza delle criticità e delle potenzialità che essa presenta. Quando l'IA viene coinvolta in tali processi, è comunque necessario ricordare che la responsabilità delle decisioni e dei loro risultati rimane sempre in capo agli esseri umani. Le decisioni prese dagli esseri umani sono spesso influenzate da vincoli di natura economica, culturale, sociale e individuale che talvolta possono introdurre elementi di arbitrarietà; al contrario, le macchine, pur essendo più razionali, possono amplificare i pregiudizi umani soprattutto se i dati utilizzati per addestrare i relativi algoritmi contengono errori o bias.

Secondo D'Acquisto, il passaggio dall'autonomia decisionale umana alla crescente implementazione di decisioni algoritmiche solleva due principali questioni: da un lato, l'autonomia delle macchine potrebbe minacciare le prerogative umane, come la responsabilità decisionale, dall'altro essa offre l'opportunità di ridurre errori associate alle scelte umane, a condizione che si riesca a tradurre i valori umani in forme comprensibili e applicabili dalle macchine. La fase cruciale di questo processo risiede nella raccolta della *raccolta dei dati*, dove l'intervento umano è determinante poiché la scelta dei dati è un processo soggettivo che può influenzare e compromettere l'oggettività della decisione finale.

Qualora i dati raccolti non fossero rappresentativi della realtà in cui l'IA dovrà operare, o non includessero campioni minoritari, o ancora, riflettessero dati discriminatori e pregiudizi oppure si verificasse un fenomeno di *data drift*, ossia i dati di input fossero troppo diversi da quelli di training, l'IA riprodurrebbe tali distorsioni.

Quanto menzionato accadrebbe perché le macchine non sono in grado di valutare criticamente il contenuto etico o morale dei dati che elaborano e dunque si potrebbe configurare il rischio che pregiudizi e stereotipi preesistenti vengano consolidati e perpetuati, rendendo difficile la loro eliminazione; è compito degli sviluppatori eseguire controlli e revisioni appropriati per prevenire eventuali criticità.

Un ulteriore rischio deriva dall'*adozione passiva delle decisioni algoritmiche*: se le decisioni prese dalle macchine vengono accettate senza un esame critico, possono produrre effetti imprevisti, in quanto non sempre allineate ai valori decisionali umani.

Le organizzazioni, considerate come "reti di decisioni" (Shrestha et al., 2019), devono essere strutturate per raggiungere efficacemente gli obiettivi prefissati. Le scelte relative alla struttura decisionale come la delega di decisioni a esperti o l'aggregazione delle decisioni di un gruppo di individui, sono da tempo ben comprese e riconosciute come determinanti per le prestazioni organizzative, l'introduzione del processo decisionale basato su algoritmi di IA pone nuove e complesse sfide che richiedono un attento bilanciamento tra automazione e supervisione umana.

2.3 Introduzione dell'Intelligenza Artificiale nel settore assicurativo

La diffusione dell'IA sta via via influenzando in modo considerevole il settore assicurativo tradizionale, favorendo un processo di digitalizzazione sempre più rilevante e capace di trasformare un ambito che, per la sua natura fortemente regolata e la consolidata redditività degli operatori storici, ha registrato pochi cambiamenti significativi nel tempo.

Oggi, l'introduzione dell'IA rappresenta una forza propulsiva che potrebbe imprimere una svolta ad un settore tradizionalmente stabile, permettendo di colmare il divario in termini di innovazione rispetto a settori tecnologicamente più avanzati come il finanziario, che ha già da tempo avviato la propria trasformazione digitale smaterializzando gran parte dei processi operativi.

L'ingresso nel mercato di giganti tecnologici come Amazon e Google, con la loro vasta disponibilità di dati, l'intensificarsi della concorrenza con startup InsurTech (contrazione di "Insurance Technology"), l'elevato costo del lavoro e le mutevoli aspettative dei clienti sono solo alcuni dei problemi che possono essere affrontati sfruttando le potenzialità dell'IA: infatti, per comprendere come gli assicuratori possono sfruttare le potenzialità dell'IA in questo settore, è necessario, per prima cosa, analizzare le *aspettative dei consumatori moderni*. Nel contesto di una forte digitalizzazione, in tutti i settori, gli utenti sentono la necessità di ottenere risposte velocemente e di fruire di servizi personalizzati; le esperienze personalizzate costituiscono un nuovo status quo e le aspettative dei clienti nei confronti dei servizi digitali sono in costante crescita. Diventa, quindi, fondamentale integrare la tecnologia digitale anche con i prodotti assicurativi tradizionali, ma un'integrazione senza l'IA risulterebbe inefficace e porterebbe le compagnie che non riescono a soddisfare tali aspettative a perdere competitività (McKinsey Global Insurance Report 2023).

Partendo dalle considerazioni esposte e analizzando la catena del valore del settore assicurativo (figura 2.3) secondo il modello di Porter, è possibile individuare le aree maggiormente interessate dall'implementazione dell'IA (Sushant K, 2020).

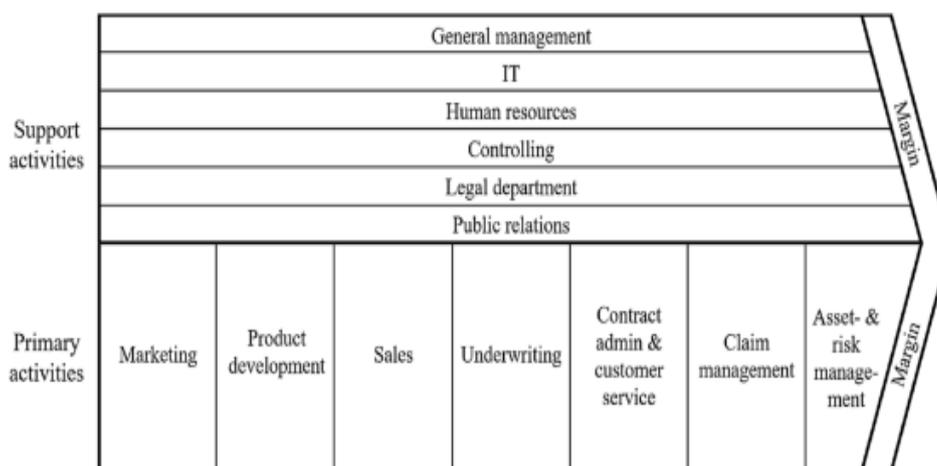


Figura 2.3 Value Chain del settore assicurativo

Nelle aree *sales, marketing e sviluppo prodotto* l'IA può offrire sia suggerimenti personalizzati sui prodotti assicurativi, adatti alle capacità finanziarie e ai profili di rischio del cliente, sia supporto agli assicuratori nell'individuare nuovi clienti e sviluppare campagne di marketing, poiché attraverso l'analisi delle tendenze di mercato e delle preferenze dei consumatori, l'IA permette infatti di ottimizzare la definizione dei prezzi. Inoltre, grazie alla capacità di analizzare i dati dei clienti, l'IA è in grado di prevedere la probabilità di abbandono del servizio, permettendo così alle compagnie assicurative di intervenire tempestivamente con offerte mirate o miglioramenti nel servizio e mantenere stabile la base di clienti.

Per quanto riguarda l'*amministrazione delle polizze*, grazie all'IA è possibile velocizzare la gestione delle richieste di risarcimento, tradizionalmente caratterizzate da procedure lunghe e complesse, riducendo grazie all'automazione dei processi quali l'estrazione, la strutturazione, l'archiviazione e il trasferimento dei dati, diminuendo la necessità di intervento manuale e, così, anche il rischio di errore.

Inoltre, automatizzando le risposte alle richieste e offrendo assistenza personalizzata tramite chatbot avanzati e assistenti virtuali, l'IA può incrementare l'efficienza del *servizio clienti*, con una notevole riduzione dei tempi di risposta e un conseguente aumento della soddisfazione e fidelizzazione del cliente.

Nell'ambito della *gestione dei sinistri*, l'IA è in grado di individuare in modo efficace potenziali frodi attraverso l'identificazione di anomalie e di elementi sospetti nelle immagini, contribuendo così alla mitigazione dei rischi e alla riduzione delle perdite. Ne consegue, quindi, anche una stima più accurata degli importi da risarcire, confrontando i dettagli dei sinistri con le coperture previste nelle polizze.

Infine, l'IA migliora la *gestione e la mitigazione dei rischi* associati ai modelli predittivi e analitici, strumenti essenziali per la previsione e gestione dei vari tipi di rischio, come quelli di credito, di mercato o operativi.

Tuttavia, alcuni player intersettoriali come Google, così come start-up innovative InsurTech, quali Lemonade, Oscar e Tractable, hanno già iniziato a invadere questo settore per impadronirsi di valore. Tractable, ad esempio, è un caso di successo in questo trend.

La startup, che è stata fondata nel 2014 a Londra ed è operante in vari paesi (USA, Canada, Giappone, UE), ha una valutazione di mercato di quasi 1 miliardo di dollari e offre i suoi prodotti e servizi a compagnie assicurative, venditori di auto, gestori di flotte, meccanici e OEM (Original Equipment Manufacturer). La startup ha sviluppato un'applicazione web che, sfruttando l'IA in cloud analizzando foto e video, fornisce una stima del danno ottenendo una riduzione dei tempi e dei costi delle parti coinvolte nella gestione dei sinistri.

Nonostante i sistemi IA presentino evidenti potenzialità, le indiscutibili limitazioni della novità della tecnologia, i pregiudizi nei risultati uniti ai cambiamenti organizzativi necessari a causa dell'assenza di un'adeguata cultura specifica, oltre all'alto livello di incertezza e ai costi elevati fanno sì che i leader di mercato siano ancora cauti con gli investimenti in questa direzione. Ess preferiscono infatti collaborare con le startup, probabilmente per osservare e imparare. Ad esempio, nel 2022, Aviva, compagnia assicurativa britannica, ha collaborato con Tractable per

automatizzare la valutazione e la stima di decine di migliaia di sinistri auto attraverso l'utilizzo dell'IA.

Con l'integrazione dell'IA negli aspetti precedentemente illustrati, tutto avverrebbe più velocemente, meglio e con meno errori; infatti, la ricerca esistente ha dimostrato come la classificazione automatizzata delle macchine possa essere più accurata del 30% rispetto alla classificazione manuale da parte dell'uomo e abbia un potenziale tale da aumentare la produttività dell'80%.

Nel settore assicurativo, l'adozione dell'IA richiede la valorizzazione della risorsa più importante a disposizione delle compagnie assicurative, i *dati*, in relazione ai quali è necessario implementare strategie nuove in funzione delle esigenze legate all'IA stessa e garantire solide misure di protezione dei dati stessi in conformità alle normative sulla privacy, anche se ciò aggiunge ulteriori complessità nell'adozione dell'IA, poiché violazioni o usi impropri dei dati personali possono comportare gravi conseguenze, tra cui sanzioni pecuniarie e danni reputazionali.

Capitolo 3

Data Quality Management

3.1 Definizioni e caratteristiche dei dati

All'interno delle organizzazioni, raccogliere, gestire e utilizzare grandi quantità di dati è essenziale per raggiungere il successo: poiché la capacità di trarre valore dai dati consente di superare i concorrenti e conservare un vantaggio competitivo nel mercato. I dati, infatti costituendo una componente fondamentale del patrimonio aziendale devono essere trattati con estrema cura e precisione nel rispetto degli obiettivi e degli scopi aziendali.

Oggi, i dati giocano un ruolo cruciale nello sviluppo e nell'implementazione di modelli di Intelligenza Artificiale all'interno dei processi aziendali, per cui disporre di grandi quantità di dati, purché essi siano di alta qualità, consente ai modelli di IA di riconoscere un'ampia gamma di situazioni, migliorando l'accuratezza e l'efficacia nelle previsioni.

Definizioni generali

Il Data Quality Management (DQM) è un insieme strutturato di processi finalizzati a garantire che i dati utilizzati siano *fitness for use*, ovvero che soddisfino i requisiti necessari per il raggiungimento di uno specifico obiettivo. È importante sottolineare che il concetto di qualità dei dati non è univoco ma varia in base alle specifiche esigenze e finalità in quanto un dato può essere considerato di qualità adeguata a uno scopo, ma non per un altro.

L'obiettivo primario della misurazione e della gestione della qualità dei dati non è semplicemente quello di garantire l'alta qualità dei dati stessi, ma piuttosto quello di aiutare un'organizzazione, come nel caso che verrà presentato nel capitolo successivo un ente assicuratore, nel raggiungimento dei propri obiettivi di business.

A tal fine, per rispondere alla domanda "come devono essere utilizzati i dati per raggiungere gli obiettivi di business?", è necessario comprendere cosa siano i dati. Di seguito vengono riportate due definizioni comunemente accettate nell'ambito dell'Information Technology (IT).

- Merriam-Webster definiscono i dati come:
 - 1) informazioni fattuali (come misurazioni o statistiche) utilizzate come base per il ragionamento, la discussione o il calcolo;
 - 2) informazioni in formato digitale che possono essere trasmesse o elaborate;

- 3) informazioni emesse da un dispositivo o organo di rilevamento che includono sia informazioni utili che informazioni irrilevanti o ridondanti e devono essere elaborate per essere significative.
- Il Cambridge Dictionary offre definizioni simili, descrivendo i dati come:
 - 4) informazioni, in particolare fatti o numeri, raccolte per essere esaminati, considerati e utilizzati per supportare il processo decisionale;
 - 5) informazioni in formato elettronico che possono essere memorizzate e utilizzate da un computer.

Assunte queste definizioni, è fondamentale chiarire la differenza tra *dato* e *informazione*.

Il *dato* è una rappresentazione originaria, cioè non interpretata di un fenomeno, evento o fatto, priva di alcun significato intrinseco: esso descrive in modo semplice un oggetto.

L'*informazione* è il risultato dell'interpretazione e contestualizzazione di un dato, o di un insieme di dati. L'informazione è ciò che si ottiene quando i dati vengono inseriti in un contesto che li rende comprensibili

Quando le informazioni vengono memorizzate o utilizzate per risolvere problemi o prendere decisioni si trasformano in *conoscenza*, che pertanto, rappresenta l'uso e l'applicazione delle informazioni. La conoscenza, sia umana sia quella appresa da un algoritmo, supporta il processo decisionale è adattiva e dinamica, cioè si evolve costantemente in base alle nuove esperienze e ai nuovi dati.

Quindi, con riferimento al concetto di qualità menzionato prima, i dati di qualità possono essere considerati come portatori di informazione, in relazione al contesto in cui vengono utilizzati.

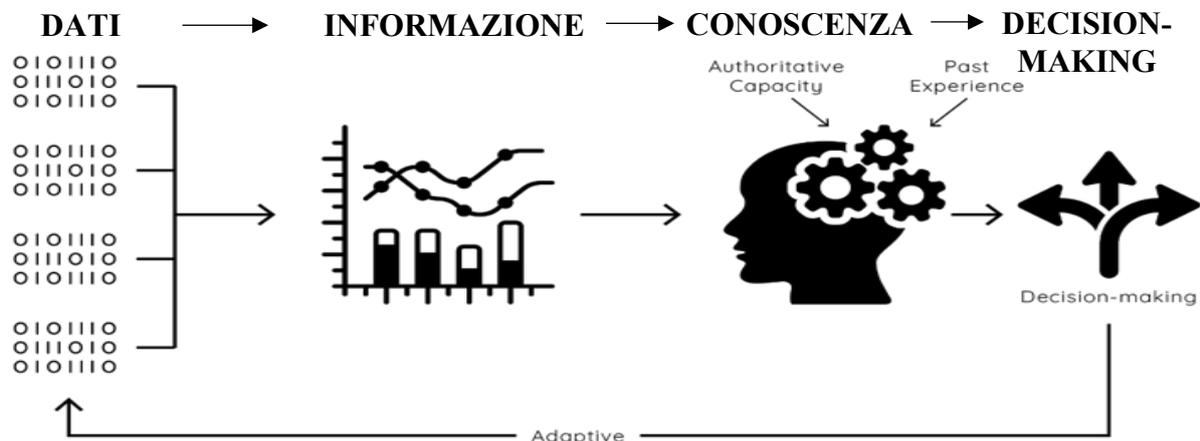


Figura 3.1 Dai dati grezzi alle decisioni

Tipologie e caratteristiche di qualità dei dati

Generalmente i dati si presentano in modo aggregato, andando a determinare un *dataset* o *set di dati*. Ogni dataset è caratterizzato da un certo *volume*, cioè da numero di dati presenti al suo interno che possono presentarsi nelle seguenti strutture (Cichy & Rass, 2019):

- *Dati strutturati*: hanno una struttura ben definita, regolare, ben organizzata e facilmente accessibile, sono solitamente utilizzati da un'entità o da un programma informatico.

I dati strutturati sono spesso memorizzati in modo tabellare, in schemi ben definiti come i database relazionali. I dati strutturati includono nomi e date, indirizzi, numeri di carte di credito, informazioni sulle azioni, geolocalizzazione e così via.

- *Dati semistrutturati*: a differenza dei dati strutturati, non sono conservati in un database relazionale ma contengono qualità organizzative che ne facilitano l'analisi. I dati semistrutturati includono documenti HTML, XML, JSON, database NoSQL e così via.
- *Dati non-strutturati*: non hanno un formato o un'organizzazione predefinita, il che li rende molto più difficili da acquisire, gestire e analizzare poiché sono costituiti principalmente da testo e informazioni multimediali. I dati non strutturati includono dati da e-mail, post di blog, documenti di elaborazione testi, file PDF, file audio, video, foto, presentazioni, pagine Web e molti altri documenti aziendali.
- *Metadati*: non si tratta di dati ordinari, ma di "dati sui dati". La principale distinzione tra "dati" e "metadati" è che i dati sono semplicemente elementi che possono essere utilizzati per classificare, misurare o persino documentare qualcosa in relazione agli attributi dei dati di un'organizzazione. I metadati, d'altra parte, spiegano le informazioni sui dati pertinenti, rendendole più significative per gli utenti dei dati. L'autore, la dimensione del file, la data generata dal documento, le parole chiave che definiscono il documento e così via sono tutti esempi di metadati.

Inoltre, affinché il dataset sia di valore per il contesto in cui andrà ad inserirsi, è indispensabile che esso presenti le seguenti *caratteristiche di qualità* (Batini 2006):

- *Completezza*: la misura in cui un dataset è sufficientemente ampio per il compito da svolgere, contiene tutti i parametri per affrontare i casi d'uso previsti, incluse le situazioni limite (edge cases). Ad esempio, nel caso di un dataset per la realizzazione di un modello di riconoscimento facciale, la completezza richiede non solo diversità nei volti, ma anche la presenza di immagini sfocate, volti parzialmente coperti o in controluce: questo assicura che il modello possa funzionare anche in condizioni meno comuni o difficili.
- *Coerenza*: la misura in cui i dati all'interno di un dataset sono uniformi, non presentano conflitti o discrepanze tra di loro compatibili tra di loro e sono presentati nello stesso formato.
- *Accuratezza*: la misura in cui i dati sono corretti, affidabili e certificati (vicini a quella che è la realtà).
- *Tempestività*: la misura in cui l'età dei dati è appropriata per l'attività da svolgere; è fondamentale che i dati, dunque, siano aggiornati, recenti e rilevanti al momento dell'uso
- *Diversità*: la misura in cui il dataset rappresenta accuratamente la varietà presente nella popolazione.

3.2 Caratteristiche del processo di DQM in riferimento al settore assicurativo

Approcciandosi ad un contesto aziendale come quello di un ente assicuratore e adottando una prospettiva orientata alla gestione della qualità dei dati (DQM) emerge l'importanza del ruolo di supporto fornito dalle strutture IT. Il processo di DQM è strutturato in una serie di passaggi iterativi, illustrati in figura 3.2, che delineano fasi operative e metodologiche del processo.



Figura 3.2 Data Quality Management nel settore assicurativo

Come si evince, i componenti base di un processo di DQM a livello “macro” (successivamente verrà approfondito un processo di DQM “micro” relativo all’implementazione di un modello di Machine Learning di classificazione) sono *controllo, garanzia e miglioramento della qualità*, i quali sono guidati da un quarto elemento essenziale, ovvero *un piano e una strategia* (Graham H. et al 2020). Tale processo, come precedentemente evidenziato, è iterativo e supportato dall’infrastruttura IT per la gestione dei dati, dalle risorse organizzative e dal contributo delle persone coinvolte.

Innanzitutto, la gestione della qualità dei dati è, fondamentalmente, una sfida strategica: senza il contributo degli stakeholder, che forniscono indicazioni sui requisiti o su un piano che leghi la qualità dei dati alla strategia generale non è possibile misurare né gestire correttamente la qualità dei dati. Per esempio, un assicuratore, prima di avviare un progetto come quello del caso che verrà presentato, dovrebbe porsi le seguenti domande:

- I dati devono essere aggregati per linea di prodotto, segmento di clientela, linea di business, area geografica/domicilio, business unit, tipo di sinistro, ecc.?
- Una volta raccolti i dati, qual è il metodo migliore per la loro archiviazione?
- Con che frequenza si dovrebbero raccogliere dati (transazionale, giornaliero, mensile, ecc.)?
- Quali processi e tecnologie è necessario implementare per consentire la raccolta, la gestione dei dati? E come deve avvenire il processo decisionale basato sui dati senza creare ostacoli legati alle sue attività quotidiane?
- Quali sono i costi delle varie strategie di dati? Tali costi sono giustificati rispetto al loro beneficio percepito e all’impatto sugli obiettivi dell’assicuratore?

Questa fase di pianificazione strategica, nonostante i notevoli costi, sia in termini di tempo che economici, deve essere collaborativa e coinvolgere tutte le funzioni aziendali, così da stabilire in maniera trasparente i requisiti che i dati devono soddisfare.

La fase del *controllo di qualità* si concentra principalmente sulla qualità dei *dati forniti* all'utente, dopo che i dati stessi sono stati acquisiti dall'azienda, essa è anche affiancata da un "monitoraggio della qualità dei dati" cioè da un processo di identificazione di eventuali criticità nel momento in cui i dati forniti all'utente finale non soddisfano i requisiti richiesti dall'utente stesso; a tal proposito è bene effettuare una revisione di regole e procedure di controllo della qualità per aiutare a prevenire il ripetersi di anomalie dei dati.

Mentre il controllo di qualità si concentra sul processo di garanzia che i dati forniti all'utente siano di qualità appropriata e sull'individuazione delle eccezioni, la *garanzia di qualità* si focalizza sul corretto funzionamento dei processi di *acquisizione* ed *elaborazione dei dati* e verifica se tali processi possono essere migliorati per aumentare la qualità complessiva dei dati. Parallelamente al costante monitoraggio della qualità dei dati, molti esperti di dati sostengono che è anche indispensabile la *misurazione* della qualità dei dati, attraverso l'utilizzo di misure soggettive, per capire in che direzione migliorare la qualità dei dati. A tal proposito ci sono diverse dimensioni di qualità che possono essere misurate e/o valutate, ma in generale, si dovrebbero valutare dimensioni specifiche relative ai valori dei dati (ad esempio, l'accuratezza), i processi che hanno prodotto i dati (ad esempio, la frequenza delle anomalie dei dati, problemi con la conversione dei dati da un formato all'altro, ecc.) e le dimensioni che influenzano l'adeguatezza e la completezza dei dati per l'uso previsto.

Infine, l'ultima componente è il *miglioramento della qualità* che comprende sia l'analisi del perché e del come viene prodotta una scarsa qualità dei dati, sia l'adozione di interventi migliorativi ai processi che hanno prodotto dati di scarsa qualità. Come notato, se un'azienda, nel caso che verrà presentato, un ente assicuratore, non adotta efficaci interventi per la gestione dei dati, ad esempio in ottica di implementazione dell'IA, è probabile che il processo esistente possa produrre, non solo dati di scarsa qualità, ma soprattutto processi inefficienti e quindi anch'essi di scarsa qualità.

Ogni iterazione del framework di Data Quality Management "ad alto livello" appena descritto, in cui la pianificazione e la strategia contribuiscono allo sviluppo di un solido processo di controllo, di garanzia e di miglioramento della qualità dei dati, è supportata dal contributo degli stakeholder, dalle risorse organizzative e dalle tecnologie disponibili.

Infatti, la sinergia di tutti gli elementi presentati non solo permette di rispondere nel miglior modo possibile ad esigenze e novità provenienti dall'ambiente esterno (nuove tecnologie, nuovi customer needs ecc) ma, soprattutto, di gestire e ottenere dati di qualità superiore, quest'ultimi fondamentali, ad esempio, per l'addestramento efficace di algoritmi IA che via via saranno sempre più integrati all'interno dei processi aziendali. Se quindi cambiano le necessità occorre rivedere il DQM e apportare cambiamenti mirati.

3.3 DQM per il Machine Learning: importanza del Data Cleaning

Una gestione ottimale dei dati a livello “macro” in un contesto aziendale può essere considerato di grande valore all’interno dell’azienda, specialmente nello scenario attuale, per uno sviluppo efficace di algoritmi di ML da integrare all’interno dei processi.

L’obiettivo della tesi è quello di colmare un gap nella letteratura, capendo in che modo i dati possano essere analizzati, selezionati, trattati e combinati tra loro per l’addestramento di un modello di Machine Learning dedicato alla classificazione di immagini (tabella 3.1), ossia un modello di Computer Vision-

Tabella 3.1 Tipologie di Machine Learning, costruzione del modello sulla base dei dati ed esempi (Vijayaraj et al., 2023)

Tipologia di Machine Learning	Costruzione del modello	Esempi
<i>Supervisionato</i>	Gli algoritmi o i modelli apprendono da dati etichettati (Task-Driven Approach)	Classificazione, Regressione.
<i>Non-supervisionato</i>	Gli algoritmi o i modelli apprendono da dati non etichettati (Data-Driven Approach)	Raggruppamenti, Associazione
<i>Semi-supervisionato</i>	I modelli sono costruiti combinando dati (etichettati + non-etichettati)	Classificazione, Raggruppamento
<i>Per rinforzo</i>	I modelli si basano su ricompense o penalità (Environment-Driven Approach)	Classificazione, Controllo

Gestione dei dati per un modello di Machine Learning di classificazione

Per lo sviluppo di un modello di ML, in particolare di Computer vision, come quello del caso studio che verrà presentato nel capitolo successivo, è necessario seguire un processo di lavoro strutturato: occorre, innanzitutto, definire il problema che si intende risolvere con il modello, dopo si analizzano i modelli esistenti e le tecnologie disponibili compatibili per la risoluzione del problema, si sceglie il modello e si definiscono delle metriche per valutarne le prestazioni. Soltanto dopo questa prima “macrofase” è possibile lavorare sui dati.

In letteratura è proposto un framework per la *gestione dei dati* articolato in più step: dopo la raccolta dei dati (*Data collection*), si passa alla preparazione dei dati (*Data preparation*) in cui il dataset di immagini viene processato per garantirne l’idoneità all’uso nel modello di ML; tale fase include le seguenti sottofasi:

- *Data analysis*: viene condotta un’analisi esplorativa circa la comprensione della natura del dataset, in particolare vengono svolte analisi statistiche delle immagini e vengono ricercate anomalie o criticità nei dati.
- *Data generation*: se il dataset è insufficiente, talvolta, è possibile ricorrere ad un’operazione di *data augmentation* per generare nuovi dati sintetici. Tecniche come rotazione, zoom, traslazione, e flipping delle immagini vengono applicate per aumentare la varietà del dataset senza alterarne la natura.
- *Data annotation*: è lo step fondamentale nei modelli di ML nell’ambito della Computer vision in quanto permette, ad esempio attraverso l’abbinamento di etichette a *maschere semantiche*, di individuare in modo specifico il soggetto dell’immagine e quindi di

apprendere in maniera più precisa e rapida. È un task che ha acquisito e acquisirà un grande valore in quanto avere dati annotati in maniera corretta rappresenta un grande vantaggio competitivo per le aziende.

- *Data organization*: i dati vengono organizzati in modo coerente. Tipicamente, si suddividono le immagini in set di training, di validazione e di testing.

Quest'ultima fase ha una grande valenza per l'addestramento del modello di ML per i motivi che verranno esposti di seguito.

I *dati di training* sono una raccolta di informazioni etichettate che permettono al modello di imparare e rappresentando la parte porzione più grande del set di dati, generalmente costituiscono almeno il 70-80%¹ dei dati totali che verranno utilizzati per la costruzione di uno specifico modello. Senza i dati di training, l'apprendimento automatico non esisterebbe. Tali dati vengono utilizzati in più cicli di addestramento per consentire al modello di imparare e migliorare nello svolgimento della propria attività con un alto livello di precisione.

In altre parole, i dati di training sono “il libro di testo” che insegneranno al modello a svolgere il compito assegnato poiché più un algoritmo “studia”, meglio risponderà al test finale che è l'applicazione nel mondo reale.

Il modo in cui si crea, formatta e si annota il set di dati di training ha un impatto diretto sul modello creato e influenza l'accuratezza; infatti, i dati mal elaborati sono uno dei motivi più comuni per cui i progetti di apprendimento automatico falliscono.

I *dati di validazione* vengono utilizzati principalmente per determinare se il modello è in grado di identificare correttamente i nuovi dati o se è in overfitting (concetto che verrà spiegato nel dettaglio nel capitolo successivo) rispetto al set di dati originale. La fase di convalida consente ai data scientist di regolare gli *iperparametri*, cioè tutti quei parametri specifici di un modello, e quindi migliorare l'accuratezza del modello stesso.

I *dati di test* vengono utilizzati sia dopo il training che dopo la convalida: essi permettono di testare l'accuratezza del modello finale rispetto agli obiettivi, ma anche di fornire un'ulteriore conferma che il modello non è in overfitting rispetto ai dati di training e di validation.

Per evitare pregiudizi di selezione, è meglio mescolare i dati in queste tre categorie in modo casuale.

Infine, l'ultima “macrofase” prevede l'addestramento e la validazione del modello ed è suddivisa in tre sottofasi:

- *Model training*: in questa fase, il modello viene addestrato utilizzando il dataset di training e “impara” dai dati.
- *Model validation*: il modello viene validato con un set separato di dati di validazione per verificare che stia generalizzando correttamente e non si stia sovra-adattando (overfitting) ai dati di training. Se le prestazioni del modello non sono soddisfacenti, si può tornare alla fase di training per migliorare ulteriormente.
- *Model testing*: se il modello supera la validazione con prestazioni simili a quelle che si vuole raggiungere, viene testato su un set di dati separato che non ha mai “visto” prima così da

¹ Fonte The Essential Guide to AI Training Data-TELUS International The Essential Guide to AI Training Data

Importanza del data cleaning

Affinché l'apprendimento dei modelli di ML risulti funzionale agli obiettivi prefissati, la disponibilità di un ampio volume di dati, come visto nel framework in figura 3.3, va gestita.

In aggiunta a ciò, è fondamentale *gestire la qualità* complessiva dei dati poiché essi potrebbero generare *incertezza* e compromettere l'accuratezza dei risultati di un modello di ML.

L'*incertezza*, in particolare, che si manifesta quando i dati sono incompleti, imperfetti e, in generale, "rumorosi" può avere diverse origini, tra cui errori nella raccolta dei dati, variabilità intrinseca dei dati o la multimodalità, ovvero l'elevata complessità dei dati stessi. Ad esempio, se i dati di training risultano distorti in qualche modo, come privi di coerenza o non strettamente correlati con gli obiettivi del modello, è probabile che l'algoritmo produca risultati imprecisi e poco affidabili.

A tal proposito (anche in relazione al caso studio che verrà presentato nel capitolo successivo), è possibile collocare nel framework presentato al capitolo 3.2 (tra le fasi di data analysis e annotation) una fase che permette di gestire e mitigare l'incertezza dei dati, cioè quella del *data cleaning* che, mirando ad aumentare la qualità dei dati correggendo errori ed omissioni, risulta necessaria quando "dati sporchi" (*garbage data*) porterebbero inevitabilmente a risultati scadenti (*garbage results*) con conseguente compromissione dell'affidabilità dei modelli (Stuart et al 2019).

Nonostante in letteratura non siano espressamente presenti dei framework di data cleaning in relazione ad un modello di ML di Computer vision, le fasi principali di tale processo possono essere sintetizzate nelle seguenti:

- *Rimozione di dati duplicati*: la duplicazione dei dati può essere causata da diversi fattori, come ad esempio una raccolta dei dati da diverse fonti. Dati duplicati non aggiungono alcun valore al modello e possono portarlo, in determinati casi, a trarre conclusioni errate
- *Rimozione di dati irrilevanti o errati*: i dati irrilevanti sono dati che non hanno alcuna utilità per l'addestramento del modello, creano "rumore" e riducono l'efficienza del modello. Probabilmente derivano probabilmente da errori nella raccolta dei dati stessi. Nonostante filtrare questa tipologia di dati riduca il volume di dati per il training, il modello comunque è in grado di apprendere da informazioni significative e migliorare le sue performance.
- *Gestione di dati mancanti*: i dati mancanti devono essere gestiti, attraverso azioni mirate, per far sì che il dataset sia più completo e affidabile

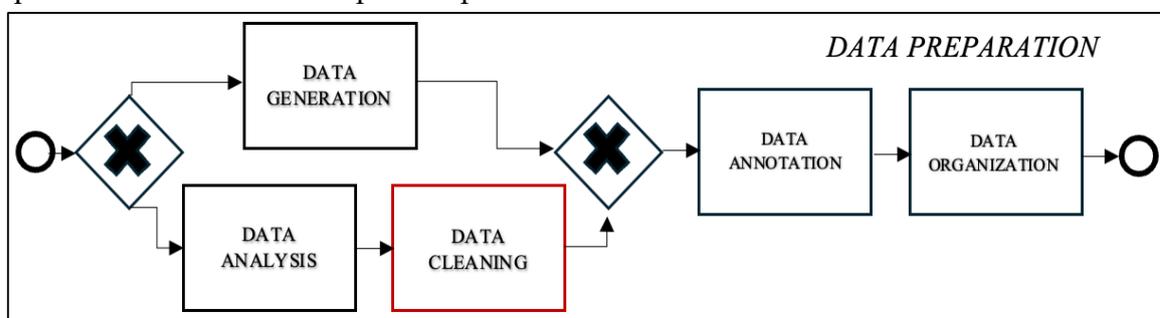


Figura 3.4 Data preparation con data cleaning

3.4 Tempo e fattori chiave nella preparazione dei dati per il Machine Learning

Una questione rilevante da considerare riguarda sia la *quantità di dati* necessari per garantire elevate prestazioni di un algoritmo sia il *tempo necessario* per adattarli in modo idoneo al training del modello.

Questi aspetti sono strettamente correlati e soggetti all'influenza di una serie di fattori, spesso sconosciuti nelle fasi iniziali o anche durante lo sviluppo, per cui risulta necessario adottare un approccio iterativo.

Nella letteratura attuale si riscontra anche una mancanza di esempi concreti di applicazioni di ML che colleghino chiaramente la quantità di dati necessari al tempo richiesto per adattarli secondo certi criteri al fine di raggiungere specifici obiettivi di performance. Pertanto, la presente tesi si propone di colmare proprio tale lacuna attraverso l'analisi di un caso studio concreto nel settore assicurativo.

Secondo il report di Telus (The Essential Guide to AI Training Data-TELUS International, n.d.), società canadese specializzata in trasformazione digitale, l'implementazione efficace di un modello di ML, così come la determinazione delle quantità di dati e delle tempistiche di adattamento, è condizionata da diversi fattori chiave che includono:

- La *complessità del modello*: dipende dal numero di parametri che il modello deve tenere in conto poiché più parametri richiedono maggiori dati di training, per cui all'aumentare del numero di parametri, cresce anche la dimensione del set di dati per addestrare il modello in modo efficace.
- Il *metodo di allenamento* influisce direttamente sulle tempistiche. Ad esempio, i modelli di ML che utilizzano l'apprendimento supervisionato hanno un limite relativamente basso per quanto riguarda i benefici derivanti dall'aggiunta di nuovi dati. Al contrario, i modelli basati sull'apprendimento non supervisionato possono migliorare autonomamente, identificando e ottimizzando i propri parametri, anche in assenza di una struttura predeterminata; tuttavia, questo approccio richiede una maggiore quantità di dati ed estende la curva di apprendimento in quanto ulteriori dati contribuiscono a migliorare le prestazioni del modello.
- Le *esigenze di etichettatura* influiscono sull'efficacia dei modelli di ML (soprattutto per quelli supervisionati) e sulla capacità di apprendere correttamente.
- La *tolleranza per gli errori* ha un grande impatto sulla quantità di dati che è necessario raccogliere e processare. Alcuni modelli devono avere un livello di prestazioni elevate e un basso tasso di errore. Si pensi alla differenza tra un modello che prevede il meteo ed uno che identifica i pazienti che sono a rischio imminente di infarto. È evidente che uno di questi ha una soglia di errore tollerabile molto più bassa rispetto all'altro, per cui più basso è il livello di rischio accettabile, più dati saranno necessari per garantire che il rischio sia mitigato e tendente a zero.
- Nel mondo reale i modelli possono imbattersi in una ampia *diversità dei contributi* e in una vasta gamma di dati di input. Una chatbot, ad esempio, deve essere in grado di comprendere

lingue diverse, espressioni formali e informali o addirittura testi grammaticalmente scorretti, deve essere in grado di comprenderli tutti per fornire un servizio clienti di alto livello. Nei casi in cui l'input del modello non sia altamente controllato, occorrerà una quantità maggiore di dati per aiutare il modello a funzionare efficacemente anche in situazioni imprevedibili.

Oltre ai fattori elencati si devono anche considerare, come precisato nel capitolo 2.2, *quali azioni si vogliono automatizzare o complementare attraverso l'uso dell'IA*: azioni ad alta incertezza e alta frequenza di errori, azioni a bassa incertezza e bassa frequenza di errori, ecc.

Per individuare le tempistiche in base alle esigenze del modello è utile fare riferimento alle *curve di apprendimento* che aiutano a identificare la relazione tra le dimensioni del set di dati, le capacità e le prestazioni del modello. Dopo un certo numero di iterazioni (approccio iterativo) si può raggiungere un punto in cui l'aggiunta di ulteriori dati offre rendimenti decrescenti: i dati, quindi, possono essere aggiunti solo se diventa necessario, poiché, Man mano che il modello inizia a produrre risultati, le esigenze relative all'addestramento potrebbero modificarsi.

Va infine sottolineato, anche come specificato precedentemente, che prima dell'addestramento i dati necessitano di una significativa elaborazione per poter essere utilizzati in modo positivo sul modello.

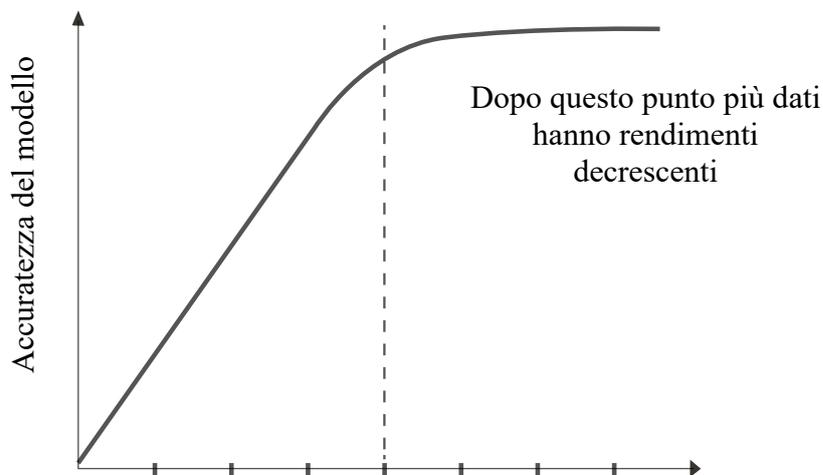


Figura 3.5 Influenza dell'aggiunta di dati sull'accuratezza del modello

3.5 Acquisizione e trasferimento del know how

L'integrazione di modelli di IA all'interno di un'organizzazione costituisce un processo complesso e articolato che richiede un considerevole investimento sia in termini di risorse finanziarie che temporali. Questo processo risulta non soltanto oneroso per lo sforzo richiesto durante la fase di implementazione, ma anche a causa delle limitazioni intrinseche nei modelli di IA per quanto concerne, in particolare, la loro capacità di acquisire e replicare la conoscenza detenuta dagli individui e da un'intera organizzazione.

Sebbene un modello di IA possa dimostrarsi estremamente efficace nell'analisi e nell'elaborazione di dati nuovi in input, esso rimane fortemente legato alle informazioni che gli vengono fornite originariamente e, di conseguenza, alla natura e alla qualità di tali informazioni. Come già sottolineato, le informazioni contribuiscono alla generazione di conoscenza, che, a livello umano può essere classificata in due categorie principali: la conoscenza esplicita e quella implicita, spesso indicata anche come conoscenza tacita.

L'efficace individuazione dei metodi di trasferimento di queste due tipologie di conoscenza è un aspetto cruciale per comprendere le sfide connesse all'addestramento e all'efficace implementazione dei modelli di IA.

Conoscenza esplicita e tacita

La *conoscenza esplicita* si riferisce a quella forma di sapere che può essere facilmente trasmessa, spiegata e codificata. Essa comprende informazioni che possono essere comunicate senza perdita di contenuto, permettendo al destinatario di raggiungere lo stesso livello di comprensione del mittente (Ambrosini e Bowman, 2001). Questo tipo di conoscenza è caratterizzato dalla propria "oggettività" intrinseca in quanto, non dipendendo da specifici contesti o dall'identità di chi la possiede; è quindi intrinsecamente più precisa e affidabile (Gamble, 2020), e non essendo idiosincratICA, può essere applicata con successo in diversi contesti organizzativi senza perdere la propria efficacia (Sobol e Lei, 1994). Tale forma di conoscenza si dimostra facilmente adattabile all'automazione e all'integrazione nei modelli di IA.

Di contro, la *conoscenza tacita* rappresenta una sfida significativa per l'integrazione nei modelli di IA: è intrinsecamente difficile da codificare in quanto include componenti cognitive complesse, quali modelli mentali, credenze personali e prospettive individuali. Si tratta di una forma di sapere non articolata, inconscia e procedurale poiché, una volta acquisita, consente di eseguire determinate azioni senza dover pensare consapevolmente ai singoli passaggi coinvolti. Per questa ragione, essa può essere definita come "know-how" (letteralmente "sapere come") poiché rimanda ad una dimensione più legata al fare, in contrapposizione al "know-what" più associato a forme di conoscenza esplicite (Spender, 1996).

Dunque, è particolarmente complesso per i detentori di conoscenza tacita spiegare in modo consapevole le motivazioni e le regole decisionali che guidano le loro azioni (Ambrosini e Bowman, 2001).

In definitiva, ciò che non può essere compreso non può essere adeguatamente modellato (Pinker, 1997; 2005).

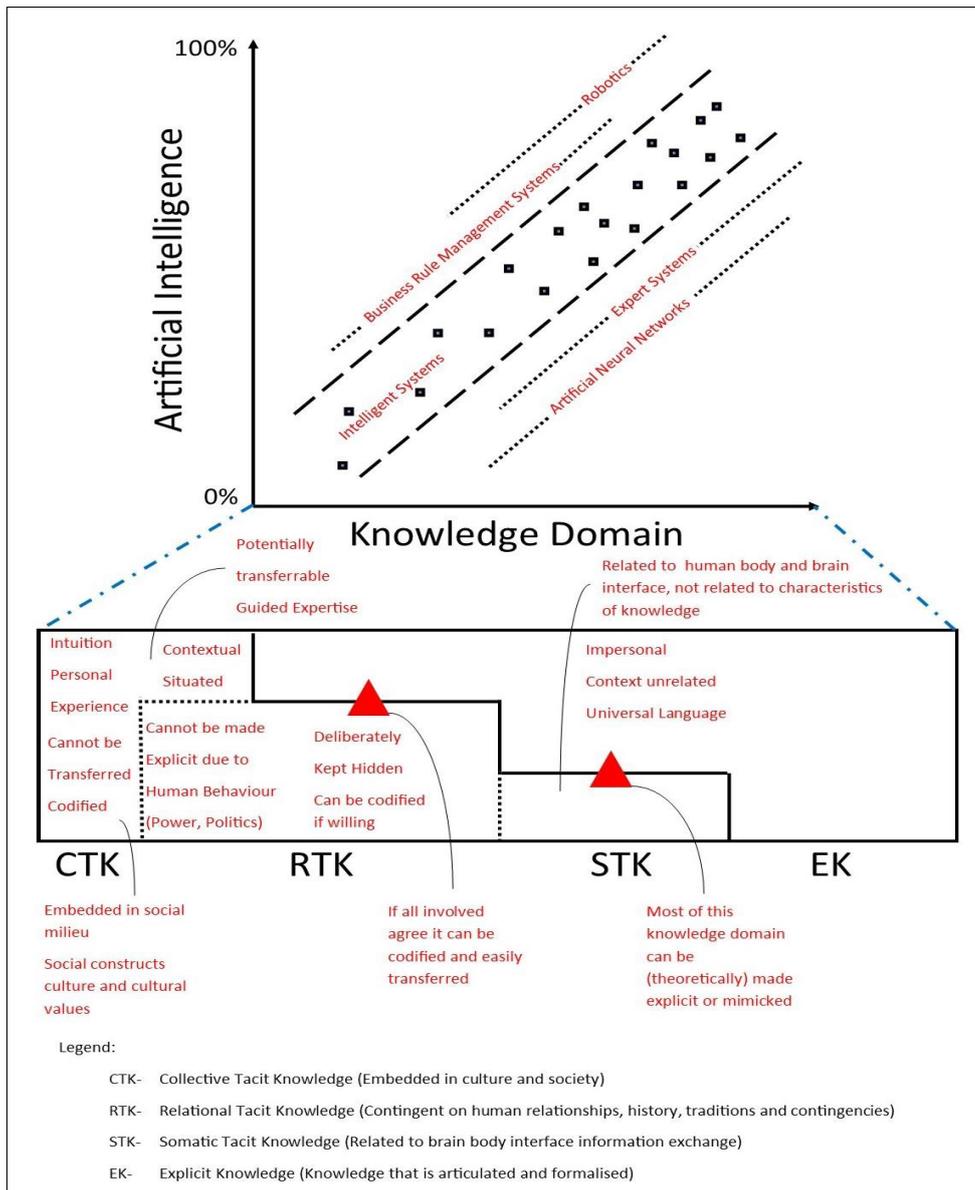


Figura 3.6 Relazione tra IA e conoscenza (Sanzogni et al., 2017)

In letteratura, vengono categorizzati differenti tipi di conoscenza, così come riportato in figura 3.6:

- *CTK, Collective Tacit Knowledge (Conoscenza tacita collettiva)*: è una forma di sapere non posseduta individualmente, ma condivisa da un gruppo, una comunità o un'organizzazione. Questa tipologia di conoscenza è profondamente radicata nei costrutti sociali, nei valori culturali e, in generale, nelle dinamiche proprie di un determinato gruppo o società e proprio a causa della sua natura intrinsecamente soggettiva e contestuale, non può essere facilmente codificata o trasmessa (es. cultura aziendale).
- *RTK, Relational Tacit Knowledge (Conoscenza Tacita Relazionale)*: riguarda la conoscenza associata alle relazioni interpersonali, alla storia condivisa, alle tradizioni e alle contingenze sociali. Risulta più trasferibile rispetto alla CTK se vi è il consenso di tutti gli individui coinvolti (es. intese sviluppate dai membri di un team senza la definizione di procedure e regole formali).

- *STK, Somatic Tacit Knowledge (Conoscenza Tacita Somatica)*: associata all'interfaccia corpo-cervello, coinvolge, ad esempio, abilità motorie che si apprendono attraverso l'esperienza (es. capacità motorie dell'atleta). È una forma di sapere che si distingue per il suo carattere impersonale e astratto e può essere teoricamente esplicitata o imitata, presentando un potenziale di trasferimento più alto rispetto alle altre forme di conoscenza tacita.
- *EK - Explicit Knowledge (Conoscenza Esplicita)*: si riferisce ad una conoscenza formalmente articolata e strutturata che quindi risulta essere facilmente trasferibile. La conoscenza esplicita può essere codificata con precisione e il suo contenuto può essere comunicato o trasmesso senza perdite significative.

Dalla figura 3.6 emerge chiaramente un legame tra l'IA e la capacità di codificare e trasferire diverse tipologie di conoscenza.

In particolare, gran parte dei successi recenti nell'ambito IA si colloca nell'area in alto a destra della figura, ossia nei contesti in cui le attività possono essere eseguite algoritmicamente attraverso forme di conoscenza codificata ed esplicita in quanto i sistemi di IA offrono una vasta gamma di opportunità per semplificare e ottimizzare attività di routine, soprattutto ripetitive e ad alto volume, come la raccolta, la classificazione, l'analisi e la presentazione di dati, che non possono essere eseguite in maniera efficiente dalle sole capacità cognitive umane.

Al contrario, l'implementazione di modelli di IA in grado di recepire e gestire forme di conoscenza tacita risulta più difficile.

Ogni individuo e, in senso più ampio, ogni organizzazione detiene conoscenza tacita, un proprio *know how* che rappresenta il bagaglio di competenze, esperienze e abilità maturate nel corso del tempo durante lo svolgimento del proprio lavoro.

Tale conoscenza costituisce a tutti gli effetti un asset immateriale di grande valore per l'organizzazione. Poiché la conoscenza tacita all'interno di un'organizzazione dipende fortemente dal modo in cui determinate azioni o procedure vengono eseguite da individui specifici, risulta fondamentale, per capire come integrare un modello di IA e per il successivo addestramento e implementazione nei processi, coinvolgere attivamente coloro che sono depositari del *know how* e quindi creare una sinergia tra tecnologia ed esperienza umana.

3.6 Influenza del DQM sulle prestazioni attese di un modello di Machine Learning

Dopo aver raccolto e filtrato i dati è necessario verificare se il modello risponde in modo adeguato e se è funzionale agli obiettivi prefissati.

Per valutare la bontà di un modello, pertanto è fondamentale assicurarsi che non si verifichino situazioni di *overfitting* e *underfitting*, poiché queste potrebbero evidentemente compromettere la capacità del modello di generare correttamente i dati attesi.

Al fine di raggiungere determinati obiettivi e poter misurare le prestazioni del modello, vengono selezionate una serie di metriche quali l'accuratezza, la precisione, il recall, l'F1 Score e il tasso di errore. È bene sottolineare che in letteratura non vi sono linee guida universalmente valide in merito al numero di metriche da adottare o ai valori ottimali che queste debbano raggiungere, poiché tali aspetti variano in base alle esigenze e agli obiettivi specifici del progetto.

Overfitting e underfitting, bias e varianza

Nel machine learning, il fenomeno dell'*overfitting* si verifica quando un algoritmo si adatta in modo eccessivamente preciso, o addirittura esatto, ai dati di addestramento (dati di training) generando un modello incapace di effettuare previsioni accurate su dati diversi da quelli utilizzati per il training. Questa criticità si presenta quando il modello è stato addestrato per un tempo troppo prolungato su dati campione o quando è caratterizzato da una complessità eccessiva, portandolo ad "imparare" anche il rumore o le informazioni irrilevanti contenute nel dataset. Di conseguenza, un modello che memorizza dettagli superflui e si adatta in modo troppo stretto al set di addestramento perde la capacità di generalizzare, cioè è incapace a "funzionare bene", risultando inadeguato per svolgere in modo efficace le attività di classificazione per cui è stato progettato.

Al contrario, l'*underfitting*, rappresenta una situazione in cui un modello non riesce a cogliere, adeguatamente, la relazione tra le variabili di input e di output generando un elevato tasso di errore sia sul set di addestramento che sui dati non visti. Tale fenomeno si verifica quando il modello è eccessivamente semplice, spesso a causa di un training insufficiente o per la mancanza di caratteristiche di input significative.

Come per l'*overfitting*, l'*underfitting* limita le prestazioni del modello e una bassa capacità di generalizzazione sui nuovi dati lo rende inefficace per attività di classificazione o di previsione. In entrambi i casi, un modello che soffre di *overfitting* o *underfitting* non può essere utilizzato efficacemente per compiti predittivi poiché non è in grado di adattarsi adeguatamente a dati nuovi e di rappresentare correttamente le informazioni rilevanti per le decisioni da supportare.

In riferimento ad *overfitting* e *underfitting* è necessario comprendere i concetti di *bias* e *varianza*.

Il *bias* quantifica l'errore introdotto da un modello quando semplifica eccessivamente un problema complesso. Un modello con elevato bias tende a fare supposizioni rigide e ad approssimare troppo semplicemente determinando *underfitting*. In tali casi, il modello non riesce a cogliere le caratteristiche fondamentali dei dati.

La *varianza*, invece, misura la sensibilità del modello ai dettagli specifici e ai piccoli cambiamenti nei dati di addestramento.

Un modello con alta varianza impara non solo il segnale effettivo presente nei dati, ma anche il rumore, ossia le informazioni irrilevanti. Questo causa *overfitting*: il modello funziona molto bene sui dati di addestramento, ma presenta scarse prestazioni sui dati nuovi o di test.

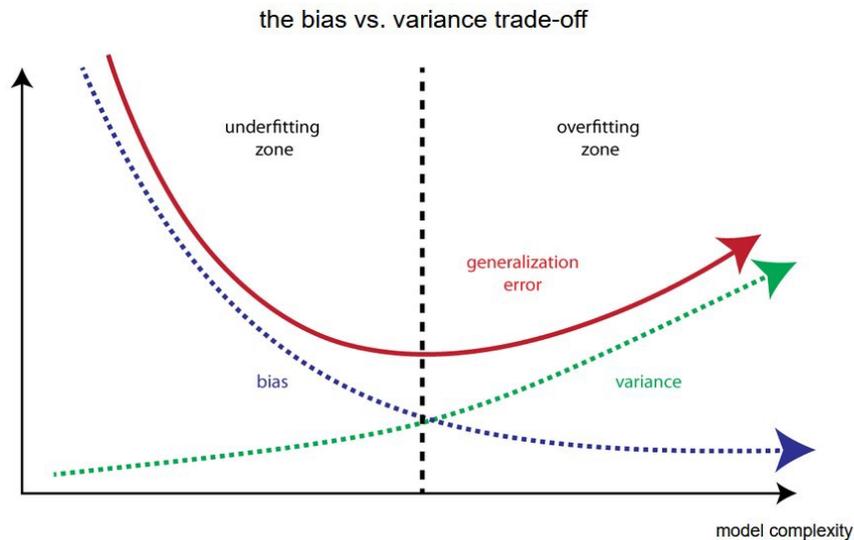


Figura 3.7 Andamento del bias e della varianza in funzione della complessità del modello

Come si evince dalla figura 3.7, bassi tassi di errore combinati con un'elevata varianza sono indicativi di *overfitting*.

Per evitare tale comportamento, è prassi comune riservare una parte del set di dati di addestramento come set di test, che viene utilizzato per verificare la presenza di *overfitting* e, se il modello presenta un basso tasso di errore sui dati di addestramento ma un alto tasso di errore sui dati di test, è un chiaro segnale di *overfitting*.

Al contrario, un elevato bias e una bassa varianza indicano la presenza di *underfitting* e questo comportamento può essere osservato direttamente durante l'addestramento del modello: i modelli soggetti ad *underfitting* sono di più agevole individuazione rispetto a quelli affetti da *overfitting*, poiché il loro scadente adattamento ai dati di addestramento è evidente fin dalle prime fasi.

Training error e validation error

In questo contesto bisogna far riferimento ai concetti di *training error* e *validation error*, nonché al loro legame con bias, varianza e complessità del modello.

Il *training error* rappresenta l'errore commesso dal modello sui dati di addestramento: man mano che la complessità del modello aumenta, l'errore di training tende a diminuire in quanto il modello diventa sempre più capace di adattarsi ai dati di addestramento grazie ad un utilizzo ottimale degli stessi.

Il *validation error* rappresenta l'errore commesso dal modello sui dati di validazione, cioè dati non "visti" durante l'addestramento. Questo errore permette di valutare quanto il modello sia in grado di generalizzare sui dati nuovi.

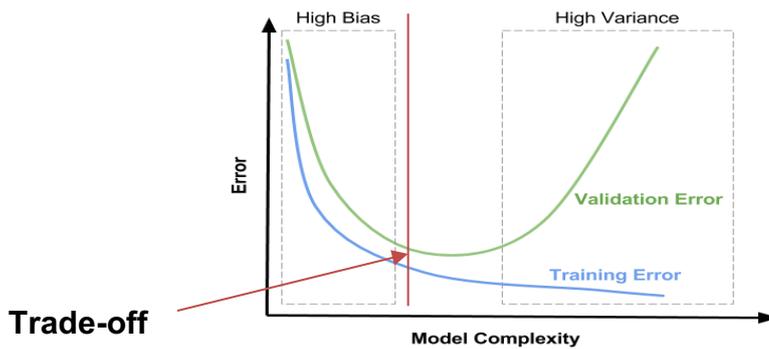


Figura 3.8 Andamento del training error e del validation error in funzione della complessità del modello

Come illustrato in figura 3.8, una bassa complessità del modello comporta valori elevati sia per l'errore di training, sia per l'errore di validazione.

Questo fenomeno si verifica poiché il modello, semplificando eccessivamente il problema, non riesce a catturare dettagli significativi nei dati, trascurando dettagli importanti e generando underfitting. Di conseguenza, la capacità di apprendimento risulta effettivamente limitata compromettendo le performance su dati nuovi.

All'aumentare della complessità del modello, si osserva una riduzione dell'errore di training, mentre l'errore di validazione inizia a crescere in quanto il modello, focalizzandosi sempre di più sui dettagli specifici dei dati di addestramento, finisce per memorizzarli perfettamente. Tale comportamento porta all'overfitting, una situazione caratterizzata da un'elevata varianza in cui il modello fatica a generalizzare e ad adattarsi ai dati nuovi.

È importante sottolineare che un modello non può presentare simultaneamente underfitting e overfitting poiché non è possibile che un modello classifichi con precisione i dati di addestramento e, al contempo, fallisca nella generalizzazione, o viceversa.

Influenza del data cleaning e della quantità di dati fornita sulle prestazioni del modello

Come precisato nel capitolo 3.3 il data cleaning, che si pone come la fase più importante nel processo di Data Quality Management, influisce significativamente sulle prestazioni del modello di ML. Questo passaggio è particolarmente critico perché consente di evitare situazioni di underfitting ed overfitting.

Nel caso in cui un modello è eccessivamente semplice rispetto all'obiettivo che si vuole raggiungere, presentando underfitting, il data cleaning aiuta a garantire che i dati forniti siano rappresentativi delle caratteristiche del problema; tuttavia, anche una pulizia accurata dei dati, unita ad una maggiore quantità in input, non assicura che un modello possa risolvere le proprie criticità, rendendosi pertanto necessari ulteriori interventi, come l'applicazione di tecniche di feature engineering, cioè arricchendo le caratteristiche informative per aumentare la capacità del modello ad apprendere, riducendo l'underfitting.

Il data cleaning ha un'influenza significativa anche sull'overfitting: la pulizia del dataset permette al modello di concentrarsi solo su caratteristiche rilevanti, inoltre, aumentando la quantità dei dati, soprattutto se diversificati, attraverso tecniche di data augmentation (capitolo 3.3), riduce la possibilità che il modello memorizzi dettagli specifici, migliorando quindi la propria generalizzazione.

Tabella 3.2 Sintesi degli elementi che contribuiscono al miglioramento delle performance di un modello di ML

	Underfitting	Overfitting
<i>Complessità del modello</i>	Aumentare la complessità del modello	Ridurre la complessità del modello
<i>Quantità di caratteristiche</i>	Aumentare il numero di caratteristiche attraverso il feature engineering, consentendo al modello di avere più informazioni per fare previsioni	Ridurre il numero di caratteristiche, eliminando le feature superflue applicando tecniche di soluzione delle caratteristiche per rimuovere quelle irrilevanti
<i>Dati</i>	Utilizzare approcci di data cleaning (discusso al capitolo 3.3), hold.out validation e cross validation). Aumentare la quantità di dati non sarà utile se il modello è strutturalmente incapace di apprendere relazioni complesse.	Utilizzare approcci di data cleaning (discusso al capitolo 3.3), hold.out validation e cross validation). Aumentare la quantità di dati sarà utile (ad esempio attraverso tecniche di data augmentation) soprattutto se i nuovi dati sono diversificati e rappresentativi, poiché consentono al modello di generalizzare meglio.

Metriche per la valutazione delle prestazioni di un modello di ML

Per la valutazione accurata di un modello di Machine Learning nella Computer Vision come quello che verrà presentato nel caso studio e per comprendere la qualità dei dati in output, è fondamentale ricorrere a metriche quantitative.

Le principali metriche sono: l'*accuratezza*, la *precisione*, il *punteggio F1* e il *recall* (o sensibilità) (Manmeet Kaur, 2024).

Si noti che nessuna singola metrica può fornire un quadro completo delle prestazioni del modello e che è necessario utilizzare una combinazione di metriche per valutare il modello in modo efficace e in base alle specifiche esigenze.

Prima di descrivere in dettaglio le metriche appena menzionate è importante definire alcuni concetti fondamentali relativi ad un modello di classificazione di computer vision:

- *TP (True Positives)*: numero di immagini correttamente classificate come appartenenti ad una certa classe;
- *TN (True Negatives)*: numero di immagini correttamente classificate come non appartenenti ad una certa classe;
- *FP (False Positives)*: numero di immagini non appartenenti alla classe corretta, ma classificate come appartenenti alla classe corretta. Questi sono comunemente chiamati anche “falsi allarmi”, poiché il modello rileva in modo errato la presenza di caratteristiche riconducibili alla classe corretta.
- *FN (False Negatives)*: numero di immagini appartenenti alla classe corretta classificate erroneamente come non appartenenti alla classe corretta. Questo tipo di errore è noto come “errore di omissione” poiché il modello non riesce a riconoscere correttamente le immagini che dovrebbero essere classificate come appartenenti ad una certa classe.

Tabella 3.3 Principali metriche per la valutazione di modelli di ML di classificazione

Metrica	Formula	Descrizione
<i>Accuratezza</i>	$Acc = \frac{TP + TN}{TP + TN + FP + FN}$	<p>L'accuratezza misura quanto un modello è generalmente corretto.</p> <p>Rappresenta la percentuale di immagini correttamente classificate (sia in senso positivo che negativo) sul totale degli esempi.</p> <p>È possibile calcolare il tasso di accuratezza come numero totale di stime corrette e dividerlo per il numero totale del set di dati. Il tasso di accuratezza peggiore è rappresentato da un valore nullo, mentre il tasso di precisione ideale è costituito da un valore unitario.</p> <p>È inoltre possibile calcolare il tasso di precisione con 1-ERR.</p>
<i>Precisione</i>	$Precisione = \frac{TP}{TP + FP}$	<p>La precisione misura la percentuale di immagini correttamente classificate ed è funzionale a raggiungere l'obiettivo di minimizzare i falsi positivi.</p> <p>La precisione risponde alla domanda "Di tutte le immagini correttamente classificate, quante effettivamente appartenevano a quella classe?"</p>
<i>Recall</i>	$Recall = \frac{TP}{TP + FN}$	<p>Il recall tiene conto dei falsi negativi, invece che dei falsi positivi come la precisione; viene chiamato anche sensibilità, in quanto all'aumentare del valore che assume, il modello diventa sempre meno preciso e classifica classi negative come positive. È particolarmente utile quando l'obiettivo è ridurre i falsi negativi (ad esempio nelle diagnosi mediche). Tale metrica risponde alla domanda "Di tutti i veri positivi, quanti sono stati correttamente rilevati come positivi?". Un valore del tasso di richiamo pari a 0 rappresenta il valore peggiore, mentre un valore tendente ad 1 rappresenta il migliore.</p>
<i>F1 Score</i>	$Recall = \frac{2}{\frac{1}{Recall} + \frac{1}{Precision}}$	<p>L'F1 Score è la media armonica tra precisione e recall.</p> <p>Combinando queste due metriche in un singolo valore. Se lo score F1 aumenta, vuol dire che il modello ha aumentato le performance per precisione, richiamo o per entrambi.</p>
<i>Tasso di errore</i>	$ERR = \frac{FN + FP}{TN + TP + FP + FN}$	<p>È possibile calcolare il tasso di errore come il numero totale di previsioni errate e dividerlo per il numero del set di dati. Il tasso di errore peggiore è rappresentato da un valore unitario, mentre il tasso di errore ideale è rappresentato da un valore nullo.</p>

Capitolo 4

Progetto D.A.R.E.

4.1 Descrizione progetto

Il progetto DARE si concentra sull'applicazione dell'Intelligenza Artificiale e in particolare di algoritmi di Machine Learning e Computer Vision per il miglioramento della gestione del rischio e della valutazione dei sinistri nel segmento “non auto” del settore assicurativo.

Mentre in altri segmenti assicurativi, dove i processi risultano più standardizzati, l'adozione di soluzioni IT ha già portato miglioramenti in termini di efficienza, nel segmento “non auto”, che include incidenti legati all'abitazione (danni da bagnamento, danni ai vetri e alle insegne, danni causati dai ladri, ecc.) i risultati finora ottenuti sono stati limitati e se si tiene in considerazione che tale segmento rappresenta una porzione ampia e rilevante del settore assicurativo, un ulteriore sviluppo in questo ambito risulta determinante per il futuro del settore stesso.

Il motivo per cui questo ambito di applicazione è particolarmente critico per le compagnie assicurative è dovuto al fatto che la gestione dei sinistri è attualmente un'attività in cui competenze specifiche e conoscenze esperienziali costituiscono i requisiti fondamentali per identificare gli elementi caratterizzanti il sinistro, classificarli, prioritizzarli e per poi produrre stime accurate e coerenti. La variabilità dei danni e la complessità delle perizie rendono indispensabile una trasformazione radicale dei processi di valutazione, con particolare focus ai *sinistri a basso valore e alta frequenza*.

Nonostante gli investimenti da parte di diverse startup siano mirati a sviluppare tecnologie IA in questo ambito, i risultati ottenuti non sono soddisfacenti: le soluzioni esistenti presentano un livello limitato di integrazione nei processi di gestione dei sinistri e una convalida aziendale molto limitata. Questi scarsi risultati sono dovuti principalmente a due fattori: le startup hanno una scarsa conoscenza delle “regole di business” e non dispongono di grandi quantità di dati, al contrario delle compagnie assicurative, per l'addestramento degli algoritmi di IA.

Il progetto D.A.R.E si focalizza, quindi, sulle *dimensioni strategiche e tecnologiche* dell'adozione dell'IA nel settore assicurativo e, attraverso la combinazione di tecnologie, dati e competenze specifiche di settore, si pone lo scopo di identificare e coordinare azioni che spaziano dalla ricerca tecnologica allo sviluppo di nuovi algoritmi, dalla selezione e integrazione dei dati alla progettazione delle fasi di sviluppo necessarie per raggiungere una completa adozione dell'IA nel processo di gestione dei sinistri, dalla definizione di nuove metriche di performance

alla chiara definizione delle dimensioni di creazione di valore per le aziende coinvolte e per il loro ecosistema. Tutti questi elementi devono garantire prestazioni adeguate, sicurezza e privacy dei dati. Esso, inoltre, mira ad ottenere benefici sia a lungo che a breve termine: infatti, si prevede che esso possa creare le basi tecniche e organizzative e le condizioni per affrontare problemi più complessi e su scala più ampia, avviando una trasformazione profonda che avrà un impatto rilevante sui processi lavorativi, sulle competenze richieste e sui meccanismi di creazione del valore.

Come precisato, l'IA è ormai divenuta indispensabile nel settore assicurativo per operare una trasformazione radicale dell'intero processo di gestione dei sinistri, a partire dall'identificazione del cliente, proseguendo con l'analisi delle immagini e dei dati forniti dai clienti stessi relativi danno, fino alla valutazione precisa e rapida del danno. Questo consente di accelerare la fase di liquidazione e di chiudere il sinistro in tempi ridotti, migliorando sensibilmente l'esperienza del cliente e garantendo maggiore efficienza operativa. Nello specifico, gli obiettivi dell'integrazione dell'IA nel progetto sono i seguenti:

- *Ridurre al minimo il fattore umano nei sinistri “ad alta frequenza” e “di basso valore”*, consentendo alle risorse di concentrarsi su danni e richieste di risarcimento più rilevanti. L'obiettivo è quello di ridurre i tempi di rimborso per i sinistri minori a meno di 5 giorni lavorativi, garantendo un adeguato livello di precisione nella valutazione dei danni: ciò è possibile trasformando in modo mirato l'intero processo di gestione dei sinistri.
- *Aumentare la soddisfazione del cliente*, permettendo una gestione più semplice ed autonoma del sinistro attraverso l'invio diretto di immagini e l'esecuzione di azioni in modo rapido. Un'esperienza rapida e senza ritardi è fondamentale, in quanto l'ultima cosa che un cliente che ha subito danni alla proprietà desidera è un ritardo nella valutazione e nel pagamento dell'assicurazione.
- *Modificare i processi aziendali interni*, le pratiche di lavoro, e sviluppare un nuovo livello di integrazione dei dati. Con l'adozione dell'IA, l'ente assicuratore potrà gestire autonomamente, senza l'intervento di esperti (e quindi di alcun costo), l'intero ciclo di vita del sinistro, dall'apertura del reclamo, alla valutazione e rimborso.

4.2 Requisiti per lo sviluppo di un algoritmo di IA

L'introduzione dell'IA nei processi aziendali dell'ente assicuratore richiede un'analisi delle procedure e attività specifiche. Le principali fasi del processo di integrazione tra le due componenti comprendono:

- *Acquisizione e trasferimento delle competenze e del know-how* degli attori coinvolti nel processo (ontologia, comprensione del flusso di lavoro, output attesi, ecc.).
- *Mappatura del processo AS-IS*: uno studio approfondito dell'attuale stato del processo permette di individuare eventuali criticità e punti di miglioramento.
- *Design del processo TO-BE*: design del futuro processo ottimizzato con l'integrazione dell'IA.
- *Identificazione e reperimento del dataset*: selezione e raccolta dei dati necessari per l'addestramento dell'algoritmo.
- *Addestramento dell'algoritmo di IA*: l'utilizzo di dataset strutturati e di grandi dimensioni e di dati etichettati permette un training efficace dell'algoritmo, assicurando il trasferimento del know-how umano agli strumenti IA.
- *Testing in ambiente reale del modello*: è fondamentale verificare le prestazioni dell'algoritmo in condizioni operative reali.
- *Utilizzo dei nuovi tool da parte dell'utente finale*.

Acquisizione e trasferimento delle competenze e del know-how

Il know-how attualmente è detenuto da periti e liquidatori le cui competenze e conoscenze risultano cruciali per il funzionamento dei processi aziendali. Il trasferimento di tale know how ai partner di progetto è stato realizzato attraverso mirate interviste che hanno permesso di comprendere dettagliatamente il flusso di lavoro e i ruoli svolti all'interno dell'organizzazione; tali informazioni sono state oggetto di un'analisi dettagliata, volta ad identificare metriche chiave e punti critici di interesse.

Studio processo AS-IS

La comprensione del processo AS-IS, identificando attori, fasi, ruoli, attività nel processo e gestione dei dati, è la base di partenza per progettare l'integrazione della tecnologia dell'IA.

Le fasi principali di tale processo sono le seguenti: denuncia sinistro, apertura sinistro, assegnazione del perito, istruttoria peritale, valutazione e liquidazione sinistro. In ognuna di queste fasi del processo operano diverse figure, alcune interne e altre esterne all'ente assicuratore che svolgono attività specifiche:

- *Call center/backoffice*: ha un ruolo di inserimento e controllo delle informazioni nella fase di apertura sinistro.
- *Network di fiduciari (periti)*: identifica e stima il danno facendo un controllo preliminare sulle garanzie colpite e a perizia conclusa ogni perito eroga una relazione finale
- *Liquidatori*: effettuano una verifica sull'intero processo, approvano o rifiutano la liquidazione

Durante le varie fasi del processo, vengono scambiati dati e documenti quali i *dati relativi all'assicurato* e le *coperture possedute*, descrizione e fotografie del sinistro e, infine, la relazione peritale finale.

Nell'analisi del processo AS-IS, tuttavia, sono stati riscontrati degli elementi che vanno attenzionati. In particolare, sono stati identificati alcuni *colli di bottiglia*: alcune attività, soprattutto quelle di verifica, vengono eseguite in maniera ridondante da più attori all'interno del processo, determinando un impiego di risorse economiche non ottimale per l'assicuratore. Ad esempio, il controllo sulla conformità delle garanzie possedute dal cliente viene effettuato sia dal backoffice che dal perito e liquidatore. Inoltre, nessun sinistro viene aperto automaticamente, poiché ogni pratica richiede un controllo e una conferma manuale da parte di un operatore, quindi il tempo di gestione dell'intero processo, dall'apertura sinistro alla liquidazione, risulta influenzato da percorsi critici non ottimizzati. Non vi è, inoltre, alcuna distinzione in termini di tempo impiegato per la gestione dei sinistri a basso impatto rispetto a quelli ad alto impatto. La causa principale di questa inefficienza risiede nell'assenza di automatismi affidabili per le parti di processo ripetitive o che richiedono un livello di controllo standardizzato.

Per quanto riguarda il *ruolo delle fotografie*, esso è attualmente limitato nel processo decisionale: le fotografie vengono utilizzate principalmente per documentare l'attività del perito e comprovare le decisioni prese; non sono, dunque, integrate nel flusso decisionale in modo significativo. Durante la fase di sottoscrizione della polizza non viene richiesta alcuna documentazione dell'immobile che potrebbe garantire lo stato attuale della proprietà in quel momento o la sua tipologia: questa carenza priva l'ente assicurativo di una documentazione visiva importanti e di informazioni utili nella gestione futura dei sinistri.

Infine, sono rilevabili delle criticità nell'attuale processo relative alla *governance e data quality*: la gestione di alcune fasi del processo, come identificazione e stima del danno, varia molto a seconda del network o del perito incaricato, limitando il controllo dell'ente assicuratore sulle perizie, specialmente se affidate in outsourcing. Ciò avviene perché non esistono procedure standardizzate per la raccolta, condivisione e archiviazione dei dati e delle fotografie; non sono state definite linee guida comuni per stabilire come le fotografie debbano essere scattate in termini di soggetti, scala, risoluzione e ordine temporale. Non vi sono, dunque KPI che misurano la qualità dei dati e in particolare delle fotografie (spesso vengono fornite fotografie in bassa risoluzione e non certificate, cioè prive di localizzazione, proprietà, soggetto ecc.).

Inoltre, i network peritali, pur seguendo le linee guida fornite dall'ente assicuratore, agiscono con una certa autonomia nella modalità di organizzazione e presentazione delle informazioni all'interno dei documenti peritali per cui ne consegue una mancanza di uniformità che rende difficile il confronto tra le perizie prodotte da diversi network, provocando di fatto, una scarsa coerenza tra i dataset raccolti che ne limita, e addirittura compromette, l'usabilità nell'addestramento dell'IA.

Studio processo TO-BE

Come precedentemente detto, con l'integrazione di un tool basato sull'IA nel processo di liquidazione dei sinistri di media e bassa complessità e alta frequenza, si vogliono ridurre bias ed

errori umani per migliorare l'efficienza e la precisione. Tale integrazione richiede cambiamenti strutturali significativi che comportano l'introduzione di nuove fasi e operazioni del processo.

Ad esempio, le fasi del processo di maggior impatto saranno quelle relative alla sottoscrizione della polizza, all'apertura del sinistro e al controllo delle garanzie fino all'analisi e stima del danno.

Oltre all'acquisizione di informazioni anagrafiche sull'utente e sull'abitazione, come indirizzo e quadratura dell'abitazione (già previste nel processo AS-IS), sarà indispensabile certificare l'autenticità e la proprietà delle immagini caricate dall'utente stesso che dovranno essere corredate da una visura catastale dell'immobile (le cui informazioni verranno verificate da un operatore di backoffice), unitamente alle coordinate GPS abbinate all'indirizzo dell'immobile e fotografie certificate e localizzate.

L'analisi automatica permetterà di fare una stima portando ad una veloce quantificazione del danno, il tutto con un nuovo sistema che sarà fruibile dal cliente finale attraverso canali digitali.

Nonostante ciò, nel processo to-be di breve periodo rimarrà comunque la figura di un decisore che valuterà gli output Explainable dell'IA e prenderà, quindi, la decisione finale.

Identificazione e reperimento del dataset

Per effettuare il training dell'IA è necessario reperire un campione significativo di perizie al cui interno sono presenti tutti i dati ed informazioni riguardanti il sinistro/danno.

In particolare, i dati grazie ai quali è stato possibile svolgere le attività di annotazione sono stati forniti da due network peritali con cui l'ente assicuratore collabora

Per quanto riguarda *l'addestramento dell'algoritmo di IA*, esso, allo stato attuale, risulta in corso d'opera. A seguito, del training e del raggiungimento di performance soddisfacenti verrà effettuato il *testing in ambiente reale del modello*, infine poi verrà osservato *l'utilizzo dei nuovi tool da parte dell'utente finale* con l'obiettivo di apportare cambiamenti via via migliorativi.

4.3 Framework per il training e test dell'IA nel progetto

Tradizionalmente i problemi di Computer Vision come, ad esempio, il riconoscimento di oggetti e la segmentazione di istanze, vengono risolti per mezzo del cosiddetto *transfer learning*, una tecnica che sfrutta modelli pre-addestrati (come quelli su ImageNet) e trasferisce le caratteristiche apprese su nuovi compiti e domini. Nonostante l'efficacia, il principale limite di questo approccio risiede nella disponibilità di dataset annotati manualmente in modo specifico per il nuovo dominio di interesse, un processo spesso lungo e costoso che implica l'intervento di personale qualificato. Ciò costituisce un problema che è particolarmente rilevante in contesti specifici o settori di nicchia dove i dataset annotati sono scarsi o addirittura inesistenti.

Per superare queste limitazioni, un approccio promettente per addestrare i modelli consiste nell'associare, come nel caso di studio, informazioni testuali, didascalie derivanti dai report peritali alle immagini riducendo la quantità di annotazioni visive necessarie. Questo tipo di apprendimento supervisionato aiuta a fornire al modello una forma di comprensione semantica più ampia e dettagliata rispetto alle sole etichette visive, semplificando e velocizzando il processo di addestramento.

In figura 4.1 viene mostrato il funzionamento del processo di apprendimento del modello, nell'ambito del progetto D.A.R.E., durante le fasi di training e test.

In generale, il modello è composto da due componenti principali: un *backbone testuale* e un *backbone visivo* che elaborano i dati forniti in input costituiti da *coppie di immagini e didascalie*.

Il backbone testuale è una rete neurale che apprende la semantica delle parole e delle frasi nella didascalia, cercando di comprendere i concetti chiave legati al danno descritto, mentre il backbone visivo è una rete neurale convoluzionale che analizza l'immagine estraendo le caratteristiche visive rilevanti, come le texture, i colori o le forme che potrebbero indicare un danno specifico. I risultati di entrambe le reti vengono combinati e utilizzati per classificare i tipi di danni.

Nel contesto del progetto D.A.R.E., l'addestramento del modello è supportato dalla creazione di *dati annotati* specifici per il dominio, cioè immagini in cui sono state applicate *maschere di segmentazione* che permettono di isolare le aree danneggiate, e descrizioni testuali.

Questo processo di annotazione è stato realizzato tramite un tool dedicato, progettato appositamente per il progetto poiché non esistono (o comunque limitatamente) dataset di immagini annotate in questo specifico settore. Nella fase di test, il modello viene valutato su immagini che rappresentano nuovi casi di danno, simili ma non esattamente presenti nei dati di addestramento. Ad esempio, un'immagine relativa ad un danno da bagnamento non presente nei dati di addestramento viene processata dal backbone visivo già addestrato; le caratteristiche estratte dall'immagine permettono al modello non solo di identificare e classificare correttamente il danno, ma anche generare una maschera di segmentazione che evidenzia l'area danneggiata.

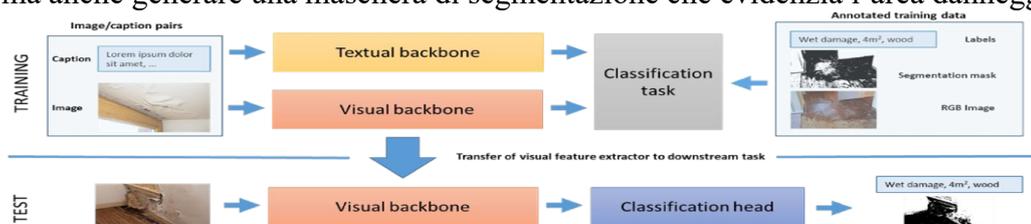


Figura 4.1 Framework per il training e test del modello di IA nel progetto

4.4 Metodologia per il controllo dell'incertezza interpretativa e per il data cleaning

Uno dei principali ostacoli nell'addestramento di modelli di ML è rappresentato dalla difficoltà di reperire set di dati sufficientemente ampi e ben annotati, indispensabili per garantire un apprendimento efficace e risultati affidabili.

Nel settore assicurativo, sebbene vi sia una vasta quantità di dati non strutturati e semi-strutturati, questi richiedono una significativa elaborazione per poter essere impiegati, come nel caso studio, nell'implementazione di modelli di IA.

Il progetto, infatti, si pone l'obiettivo di trasformare i dati grezzi a disposizione in risorse strutturate e utili per controllare l'incertezza interpretativa e quindi per addestrare più efficacemente i modelli di ML.

Per raggiungere tale scopo, è stato adottato un approccio metodologico articolato in diverse fasi, ciascuna delle quali mira a risolvere specifiche criticità legate ai dati e alla loro applicazione, tali fasi sono:

- *Raccolta e analisi dei dati;*
- *Definizione di un'ontologia di dominio;*
- *Data cleaning e segmentazione semantica per il training del modello;*
- *Statistiche del campione analizzato;*

In ultima analisi, il progetto si prefigge di delineare in maniera chiara e dettagliata l'intero processo di gestione dei sinistri, specificando ogni fase operativa e comprendendo in sé l'integrazione di tutte le tecnologie necessarie per la completa digitalizzazione dello stesso.

Raccolta e analisi dei dati

Come illustrato in precedenza, i dati acquisiti riguardanti circa 18mila perizie sono stati forniti da due network peritali con cui l'ente assicuratore collabora.

Tali dati sono costituiti principalmente da fotografie (in formati standard contenenti dettagli relativi a danni da bagnamento), file Excel e Word, PDF e CSV, contenenti tutte le informazioni rilevanti delle perizie, tra cui coperture assicurative coinvolte, descrizione del danno e del contesto del sinistro, valutazione economica del danno e tipologia dell'immobile interessato. Ciascun network peritale presenta caratteristiche specifiche riguardanti l'organizzazione e la struttura dei dati forniti.

Per ogni perizia gestita dal *network peritale 1*, è presente una cartella associata ad un codice identificativo univoco. All'interno della cartella sono presenti le immagini relative al sinistro. La relazione peritale, invece, non è inserita direttamente nella cartella ma è disponibile in un file Excel contenente tutte le relazioni peritali per cui, per accedere alla relazione specifica di una perizia, è necessario inserire un codice univoco identificativo. All'interno della relazione peritale è possibile reperire le seguenti informazioni:

- *Le circostanze del sinistro:* viene descritta la dinamica del sinistro, ricostruita sulla base delle dichiarazioni dell'assicurato e degli accertamenti peritali.
- La presenza o meno di una *coassicurazione*.
- La *conformità del rischio*.

- La *conclusione del sinistro*: viene riportato se il danno accertato è riconducibile o meno ad una garanzia della polizza e, in caso affermativo, viene indicato il valore dell'indennizzo.

Nel caso *del network peritale 2* per ogni perizia è presente una cartella che contiene sia tutte le immagini del sinistro, sia un file Excel correlato, organizzato in quattro fogli, ognuno dei quali riporta informazioni specifiche:

- Foglio 1: le informazioni presenti sono le seguenti:
 - *Dati del sinistro*: si riportano le informazioni relative all'assicurato (es. numero identificativo, ubicazione, ecc.) alla data di denuncia del sinistro, alla garanzia colpita, alla data del sopralluogo effettuato.
 - *Definizione del sinistro*: si riportano le informazioni in merito all'entità del danno richiesto, al danno accertato e alla franchigia/scoperto
 - *Descrizione del rischio e sua valutazione*: vengono descritte le caratteristiche costruttive del fabbricato e le conformità o meno del rischio alle descrizioni contrattuali
 - *Notizie sulla dinamica del sinistro*
 - *Garanzie interessate ed elementi di polizza*
 - *Operazioni peritali*
- Foglio 2: contiene una *stima del danno*
- Foglio 3: riporta le osservazioni e conclusioni del perito
- Foglio 4: fornisce una sintesi delle informazioni contenute nei fogli precedenti, offrendo una visione d'insieme della perizia.

Definizione di un'ontologia di dominio

Al fine di ottenere una maggiore omogeneità e qualità dei dati acquisiti dalle diverse fonti e schematizzare il know how dei periti (così come discusso nel capitolo 3.5), è stata definita un'opportuna *ontologia di dominio*.

Un'ontologia è una rappresentazione strutturata di concetti che mira a definire formalmente le entità, le proprietà, le funzioni e le relazioni all'interno di uno specifico dominio, rappresenta una base semantica condivisa e formale per comprendere, analizzare e categorizzare con precisione oggetti e caratteristiche delle immagini acquisite in tempi e da soggetti diversi, evitando quindi così incertezza interpretativa.

La creazione di un'ontologia è un processo complesso in cui gli esperti forniscono basi concettuali e conoscenze. Successivamente tali elementi vengono tradotti in una struttura formale.

Il primo passo per la creazione di un'ontologia è definire chiaramente gli obiettivi e l'ambito del dominio da rappresentare. Nel caso del progetto D.A.R.E, danni da bagnamento.

Una volta stabiliti questi elementi si procede con l'analisi della documentazione disponibile, come le perizie assicurative, per individuare concetti ricorrenti, terminologia e strutture comuni, una fase questa che consente di comprendere come i concetti vengono utilizzati e relazionati nel dominio specifico.

Successivamente, si conducono interviste con esperti del settore per risolvere ambiguità e per raccogliere in generale ulteriori informazioni, soprattutto provenienti da conoscenze tacite che non emergono chiaramente dall'analisi della documentazione. L'esperienza degli esperti si rivela cruciale per costruire una rappresentazione ontologica completa e precisa.

Una volta raccolte le conoscenze necessarie, si definisce un modello concettuale generale in cui vengono definiti concetti principali, classi e proprietà. Tale modello viene dunque trasformato in una struttura formale ove vengono definite classi, sottoclassi, ulteriori proprietà e relazioni.

Al fine di garantire che l'ontologia possa rispecchiare la realtà del dominio, essa viene aggiornata correggendo errori e omissioni e acquisendo continui feedback degli esperti di dominio.

Nel caso del progetto D.A.R.E. le classi principali definite sono quelle riportate sotto nella tabella 4.1 e rappresentano, quindi, gli elementi che effettivamente vengono considerati d'interesse nel processo di data cleaning e di segmentazione semantica.

L'ontologia, in ultima analisi, permette la standardizzazione dei dati e, nel caso in questione, la *standardizzazione del soggetto dell'immagine*: infatti, in assenza di un'ontologia, due immagini che rappresentano lo stesso fenomeno potrebbero essere classificate in modo diverso, compromettendo la coerenza e l'affidabilità del dataset.

Tabella 4.1 Esempi di elementi contenuti nell'ontologia.

Causa	Effetto	Elemento danneggiato
Rottura tubo mandata	Macchia chiara	Plafone
Rottura tubo scarico	Macchia scura	Parete
Rottura tubo scarico acque grigie	Muffa chiara	Pavimento
Rottura tubo scarico acque nere	Muffa scura	Infissi
Rottura tubo carico	Fughe bagnate	Piastrelle
Rottura tubo carico acqua calda	Spargimento di liquidi	Porte
Rottura tubo carico acqua fredda	Piastrelle gonfie	Battiscopa
Rottura tubo riscaldamento classico	Distaccamento piastrelle	Tinteggiatura
Rottura tubo riscaldamento a pavimento	Sollevamento parquet	Intonaco
Rottura tubo condensa condizionatori	Gocciolamento	
Rottura tubo altro	Scrostamento	
Rottura grondaie	Distacco tinteggio	
Rottura pluviali		
Infiltrazione		
Rottura guarnizione		
Occlusione		
Rottura cassetta wc		
Causa non identificabile		

Data cleaning e segmentazione semantica per il training del modello

Per un efficace implementazione dell'IA, in analogia con il capitolo 3, è fondamentale disporre di dati di alta qualità, ben strutturati e rappresentativi della realtà: i dati raccolti, quindi, sono stati sottoposti inizialmente ad un processo di *data cleaning*, specifico per il progetto, che va ad integrare il framework presentato al capitolo 3.2 e consiste nel rimuovere dati inaccurati, falsi, incompleti, corrotti o irrilevanti dal dataset.

I dati inizialmente analizzati sono stati sottoposti, in prima battuta, ad un *processo di anonimizzazione* che ha permesso di eliminare informazioni relative ai titolari delle polizze, come nome, cognome, codice fiscale, IBAN, mantenendo, tuttavia, l'indirizzo generico dell'immobile (comune o via senza numero civico) dell'assicurato per condurre statistiche e indagini relative a frodi e bias. Tale approccio prende il nome di *Privacy by Design* (PbD) e mira a proteggere la privacy e i dati personali degli utenti al fine di prevenire rischi connessi alla manipolazione di tali dati sensibili e al loro trattamento in conformità alle procedure e alla normativa vigente GDPR.

Successivamente, le immagini di ogni perizia sono state analizzate e classificate singolarmente sulla base dell'ontologia grazie all'utilizzo di uno specifico tool, appositamente realizzato per il progetto, che ha anche permesso di applicare *maschere semantiche* (processo di *segmentazione semantica*) sulle immagini più rappresentative del danno, isolandolo; in tal modo è stato possibile delimitare in modo più rigoroso il soggetto vero e proprio dell'immagine.

Inoltre, ogni informazione e caratteristica relative alla singola perizia sono state riportate all'interno di un file Excel su cui poi sono state elaborate delle analisi e statistiche, riportate nel capitolo successivo.

In particolare, le perizie sono state “ripulite” e annotate nel seguente modo:

- Ogni immagine presente all'interno della singola perizia è stata classificata come *in contesto*, *fuori contesto* o *ambigua*.
- Per le foto *fuori contesto* è stato possibile specificare una tra le seguenti sottoclassi:
 - *Building*: foto che rappresentano edifici o porzioni di essi che non contengono dettagli pertinenti al danno.
 - *Human interaction*: foto di persone coinvolte in attività o interazioni non rilevanti.
 - *Selfie*: foto-autoscatto che ritraggono molto spesso gli assicurati durante le videoperizie.
 - *Screen capture*: foto-screenshot che contengono chat o altro.
 - *Documents*: foto di documenti cartacei o in formato digitale.
 - *Random stuff*: foto che ritraggono oggetti o elementi casuali o qualsiasi foto fuori contesto che non rientra nelle altre categorie specifiche.

Ad esempio, le immagini appartenenti alla categoria *building*, che ritraggono edifici interi o porzioni di essi (es. facciate), sono state considerate come *fuori contesto* in quanto trattasi di foto che vengono richieste e documentate per dimostrare e controllare l'ubicazione e la conformità del rischio dell'assicurato rispetto a quanto dichiarato in polizza. Pertanto, ragionando in ottica futura, al momento della sottoscrizione della polizza, l'assicurato dovrà fornire informazioni più dettagliate sull'abitazione in modo tale che con i nuovi tool

le immagini possano essere arricchite e autenticate anche con la posizione GPS. Questo overload di foto verrà quindi ridotto.

- Per le *foto in contesto* è stato possibile selezionare tra:
 - *Relevant*: immagine in contesto e qualitativamente buona
 - *Blurred*: immagine in contesto ma sfocata
 - *Underexposed*: immagine in contesto ma sottoesposta
 - *Overexposed*: immagine in contesto ma sovraesposta
- Le immagini in cui non è stato possibile stabilire se appartenessero alla categoria in contesto o fuori contesto sono state identificate come *ambigue*
- Per ogni foto, inoltre, è stato indicato il *timing* relativo alla fase di lavorazione della parte danneggiata. Per le foto è stato possibile selezionare tra:
 - *Pre intervention*: se l'immagine ritrae un danno nel suo stato originale (es. fughe bagnate)
 - *During intervention*: se l'immagine ritrae lavori d'intervento (es. piastrelle spaccate)
 - *Post intervention*: se l'immagine ritrae il ripristino dello stato della parte precedentemente danneggiata
 - *Unknown*: questa sottoclasse è stata selezionata nel caso in cui non sia stato possibile stabilire, attraverso l'osservazione del soggetto riportato in foto, il timing della fase di lavorazione oppure nel caso in cui ci si sia imbattuti in una foto fuori contesto o ambigua.

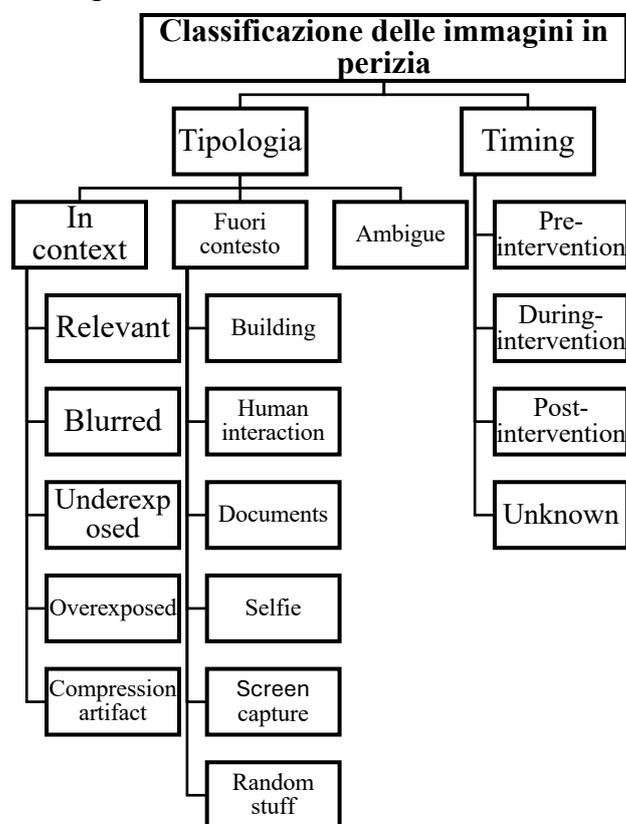


Figura 4.2 Schema di classificazione delle immagini in perizia

- Ove possibile è stata indicata la *location della fotografia*.
- Nella parte inferiore, ove possibile, è stata riportata tra parentesi quadre, una breve didascalia dell'immagine con termini dell'ontologia e/o presenti nella descrizione del sinistro.

Questo step che ha una grande valenza in quanto contribuisce in maniera significativa a migliorare la qualità complessiva del dataset, consente al modello, durante la fase di addestramento, di apprendere in modo più rapido e preciso grazie all'abbinamento delle informazioni testuali con le fotografie. Questa integrazione tra immagini e testo permette al modello di avere una comprensione più ricca e contestualizzata dei dati, conseguendo un apprendimento più efficace.

- Sulla destra del tool è stata scelta l'etichetta opportuna per procedere alla *segmentazione delle porzioni di immagini di interesse*. In particolare, le sole immagini appartenenti alle classi in contesto sono state etichettate in quanto al loro interno è possibile constatare la presenza del danno e quindi risultano funzionali all'addestramento dell'IA.

Nel danno da bagnamento considerato è stato possibile le seguenti etichette: *macchia chiara, macchia scura, muffa chiara, muffa scura, fughe bagnate, spargimento di liquidi, gocciolamento, scrostamento, rialzamento, rottura tubo*. Inoltre, è stato possibile etichettare l'*area di intervento* e anche i residui come *tubo, materiale elettrico, detriti e altro*.

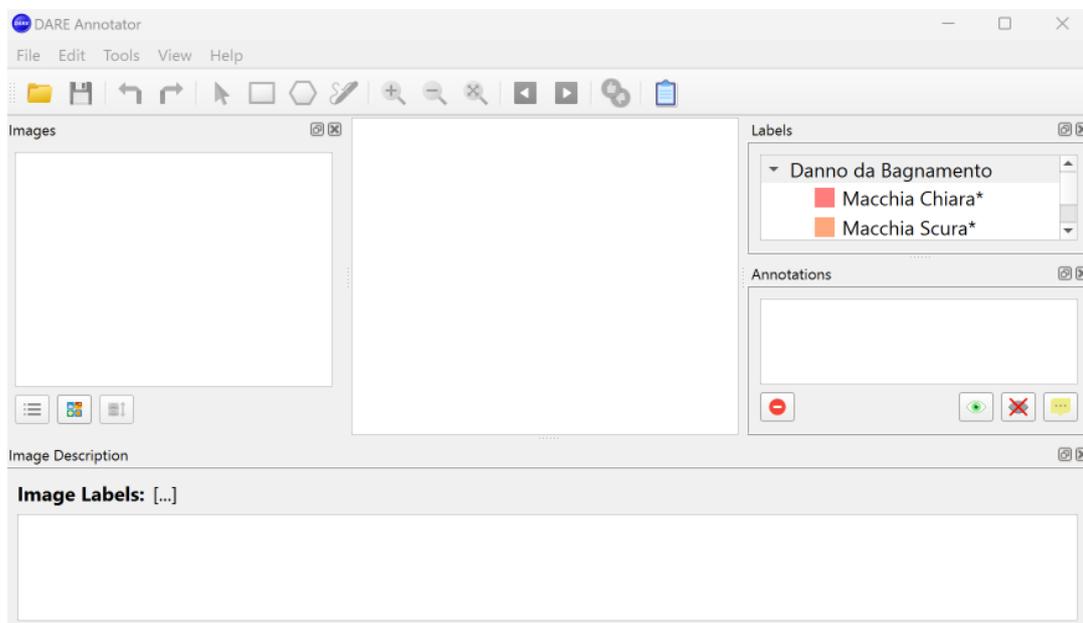


Figura 4.3 Schermata principale del tool di annotazione

Il processo di data cleaning, essendo iterativo, è stato via via ottimizzato grazie all'analisi delle immagini che ha permesso di individuare nuove classi, di aggiornare l'ontologia e, conseguentemente, il tool di annotazione. Ad esempio, nell'ultima versione del tool, le classi troppo ampie e generiche sono state ristrette per consentire una classificazione più precisa delle immagini e per migliorare l'efficacia dell'apprendimento del modello nel processo di classificazione: infatti, in precedenza, immagini raffiguranti gli elementi più disparati venivano classificate come random stuff, cosa che non consentiva al modello di distinguere con precisione determinati dettagli. Pertanto, è stato possibile definire in modo più rigoroso diversi aspetti relativi delle immagini, quali l'ambientazione, la presenza di persone, la qualità stessa delle immagini e l'evidenza di eventuali danni.

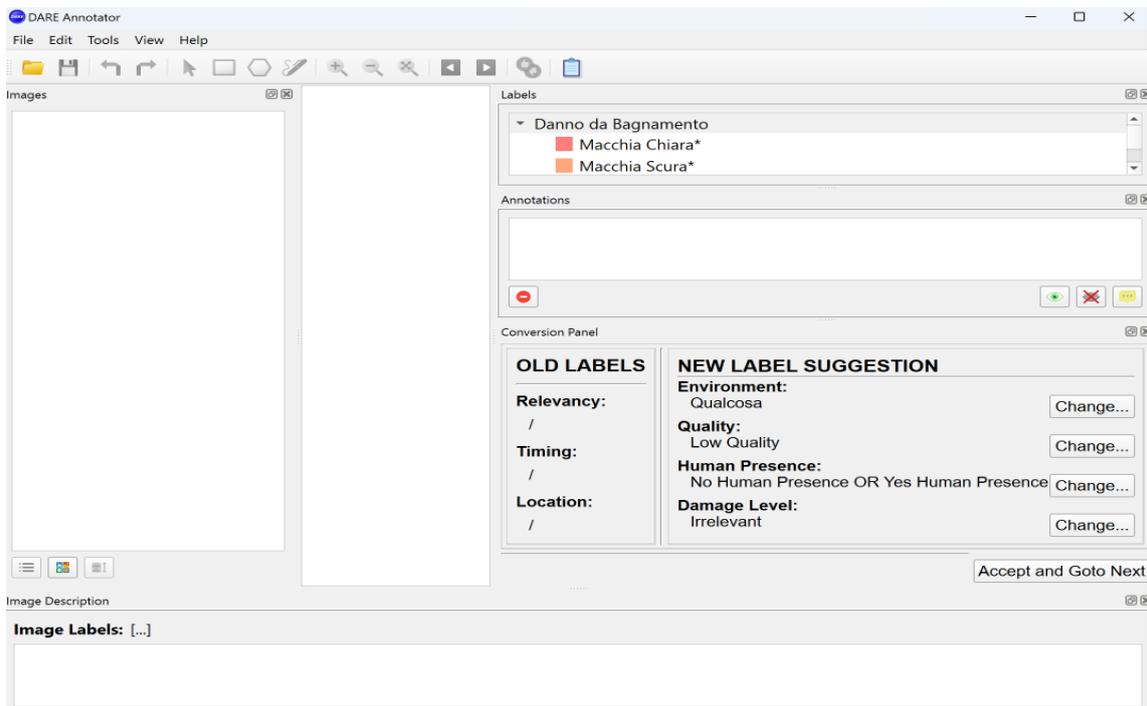


Figura 4.4 Schermata principale del tool con l'ultimo aggiornamento

Statistiche del campione analizzato

Le caratteristiche delle perizie analizzate sono state riportate all'interno di un file Excel consentendo di fare delle valutazioni per miglioramenti da apportare al tool e delle riflessioni sulla soggettività con cui le immagini sono state annotate. L'analisi del campione verrà riportata nel dettaglio nel capitolo successivo.

Capitolo 5

Analisi dei dati peritali

Nell'analisi del campione del dataset analizzato, costituito da un totale di 700 perizie sono state coinvolte quattro risorse, tre delle quali hanno analizzato ciascuna un campione di 200 perizie e una un campione di 100 perizie.

Nel presente capitolo vengono presentati i risultati ottenuti relativi ad un campione di 200 perizie, prima in modo aggregato, poi più in dettaglio distinguendo i risultati per i due network peritali al fine di evidenziarne similitudini e differenze.

Infine, vengono riportati anche i valori ottenuti dalle altre risorse evidenziando analogie e differenze nei risultati.

5.1 Risultati e statistiche del campione analizzato di 200 perizie (foto)

La struttura del campione su cui è stata svolta l'attività di analisi e annotazione si può schematizzare come segue: 103 perizie sono relative al primo network peritale, mentre le altre 97 sono relative al secondo.

Tabella 5.1 Risultati e statistiche del campione analizzato (200 perizie)

Risultati del campione	Statistiche del campione
Perizie totali analizzate = 200 Foto totali = 4172, media foto per perizia → circa 21 foto <ul style="list-style-type: none">• 2861 foto in contest, di cui:<ul style="list-style-type: none">○ 2534 Rilevanti○ 257 Blurred○ 28 Underexposed○ 1 Overexposed○ 0 Compression artifact• 103 foto ambigue• 1173 foto fuori contest, di cui:<ul style="list-style-type: none">○ 273 Building○ 297 Human interaction○ 23 Selfie○ 123 Screen capture○ 185 Documents○ 274 Random stuff• 1812 Pre intervention• 773 During intervention• 27 Post intervention• 1532 Unknown• 1835 Segmentate	<ul style="list-style-type: none">• 68,67% Foto in contesto, di cui<ul style="list-style-type: none">○ Rilevanti 88,45%○ Blurred 8,97%○ Underexposed 0,98%○ Overexposed <1%• 64% delle foto in contesto segmentate• 2,47% Foto ambigue• 28,12% Foto fuori contesto, di cui<ul style="list-style-type: none">○ 23,27% Building○ 25,32% Human interaction○ 1,96% Selfie○ 10,49% Screenshot○ 15,77% Documents○ 23,36% Random stuff• 43,43% Pre intervention• 18,53% During intervention• 0,65% Post-intervention• 36,72% Unknown

5.1.1 Analisi del campione analizzato di 200 perizie (foto)

L'analisi delle immagini relative alle perizie rivela che la *media di foto per perizia* include circa 21 fotografie, tuttavia, tale valore per perizia varia in modo considerevole: il caso con il maggior numero di immagini ne contiene 136, mentre quello con il minimo ne presenta solo 2. Questa significativa variabilità riflette, quindi, differenti livelli di dettaglio nella documentazione fotografica a corredo di ogni singola perizia.

Una parte consistente delle immagini è in contesto e, quindi, presumibilmente funzionale all'analisi del sinistro. Di queste, l'88,45% è considerato rilevante, mentre il restante 11% comprende immagini di qualità subottimale come foto sfocate, sottoesposte e sovraesposte. Questa suddivisione indica che, sebbene si disponga di buona quantità di foto utili, alcune immagini di bassa qualità (blurred, underexposed, overexposed e compression artifact), che non soddisfano gli standard di qualità ottimali necessari per una valutazione accurata dei danni, potrebbero essere rielaborate utilizzando dei programmi specifici per migliorarne la qualità visiva e rendere queste immagini idonee all'inclusione nel campione, ottimizzando così il dataset per l'addestramento del modello di Machine Learning.

Le *foto ambigue*, che risultano essere circa il 3%, sono immagini che non rientrano in una specifica classe ontologica in quanto o interessano parti non coinvolte nella polizza (es. danno alla tinteggiatura della facciata esterna di un palazzo condominiale a seguito della rottura di una tubazione interna collocata in un certo punto di un'abitazione) oppure si tratta di immagini per cui è stato poco chiaro definire la loro utilità.

Nonostante l'applicazione di un processo di anonimizzazione preliminare, in linea con l'approccio *privacy by design* del progetto (che ha previsto la rimozione di dati personali come nome e cognome, codice fiscale, Iban) il campione è stato ulteriormente "filtrato" per escludere foto fuori contesto.

Dalle 200 perizie analizzate, il 28,12% delle immagini è stato classificato come *fuori contesto*, includendo foto di edifici o parti di essi (es. facciate), interazioni umane, selfie, screenshot, documenti ed oggetti casuali. Questo genere di immagini che viene generalmente incluso nella documentazione fotografica per effettuare indagini antifrode, costituisce comunque una porzione rilevante del campione che non è pertinente all'analisi dei danni in quanto rappresenta elementi accessori.

Le foto sono inoltre classificate anche in base al momento di cattura rispetto all'intervento peritale: la maggior parte delle immagini documenta le fasi precedenti (43,43%) o nel corso dell'intervento (18,53%), mentre le immagini post-intervento risultano meno rappresentate (0,65%). Circa il 37% delle foto per le quali non è stato possibile identificare il momento esatto della cattura è stato classificato come unknown.

Infine, 1835 foto, pari al 64% delle foto in contesto, sono state sottoposte a un processo di segmentazione semantica che ha consentito, come precisato nel capitolo 4, di isolare specifiche aree delle immagini che ritraggono il danno al fine di aiutare il modello di ML a imparare in modo più mirato e veloce.

5.2 Confronto tra i due network peritali

Di seguito viene presentata, sulla base dei risultati e statistiche del campione analizzato, un'analisi approfondita dei due network peritali, con l'obiettivo di fornire una visione critica e comparativa dei diversi approcci peritali, mettendo in luce punti di forza e criticità di ciascun network.

5.2.1 Risultati analitici e statistiche del campione analizzato (foto)

Tabella 5.2 Risultati analitici: confronto tra i due network

Network 1	Network 2
Perizie totali analizzate = 103 Foto totali = 2274 /media foto per perizia → circa 22 foto <ul style="list-style-type: none"> • 1758 foto in contest, di cui <ul style="list-style-type: none"> ○ 1543 Rilevanti ○ 183 Blurred ○ 8 Underexposed ○ 1 Overexposed ○ 0 Compression artifact • 70 foto ambigue • 437 foto fuori contesto, di cui <ul style="list-style-type: none"> ○ 79 Building ○ 174 Human interaction ○ 8 Selfie ○ 63 Screen capture ○ 53 Documents ○ 62 Random stuff • 1082 Pre intervention • 486 During intervention • 14 Post intervention • 690 Unknown • 1120 Segmentate 	Perizie totali analizzate = 97 Foto totali = 1898, media foto per perizia → circa 20 foto <ul style="list-style-type: none"> • 1103 foto in contesto <ul style="list-style-type: none"> ○ 991 Rilevanti ○ 74 Blurred ○ 20 Underexposed ○ 1 Overexposed ○ 0 Compression artifact • 33 foto ambigue • 736 foto fuori contesto, di cui <ul style="list-style-type: none"> ○ 194 Building ○ 123 Human interaction ○ 15 Selfie ○ 60 Screen capture ○ 132 Documents ○ 212 Random stuff • 730 Pre intervention • 287 During intervention • 13 Post intervention • 842 Unknow • 715 Segmentate

Tabella 5.3 Statistiche analitiche: confronto tra i due network

Statistiche Network 1	Statistiche Network 2
<ul style="list-style-type: none"> • 77,31% Foto in contesto, di cui <ul style="list-style-type: none"> ○ Rilevanti 87,7% ○ Blurred 10,4% ○ Underexposed <1% ○ Overexposed ~0% • 63,7% delle foto in contesto segmentate • 3,08% Foto ambigue • 19,22% Foto fuori contest, di cui <ul style="list-style-type: none"> ○ 18,08% Building ○ 39,82% Human interaction ○ 1,83% Selfie ○ 14,42% Screencapture ○ 12,13% Documents ○ 14,15% Random stuff • 47,58% Pre intervention • 21,37% During intervention • 0,61% Post-intervention • 30,34% Unknown 	<ul style="list-style-type: none"> • 58,11% Foto in contesto, di cui <ul style="list-style-type: none"> ○ Rilevanti 89,84% ○ Blurred 6,7% ○ Underexposed 1,81% ○ Overexposed 0% • 64,82% delle foto in contesto segmentate • 1,74% Foto ambigue • 38,78% Foto fuori contest, di cui <ul style="list-style-type: none"> ○ 26,36% Building ○ 16,71% Human interaction ○ 2,04% Selfie ○ 8,15% Screencapture ○ 17,93% Documents ○ 28,80% Random stuff • 38,46% Pre intervention • 15,12% During intervention • 0,68% Post intervention • 44,36% Unknown

5.2.2 Analisi analitica del campione analizzato (foto): confronto tra i due network

L'analisi dei dati evidenzia che il *numero medio di foto per perizia* del network 1 è leggermente più alto nel primo network, con una media di 22 immagini, rispetto al secondo network che registra una media di 20 immagini per perizia. Tuttavia, il valore medio di fotografie per perizia non è particolarmente significativo poiché il campione esaminato presenta una notevole variabilità in considerazione di perizie contenenti meno di 5 immagini e altre con oltre 50 immagini.

Considerando il numero di *foto in contesto*, ovvero quelle pertinenti al soggetto o alla situazione rilevante per la perizia, il network 1 presenta il 77,31% delle foto totali classificate come in contesto, con una media quindi di 17 immagini di tale tipologia, mentre nel secondo network le foto in contesto sono il 58,11%, pertanto le perizie contengono una media di circa 11 foto in contesto; tra le foto in contesto, la percentuale di immagini rilevanti è pari all'87,7% per il primo network e all'89,8% per il secondo network, quindi valori in termini percentuali molto simili.

Le *immagini segmentate*, fondamentali per l'addestramento di modelli di ML, sono una porzione delle immagini in contesto. In particolare, per il primo network ne sono state etichettate 1120 su un totale di 1758 delle foto in contesto, ossia circa il 64%, mentre per il secondo sono state etichettate 715 sulle totali 1103 in contesto, che corrispondono a circa il 65%.

Analizzando i dati in termini medi, per il primo network, il numero di immagini in contesto come detto, è di 17 immagini, di cui circa 11 risultano etichettate; nel secondo network, invece, il numero medio di foto in contesto è pari a 11 e quindi circa 7 immagini sono state etichettate.

Le foto in contesto, ma *blurred*, *underexposed* e *overexposed* rappresentano una porzione limitata del dataset, infatti, nel caso del primo network peritale sono 192, pari al 10,92% delle 1758 foto in contesto totali e all'8% circa rispetto alle foto totali: mediamente per ogni perizia è presente una foto delle tipologie appena citate. Analizzando in modo più specifico le perizie contenenti tale tipologia di foto in contesto, esse sono presenti in 34 perizie a fronte delle 103 analizzate. Per il secondo network, il totale di foto in contesto blurred, underexposed, overexposed e compression artifact, seppur la differenza di perizie analizzate tra primo e secondo network, è di 94, rappresentando l'8,5% delle 1103 foto totali in contesto e il 4,9% circa rispetto alle foto totali. Anche in questo caso analizzando in modo più specifico le perizie contenenti tale tipologia di foto in contesto, esse sono presenti in 33 perizie a fronte delle 97 analizzate.

Per quanto riguarda le *foto ambigue*, per il primo network risultano essere pari a 70 contenute in totale in 23 perizie su 103, mentre per il secondo network, le immagini ambigue risultano essere pari a 33 relative ad un totale di 13 perizie su 97.

Le foto fuori contesto sono il 19,34%, per il primo network e 38,78 % per il secondo.

Sebbene la numerosità delle perizie analizzate presenti una lieve differenza tra i due network (103 perizie analizzate e annotate per il primo, rispetto alle 97 del secondo) il network 1 appare maggiormente orientato verso un focus diretto e specifico sull'intervento oggetto della perizia, privilegiando la raccolta di immagini pertinenti e rilevanti per la valutazione puntuale del caso, mentre il network 2 mostra una propensione più marcata verso la documentazione del contesto circostante.

5.2.3 Osservazioni analitiche sul campione analizzato (testo perizia con importo liquidazione)

Per quanto riguarda il *network 1* sul totale di 103 perizie analizzate risultano 54 perizie liquidate, 48 non liquidate e una perizia ambigua per la quale non è chiaro se vi sia stato un importo effettivamente liquidato o meno.

A seguito all'analisi dei report peritali, si riportano le principali cause di mancata liquidazione individuate con la relativa frequenza.

Tabella 5.4 Cause di mancata liquidazione per il network 1 e relativa frequenza

Causa di mancata liquidazione	Frequenza
Mancata individuazione del danno (con lavori svolti)	18
Non sono stati svolti i lavori di accertamento del danno	12
Il danno interessava parti esterne alla parte assicurata	2
L'assicurato non accetta la cifra proposta	1
Mancata trasmissione della documentazione fotografica	1
Danni dovuti ad altra causa che non rientra nell'assicurazione	8
Non sono stati conservati gli elementi residui del sinistro e nessuna documentazione fotografica è stata fornita a supporto delle riparazioni eseguite. Non è pertanto possibile alcun riscontro sulle cause del danno.	5
La spesa per il ripristino è stata esigua	1
TOTALE	48

Tabella 5.5 Cause di ambiguità nell'analisi della liquidazione delle perizie per il network 1 e relativa frequenza

Causa ambiguità	Frequenza
All'interno del file Excel non è riportata nessuna descrizione sul sinistro	1

L'*importo medio liquidato* risulta essere di circa 1100 euro, tuttavia, la varianza tra gli importi è significativa: in particolare l'importo massimo liquidato è risultato essere pari a 10000 euro, mentre l'importo minimo liquidato di 75 euro.

Per quanto riguarda il *network 2*, risultano sul totale di 97 perizie analizzate, 47 perizie liquidate, 39 non liquidate e 8 perizie ambigue per le quali non è possibile rilevare effettivamente se vi sia stato un importo liquidato o meno. In sole 3 perizie, per via della mancanza del file Excel, non è stata possibile stabilire se una perizia sia stata liquidata o meno.

Tabella 5.6 Cause di mancata liquidazione per il network 2 e relativa frequenza

Causa	Frequenza	
L'ammontare del danno non è stato concordato	Non è stata accertata la causa del sinistro	2
	L'Assicurato non accetta le riserve elevate	3
	Non si è in possesso dell'incarico con la posizione RC proprietà	1
	Garanzia di risarcimento da parte della Filiale	2
	Gli interventi di ricerca inizieranno a breve	1

	L'ammontare non ancora concordato con la Controparte	1
	Si è in attesa dell'intervento di altre coperture assicurative	1
Il danno risulta essere coperto interamente dalla franchigia		7
Le risultanze della perizia sono state comunicate all'Assicurato, al quale è stato inviato un atto che, ad oggi, non è pervenuto sottoscritto		9
Al momento della videoperizia non risulta ancora effettuata la ricerca del guasto e pertanto non si dispone degli elementi necessari per stimare l'indennizzo spettante		1
L'Assicurato non è in grado di fornire tempi certi di attuazione dell'intervento di ricerca del guasto		3
Accettazione della riserva		2
Danno dovuto ad altra causa, quindi si eleva riserva		2
Non sussiste operatività di alcuna garanzia di polizza		3
L'assicurato non accetta quanto proposto in quanto l'ammontare della franchigia contrattuale non gli era stata comunicata in sede di stipula della polizza		1
TOTALE		39

Dall'analisi effettuata è emerso, spesso tra i dati presenti nei file Excel emergano delle ambiguità che ne compromettono l'effettiva comprensione circa l'avvenuta liquidazione o meno del danno.

In particolare, le perizie ambigue risultano essere 8 rispetto alle quali si riportano di seguito le cause di ambiguità con la relativa frequenza.

Tabella 5.7 Cause di ambiguità nell'analisi della liquidazione delle perizie per il network 2 e relativa frequenza

Cause ambiguità	Frequenza
All'interno di alcuni fogli presenti nel file Excel viene riportato l'importo di liquidazione; tuttavia, nonostante il danno sia stato concordato telefonicamente con l'assicurato, l'atto inviato non è stato sottoscritto.	4
Il danno non è stato concordato (es. l'assicurato non si è reso disponibile nell'effettuare videochiamata) tuttavia viene indicato un importo liquidato	2
Altra causa: la riserva sollevata sull'indennizzabilità del danno è stata comunicata all'assicurato. L'assicurato ha replicato di non essere d'accordo con la riserva in quanto il tratto di tubazione oggetto del sinistro è comunque di pertinenza e di proprietà dell'abitazione assicurata.; tuttavia emerge un importo liquidato)	2

Infine, per sole 3 perizie mancanti dei file Excel, non è stato possibile riportare né la descrizione all'interno del tool per supportare la segmentazione delle immagini in modo più dettagliato, né capire l'effettiva liquidazione del danno.

Per quanto riguarda l'importo medio totale liquidato, questo risulta essere di circa 750 euro, ma si rileva una varianza tra gli importi significativa: in particolare, l'importo massimo liquidato risulta essere pari a 2660 euro, mentre l'importo minimo liquidato pari a 77 euro.

5.3 Risultati e statistiche del dataset complessivo di 700 perizie (foto)

Tabella 5.8 Risultati a confronto delle quattro risorse

	Risorsa 1 (100 perizie)	Risorsa 2 (200 perizie)	Risorsa 3 (200 perizie)	Risorsa 4 (200 perizie)	TOTALE	
Foto totali	2247	5460	4523	4172	16402	
Foto in contesto	1319	2644	3348	2865	10176	
-Rilevanti	906	1662	3191	2534	8293	
-Blurred	329	687	111	257	1384	
-Underexposed	43	286	31	28	388	
-Overexposed	14	29	-	1	44	
-Compression artifact	27	-	-	-	27	
Foto ambigue	89	304	29	103	525	
Foto fuori contesto	835	2488	1146	1173	5642	
-Building	75	337	279	273	1052	
-Human-interaction	75	514	301	297	1187	
-Selfie	40	18	7	23	88	
-Screencapture	43	452	139	123	757	
-Documents	20	125	113	185	453	
-Random stuff	484	1046	308	274	2112	
Timing	-Pre intervention	643	1424	1315	1835	5217
	-During intervention	421	945	651	1812	3829
	-Post intervention	91	165	94	773	1123
	-Unknown	1018	2926	2371	27	6342
Foto etichettate	674	1936	2399	1532	6541	

Tabella 5.9 Statistiche a confronto delle quattro risorse

	Risorsa 1	Risorsa 2	Risorsa 3	Risorsa 4	TOTALE	
Foto medie per perizia	22,47	27,3	22,615	20,86	23,31	
%Foto in contesto rispetto alle foto totali	58,70%	48,42%	74,02%	68,67%	62,04%	
- %Foto rilevanti rispetto alle foto in contesto	68,70%	62,86%	95,31%	88,45%	81,50%	
- %Foto blurred rispetto alle foto in contesto	24,94%	25,98%	3,32%	8,97%	13,60%	
- %Foto underexposed rispetto alle foto in contesto	3,26%	10,82%	0,93%	0,98%	3,81%	
- %Foto overexposed rispetto alle foto in contesto	1,06%	1,10%	-	0,03%	0,43%	
-%Foto compression artifact rispetto alle foto in contesto	2,05%	-	-	-	0,27%	
%Foto ambigue rispetto alle foto totali	3,96%	5,57%	0,64%	2,47%	3,20%	
%Foto fuori contesto rispetto alle foto totali	37,16%	45,57%	25,34%	28,12%	34,40%	
-%Foto Building rispetto alle foto fuori contesto	16,71%	13,55%	25,35%	23,27%	18,65%	
-%Foto Human-Interaction rispetto alle foto fuori contesto	9,65%	20,66%	26,27%	25,32%	21,04%	
-%Foto Selfie rispetto alle foto fuori contesto	1,88%	0,72%	0,61%	1,96%	1,56%	
-%Foto Screencapture rispetto alle foto fuori contesto	7,76%	18,17%	12,13%	10,49%	13,42%	
-%Foto Documents rispetto alle foto fuori contesto	4,94%	5,02%	9,86	15,77%	8,03%	
-%Foto Random stuff rispetto alle foto fuori contesto	59,96%	42,04%	26,88%	23,36%	37,43%	
Timing	%Foto Pre intervention rispetto alle foto totali	28,62%	26,08%	29,07%	43,43%	31,67%
	%Foto During intervention rispetto alle foto totali	18,74%	17,31%	14,39%	18,53%	17,01%
	%Foto Post intervention rispetto alle foto totali	4,05%	3,02%	2,08%	0,65%	2,30%
	%Foto Unknown rispetto alle foto totali	45,30%	53,59%	52,42%	36,72%	47,84%
%Foto segmentate rispetto alle foto totali	30,00%	35,46%	53,04%	43,98%	41,73%	
%Foto segmentate rispetto alle foto in contesto	51,10%	73,22%	71,65	64,05%	67,26%	

5.3.1 Analisi del dataset complessivo di 700 perizie: confronto tra le risorse

Le considerazioni presentate di seguito sono state elaborate isolando la *casualità* con cui le perizie, e quindi le immagini, sono state assegnate alle diverse risorse. È stata dunque effettuata un'analisi della *soggettività* con cui le risorse, pur disponendo di linee guida, hanno classificato e annotato le immagini presenti nella documentazione fotografica delle perizie. Si precisa che, dato che alla Risorsa 1 è stato assegnato un campione composto da 100 perizie, quindi di dimensioni diverse rispetto ai campioni delle altre risorse, esso, proprio perché più piccolo, tende ad essere meno rappresentativo della popolazione complessiva e a presentare maggiore variabilità nei risultati (es. la presenza di outlier può avere un impatto maggiore).

Per quanto riguarda le *foto in contesto*, la Risorsa 3 presenta la percentuale più alta di foto in contesto (74,02%), seguita dalla Risorsa 4 (68,67%) e dalla Risorsa 1 (58,70%), mentre la Risorsa 2 presenta il valore più basso (48,42%) che suggerisce un probabile approccio più restrittivo nella definizione di foto in contesto.

Costituendo le *foto segmentate* una porzione delle foto in contesto, quello che emerge è che la Risorsa 3 ha la più alta percentuale di segmentazione delle foto, seguita dalla Risorsa 4: ciò indica probabilmente una maggiore attenzione alla segmentazione delle immagini.

Per quanto riguarda le *foto rilevanti* la Risorsa 3 classifica il 95,31% delle foto in contesto come rilevanti e ciò suggerisce un approccio inclusivo rispetto alle altre risorse, specialmente rispetto alla Risorsa 2 che presenta, invece, la percentuale più bassa di foto rilevanti (62,86%), evidenziando quindi una classificazione più severa.

Le Risorse 1 e 2, seguite dalla risorsa 4, segnalano la maggior percentuale di *foto sfocate o sottoesposte*, rilevando una maggiore attenzione ai difetti di qualità, mentre la Risorsa 3 presenta i valori più bassi per le immagini con difetti (3,32%), suggerendo l'attribuzione di una rilevanza minore a questi aspetti.

Circa le *foto fuori contesto* la Risorsa 2 registra la percentuale più alta di foto fuori contesto (45,57%), mentre la Risorsa 3 la più bassa (25,34%), differenza che potrebbe riflettere, così come anche riportato sopra, criteri diversi nella classificazione tra contesto e fuori contesto.

Le foto fuori contesto tra la Risorsa 3 e la Risorsa 4 risultano abbastanza allineate, mentre la Risorsa 1 presenta il valore più alto di foto di tipo Random stuff (59,96%), seguita dalla Risorsa 2 (42,04%).

Infine, con riferimento al timing delle foto la Risorsa 4 mostra una forte presenza di foto pre intervention (43,43%), mentre tutte le altre risorse condividono una percentuale significativa di foto classificate come unknown, segnalando, quindi, una certa incertezza nel timing.

Come ci si attendeva, la soggettività induce una forte variabilità delle rilevazioni come mostrato dall'approccio e dai criteri di valutazione applicati dalle risorse: la Risorsa 1 classifica in modo più equilibrato, evidenziando un maggiore focus sulla qualità dell'immagine e una segmentazione meno frequente; la Risorse 2 adotta un approccio molto stringente specialmente per quanto riguarda le foto in contesto ed è altresì attenta alla qualità dell'immagine; la Risorsa 3 risulta più inclusiva nel considerare un maggior numero di foto in contesto e rilevanti e probabilmente per questo motivo presenta la più alta percentuale di segmentazione (53,04%); infine la Risorsa 4 è anch'essa molto inclusiva nel considerare foto in contesto e rilevanti e presenta una particolare attenzione per le immagini pre intervention.

Capitolo 6

Discussione dei risultati peritali

Partecipare attivamente al caso studio ha offerto un'opportunità diretta per comprendere l'importanza e l'influenza del processo di *data cleaning* sull'apprendimento del modello di classificazione. I continui affinamenti dell'ontologia (descritta al capitolo 4.4), che hanno consentito di mitigare progressivamente l'incertezza interpretativa nella classificazione dei danni, e dei successivi aggiornamenti del tool di annotazione hanno permesso di "pulire" i dati in modo sempre più rigoroso e con un impatto significativo sulle prestazioni del modello.

Infatti, è emersa, grazie ad una sempre più precisa analisi di nuovi dati, una variazione nella metrica F1 score (una delle metriche più utilizzate per la valutazione dei modelli di classificazione). Come detto nel capitolo 3.6 l'F1 Score è una media armonica tra altre due metriche: precisione e recall. Quindi, combinando la capacità di identificare correttamente le immagini positive, cioè la precisione, con quella di individuare tutti le immagini effettivamente positive, cioè il recall, una sua variazione indica un miglioramento dell'affidabilità del modello nel classificare correttamente i dati.

L'ampliamento di specifiche classi, realizzato attraverso un'attenta analisi delle matrici di confusione², ha permesso di identificare i punti deboli del modello e di introdurre, come detto, miglioramenti nell'ontologia e nel tool di annotazione, grazie ad un approccio che, unito ad ulteriori immagini analizzate ed annotate, ha portato ad un aumento dell'efficacia del modello nella classificazione delle immagini, come evidenziato in figura 6.1, e quindi del passaggio dell'F1 Score da un valore iniziale di 0,729 a un valore successivo di 0,842.

Pertanto, attraverso queste continue azioni migliorative il modello si è avvicinato, e continuerà ad avvicinarsi sempre più, ad un F1 score prossimo all'unità, dimostrando una capacità sempre più elevata nel classificare le immagini.



Figura 6.1 Evoluzione dell'F1 Score

² La *matrice di confusione* è uno strumento utile per visualizzare l'accuratezza della classificazione dei modelli visualizzando il numero di previsioni corrette e previsioni errate per tutte le classi una accanto all'altra.

Inoltre, è possibile formulare delle considerazioni sulle modalità dell'analisi condotta sul campione di 200 perizie, così come sull'intero dataset perizie, che, oltre all'inevitabile e fisiologica oggettività con cui le stesse sono state analizzate ed annotate, va evidenziata l'assenza di *procedure operative standardizzate* che regolino in modo omogeneo la documentazione visiva. Attualmente, nella pratica comune, la qualità delle immagini e la coerenza generale delle perizie variano a seconda del network peritale di riferimento per l'ente assicuratore, della competenza e dell'attenzione con cui ciascun perito o operatore svolge il proprio lavoro. Questa disomogeneità comporta incertezza interpretativa nell'analisi dei dati raccolti e incide negativamente sulla coerenza e sulla precisione del dataset per training del modello di ML.

Per ovviare a tale criticità, il progetto prevede un approccio strutturato che mira, parallelamente alla *standardizzazione dei dati*, ossia del soggetto dell'immagine attraverso la formalizzazione di un'ontologia (che definisce una rappresentazione uniforme e condivisa degli elementi rilevanti) a regolare rigorosamente le procedure operative per la realizzazione di immagini.

Non esistono, ad esempio, procedure univoche che obblighino chi scatta le immagini a produrne un numero standard e utile per la corretta visualizzazione del danno, così come mancano, altresì, direttive specifiche che impongano l'adozione di parametri tecnici minimi come una corretta esposizione e risoluzione delle immagini, un'adeguata inquadratura o una certa distanza dal danno, elementi che sarebbero necessari per garantire una documentazione chiara e dettagliata.

Proprio con riferimento all'assenza di linee guida specifiche sul numero standard di fotografie da acquisire per documentare i danni è emerso, ad esempio, che all'aumentare della complessità del danno (e quindi potenzialmente dell'importo liquidato) chi è incaricato di effettuare la documentazione fotografica incontra difficoltà nel selezionare accuratamente ciò che deve essere ripreso, tendendo così a produrre un numero di immagini che risulta eccessivo (*data overload*) per le finalità peritali. In conseguenza di tale osservazione, il dataset delle 700 perizie è stato analizzato isolando solo quelle liquidate, correlando l'importo liquidato (tendenzialmente legato alla complessità del danno) al numero di fotografie presenti nella documentazione.

Dall'analisi grafica dei dati, rappresentata nel grafico sottostante (importo liquidato in funzione del numero di foto in perizia) emerge una chiara variabilità.

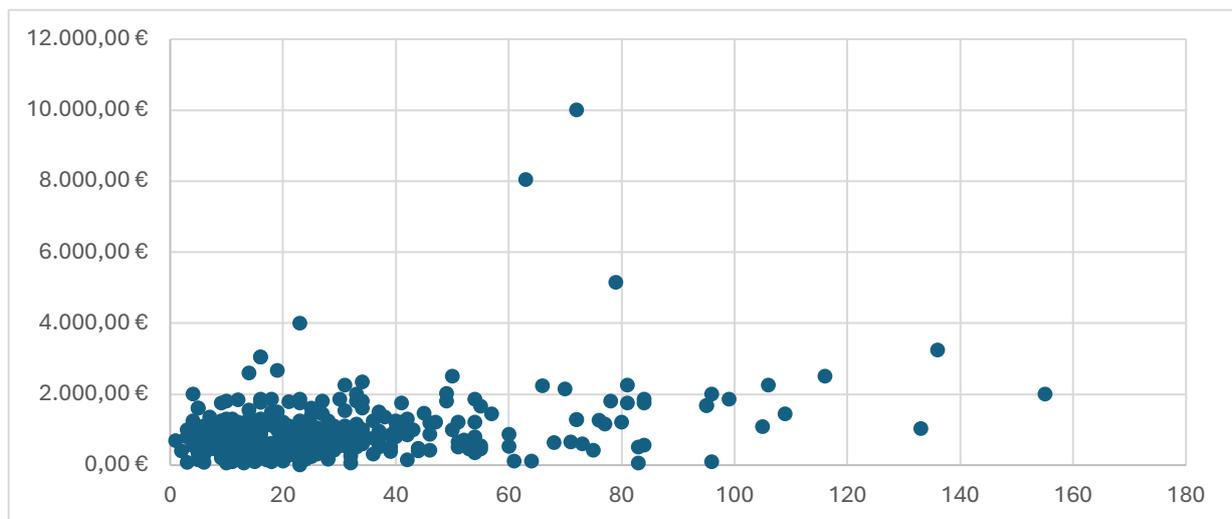


Figura 6.2 Importo liquidato in funzione del numero di foto presenti in perizia

Si nota infatti una significativa concentrazione di immagini nella parte inferiore sinistra del grafico, che corrisponde ai danni di minore entità (con importi inferiori a 1000 euro), o di pronta liquidazione, dove la semplicità della situazione consente di limitare il numero di fotografie. Parallelamente, si evidenzia una tendenza crescente che suggerisce, come all'aumentare dell'importo liquidato cioè nei casi più complessi, la documentazione fotografica sia più estesa, con un numero di immagini spesso superiore a quello effettivamente necessario per la valutazione (data overload).

Quindi, al fine di controllare il data overload (molte foto e poco chiare possono determinare incertezza interpretativa nel classificare i danni stessi) e migliorare l'apprendimento del modello di IA, sono state definite delle nuove procedure per l'acquisizione, l'elaborazione e l'annotazione dei dati *ex-ante* che riguardano principalmente tre aspetti: la *qualità dell'immagine*, la *selezione dei soggetti di interesse per l'annotazione* e l'*eliminazione delle foto non utili*.

Per quanto riguarda la qualità dell'immagine, è necessario che vi siano dei requisiti minimi che le immagini dovranno rispettare in fase di realizzazione: al fine di conservare i dettagli e garantire chiarezza del danno, bisognerà assicurarsi di aver scattato foto con risoluzione massima e assenza di compressione, sfocature, sovraesposizioni o sottoesposizioni, prive di segni o simboli sovrapposti, come scritte o frecce, che potrebbero interferire con il soggetto da annotare.

Per documentare l'evento dannoso bisognerà disporre di determinate foto per selezionare uno o più soggetti di interesse da annotare, per cui, in particolare, saranno necessari almeno una foto per apprezzare gli effetti visibili del danno (ad una distanza tale da permettere sia una visione ottimale del danno, sia l'area che secondo l'esperienza del perito dovrà essere ripristinata), almeno una foto relativa alla causa del danno (es. rottura tubo di scarico), almeno una foto per ogni elemento danneggiato (es. porte, infissi, ecc.), almeno una foto dei lavori di ricerca, oppure durante l'intervento, almeno una foto di residuo, in caso di tubatura danneggiata, e almeno una foto dopo il ripristino cioè post-intervento.

Occorrerà, infine, procedere con l'eliminazione delle foto non utili al processo di annotazione, pertanto, non dovranno essere incluse tutte quelle immagini non rappresentative del danno, come immagini che ritraggono l'esterno dell'edificio, che non mostrano chiaramente i danni o che contengono persone o animali, compromettenti la privacy e immagini contenenti segnali stradali, documenti o testi.

In definitiva, con una classificazione ben strutturata delle immagini *ex ante* ci si aspetta di conseguire benefici non indifferenti come una diminuzione significativa del data overload, una riduzione delle tempistiche di data preparation per il training degli algoritmi IA, un miglioramento dell'accuratezza dei modelli, una maggiore uniformità delle decisioni con conseguente riduzione della discrezionalità nell'importo liquidato.

Capitolo 7

Esperimento per valutazione comparativa

7.1 Introduzione

Nell'ambito delle attività svolte nel corso del progetto D.A.R.E., è emersa la necessità di indagare, se e in quale misura strumenti basati sull'IA possano effettivamente supportare e migliorare il lavoro umano, rafforzando in particolare la qualità dei processi decisionali.

Per conferire rigore alla presente ricerca è stata condotta una *literature review* finalizzata a comprendere lo stato dell'arte riguardo alle questioni precedentemente delineate. Tale revisione è stata effettuata mediante una query specifica su Scopus con l'impiego delle parole chiave *Artificial Intelligence* e *Experiment* e l'applicazione di un filtro mirato ad includere articoli di comprovata autorevolezza provenienti da riviste quali *Strategic Management Journal*, *Information Systems Research*, *Organization Studies*, *Management Science* e *Organization Science*. La selezione di queste fonti è motivata dal loro focus su argomenti afferenti ai settori economico e manageriale.

In un primo momento gli articoli considerati di interesse sono stati pari a 65 tuttavia dopo un'analisi accurata che ha portato ad individuare gli studi con impianto sperimentale, il corpus finale è stato ridotto ad un numero di 8 articoli.

In letteratura emerge che tesi per cui la sinergia tra IA e gli individui genera nuove competenze di valore, diverse (Krakosky et al., 2022) superano le possibilità dei singoli elementi, offrendo prospettive rilevanti per il miglioramento dell'accuratezza analitica e della precisione decisionale (*metafora dei centauri*). Dell'Acqua et al. (2024) approfondiscono l'interazione uomo-IA in un ambiente di consulenza strategica rilevando che l'adozione di IA, come GPT-4 non solo potenzia la produttività ma riduce la variabilità delle prestazioni tra consulenti con diversi livelli di esperienza. Da questa interazione, tuttavia, emerge un calo della diversità delle idee generate, sollevando questioni su come bilanciare standardizzazione e creatività nell'uso della tecnologia. Parallelamente, in un contesto di customer service, Wang et al. (2023) hanno investigato l'introduzione di un sistema vocale IA in un call center dimostrando l'efficacia dell'IA in contesti semplici, ma limitata in contesti complessi ove sono richieste competenze umane avanzate.

Nonostante, quindi, la letteratura scientifica evidenzi le potenzialità dell'IA, emerge, tuttavia, una limitata disponibilità di studi empirici caratterizzati, per lo più, da approcci concettuali semplificati, riflettendo sia la novità della tecnologia sia le sfide metodologiche connesse alla dimostrazione dei benefici concreti dell'IA in contesti complessi con molteplici variabili e alti livelli di incertezza.

A tal proposito, nella classificazione teorica proposta da Perrow, l'IA può essere particolarmente efficace come complemento al lavoro umano in quei compiti che combinano alta

incertezza, cioè bassa analizzabilità, e bassa complessità, quest'ultima intesa come deviazione minima rispetto alle normali procedure.

Dunque, il fulcro di questa ricerca è quello di approfondire ulteriormente la sinergia tra IA e competenze umane in un settore contraddistinto da situazioni variabili, non standardizzabili e non ripetitive, come quello assicurativo, dove sono richieste decisioni rapide, precise e basate su valutazioni approfondite.

All'interno di questo settore, l'attenzione è rivolta, in particolare, alla *gestione dei sinistri*, un ambito contraddistinto da processi analitici complessi e da una forte dipendenza dalla qualità delle decisioni: per ogni sinistro, infatti, non occorre semplicemente verificare la presenza di un danno, ma bisogna identificarne e valutarne accuratamente le cause, distinguendo quelle ammissibili ai fini del rimborso da quelle escluse, determinare l'estensione e la gravità dei danni per formulare una stima precisa ed equa dell'importo risarcibile.

Attualmente, questo processo di valutazione, intrinsecamente complesso, viene eseguito con metodologie che consentono un ampio margine interpretativo agli operatori, sollevando questioni rilevanti riguardo all'introduzione dell'IA.

La ricerca intende dunque esaminare se l'adozione di strumenti basati sull'IA possa contribuire a ridurre la soggettività e la discrezionalità delle valutazioni, riducendo i margini interpretativi e favorendo maggiore coerenza e standardizzazione nei giudizi.

Questo studio, quindi, è finalizzato a colmare un gap nella letteratura proponendo un *esperimento di valutazione comparativa* volto ad indagare l'influenza dell'IA nel processo decisionale umano in un ambiente incerto e con molte variabili, concentrandosi in particolare sulla capacità di tale tecnologia di facilitare la selezione e l'analisi delle informazioni più rilevanti per la valutazione dei sinistri. Inoltre, si vuole comprovare il ruolo dell'ontologia come strumento efficace per mitigare l'incertezza interpretativa e la soggettività, due variabili particolarmente rilevanti nell'analisi della documentazione fotografica delle perizie che possono generare disomogeneità e discrepanze nei risultati finali. L'ontologia, infatti, definita come un insieme strutturato di relazioni e concetti specifici di un dominio, si presenta come uno strumento che può guidare l'interpretazione dei dati visivi, consentendo uniformità e oggettività nelle decisioni finali.

Un ulteriore aspetto di rilievo in questa ricerca riguarda l'analisi delle capacità dell'IA di ridurre il divario di performance tra utenti con diversi livelli di esperienza.

Infatti, in molti contesti decisionali complessi le competenze giocano un ruolo determinante nei risultati finali e quindi attraverso l'uso combinato di IA e strumenti basati sull'ontologia, questa ricerca accerterà se sia possibile livellare tali differenze, rendendo l'efficacia delle decisioni meno dipendente dall'esperienza e più aderente a criteri oggettivi e strutturati.

Infine, lo studio intende approfondire l'influenza della spiegabilità dei suggerimenti forniti dall'IA (explainable AI) con l'obiettivo di comprendere come la comprensibilità degli output della macchina incidano sull'efficacia del processo di integrazione della tecnologia IA all'interno dei tradizionali processi decisionali. Infatti, una comprensione approfondita dell'impatto della spiegabilità consentirebbe di determinare se la naturale diffidenza degli individui possa evolversi in un rapporto di fiducia verso la tecnologia, agevolando così una collaborazione informata e consapevole.

Tabella 7.1 Riepilogo dei Gap, delle Research questions e degli obiettivi dell'esperimento

Gap	<ul style="list-style-type: none"> • Esistono pochi studi sperimentali relativi ad un contesto aziendale reale che si occupano di analizzare l'interazione tra uomo ed IA • Esistono pochi studi sperimentali caratterizzati da elevata incertezza e quindi elevata variabilità
Research questions	<ul style="list-style-type: none"> • Come influisce l'IA nel processo decisionale per la valutazione dei sinistri? <ul style="list-style-type: none"> • L'IA permette al decisore di selezionare e, quindi, analizzare solo le informazioni più rilevanti per la valutazione dei sinistri? • L'IA può avvicinare le performance di un utente non esperto a quelle di un utente esperto? • Di quanto l'IA può migliorare le performance di un utente esperto? • L'ontologia può mitigare l'incertezza interpretativa e la soggettività nelle valutazioni? • Qual è l'influenza della spiegabilità nell'efficacia del processo di integrazione?
Obiettivi	<p>Realizzare un <i>esperimento di valutazione comparativa</i> finalizzato a:</p> <ul style="list-style-type: none"> • Indagare l'influenza dell'IA nel processo decisionale umano sia negli utenti esperti sia non esperti nel processo di valutazione dei sinistri • Comprovare il ruolo dell'ontologia di dominio

7.2 Risorse e strumenti operativi

Per garantire una comprensione completa e approfondita dell'esperimento condotto, risulta fondamentale fornire una panoramica sia degli strumenti utilizzati, che dei partecipanti coinvolti, questi ultimi classificabili come esecutivi e non esecutivi.

Lo studio ha fatto uso di un campione di 500 perizie fornite dai network peritali con cui la compagnia assicurativa partner del progetto collabora. Tali perizie sono state selezionate casualmente da un insieme più ampio di documenti, previamente filtrato per garantire l'assenza di criticità interpretative che potessero inficiare l'accuratezza dei risultati.

Nel corso dell'esperimento, per assistere l'attività umana e valutarne la loro efficacia operativa in contesti decisionali reali sono stati utilizzati due strumenti di supporto opportunamente progettati. In particolare, nella fase iniziale dello studio è stato impiegato un *tool dell'ontologia*, sviluppato specificatamente all'interno del progetto D.A.R.E., che consente di classificare in modo dettagliato i danni visibili nelle immagini contenute nella documentazione fotografica peritale. Questo strumento consente di assegnare etichette alle immagini e di isolare le aree più rappresentative del danno tramite l'uso dell'applicazione di maschere semantiche, fornendo così un supporto strutturato per l'analisi delle immagini.

Nella fase successiva, l'esperimento ha previsto l'utilizzo di un secondo strumento, un *tool basato su IA*, che automatizza molte delle funzioni svolte manualmente attraverso lo strumento di ontologia, classificando i danni e applicando maschere semantiche alle aree di interesse immediatamente all'apertura della cartella contenente la documentazione fotografica della perizia. Inoltre, il tool fornisce una probabilità della copertura assicurativa del danno e una stima dell'importo liquidabile. Ad esempio, qualora venga rilevato un danno ad una parete, il tool suggerisce che tale danno, con una probabilità del 90%, è coperto dalla garanzia e stima un importo approssimativo da liquidare. Quanto descritto, tuttavia, richiede una successiva valutazione e approvazione finale da parte del decisore umano. Un elemento chiave di questo strumento è la funzione di spiegabilità, la quale genera una spiegazione testuale del processo

decisionale dell'IA, dettagliando i passaggi che hanno portato alla formulazione delle conclusioni presentate.

Per quanto riguarda i *partecipanti esecutivi* all'esperimento, sono stati coinvolti sei studenti tesisti, considerati utenti *principianti* e sei periti, considerati utenti *esperti*.

I principianti, pur non disponendo di un'esperienza paragonabile a quella dei professionisti esperti, conoscono limitatamente, grazie al lavoro di tesi, il dominio assicurativo e peritale, l'ontologia e il tool di classificazione. È opportuno precisare che la figura del principiante è stata selezionata in quanto la finalità dell'esperimento è proprio quella di riprodurre un contesto il più possibile aderente alla realtà operativa: infatti, proprio un utente con competenze limitate, ma con una conoscenza di base dell'ontologia, può trarre maggiore beneficio da strumenti strutturati come il tool di classificazione o il tool IA. Al contrario, coinvolgere soggetti totalmente inesperti non sarebbe stato realistico in un dominio di conoscenza così specifico come quello assicurativo in quanto risulterebbe inverosimile che un individuo privo di conoscenze tecniche possa ricoprire un ruolo decisionale, anche se supportato da tool avanzati.

Per quanto riguarda gli utenti esperti, sono stati coinvolti sei periti professionisti, messi a disposizione dai network peritali affiliati alla compagnia assicurativa partner del progetto, che quindi possiedono una consolidata esperienza e conoscenza approfondita del dominio peritale e assicurativo.

È significativo notare che i periti esperti non hanno alcuna familiarità con l'ontologia, il tool dell'ontologia e il tool IA poiché questi ultimi, come menzionato, sono stati sviluppati internamente al progetto D.A.R.E. e non fanno parte degli strumenti di lavoro abitualmente utilizzati nella valutazione dei sinistri. Tale aspetto risulta di fondamentale importanza poiché permette di isolare il contributo effettivo degli strumenti nella precisione delle decisioni assunte, consentendo così di misurare il miglioramento delle prestazioni derivante dal loro utilizzo.

È stata coinvolta anche una figura con funzione di *revisore*, incaricata di confrontare le valutazioni prodotte dai principianti e dagli esperti con quelle di un altro attore chiave, il cosiddetto *Ground truth*, e assegnare dei punteggi.

Il *Ground truth* è una figura che vanta un'esperienza decennale nel settore peritale, dotata di una conoscenza capillare di dominio, che ha revisionato e validato tutte le perizie del campione sperimentale utilizzando gli strumenti disponibili. La conoscenza dei tool e del dominio gli ha permesso di generare annotazioni e output considerati come totalmente corretti e affidabili che sono stati utilizzati come baseline per il confronto e la valutazione della precisione dei risultati ottenuti dai principianti e dagli esperti nell'esecuzione dell'esperimento.

7.3 Criteri e metriche di valutazione

Per una valutazione oggettiva dell'efficacia del processo decisionale supportato da tool tecnologici, risulta fondamentale misurare l'accuratezza delle risposte fornite dai partecipanti rispetto ad un riferimento oggettivo e vero, il cosiddetto *Ground truth*.

In una prima fase, l'attenzione si concentra sui risultati ottenuti dai principianti per verificare se, nonostante il loro livello di competenza iniziale limitato, abbiano comunque le capacità per pervenire ad un risultato corretto e paragonabile a quello ottenuto dagli esperti, grazie all'utilizzo di strumenti strutturati. Parallelamente, si vuole esaminare l'accuratezza dei risultati degli esperti sia senza l'ausilio di strumenti, per valutarne le loro capacità indipendenti, sia con il supporto di questi ultimi allo scopo di comprendere se il loro impiego possa migliorare ulteriormente le loro prestazioni, riducendo l'influenza della soggettività e avvicinandosi quanto più possibile al *Ground truth*.

Un secondo parametro che si vuole misurare è rappresentato dal tempo impiegato per completare il processo decisionale, con e senza l'utilizzo di strumenti. Questo parametro permette di valutare l'efficacia dei tool nella riduzione dei tempi decisionali senza compromettere l'accuratezza, fornendo, al contempo, una misura del rapporto tra rapidità ed esattezza nel processo decisionale. Sarà possibile anche valutare eventuali variazioni nella rapidità della curva di apprendimento dovuta all'utilizzo del tool da parte dei partecipanti non esperti, come suggerito dalla letteratura (Brynjolfsson et al., 2023).

I partecipanti saranno poi chiamati a compilare, a seguito della visualizzazione e annotazione delle immagini presenti nella documentazione fotografica una tabella Excel, come quella riportata in figura 1, in cui si devono riportare le cause e gli effetti del danno, gli elementi danneggiati e l'ubicazione del danno all'interno dell'abitazione. I partecipanti sono chiamati ad esprimere su una scala da 0 a 100 la probabilità di copertura del danno e una stima dell'importo indennizzabile. In particolare, si è scelto di esprimere la copertura del danno mediante una scala graduata, al fine di cogliere con maggiore precisione il grado di confidenza che un partecipante manifesta nel formulare la propria decisione. Tale obiettivo risulterebbe irrealizzabile avvalendosi unicamente di una scala binaria (SI/NO).

Esperto (Ground Truth)

ID	N° Sinistro	codice / nome file FOTO	Effetto 1	Effetto 2	Effetto 3	Causa 1	Causa 2	Causa 3	Elemento danneggiato 1	Elemento danneggiato 2	Elemento danneggiato 3	Ubicazione foto	Copertura	Valore indennizzabile
1	20223450000789	C:\foto\foto.png	macchia_chiara	scrostamento		rottura_pluviali			Parete	Pavimento		Cucina	NO	€ -

NON esperto

ID	N° Sinistro	codice / nome file FOTO	Effetto 1	Effetto 2	Effetto 3	Causa 1	Causa 2	Causa 3	Elemento danneggiato 1	Elemento danneggiato 2	Elemento danneggiato 3	Ubicazione foto	Copertura	Valore indennizzabile
1	xxxxxxxx	C:\foto\foto.png	macchia_chiara	scrostamento		rottura_tubo_altr			Parete	Pavimento	Soffitto	Cucina		€ 900,00

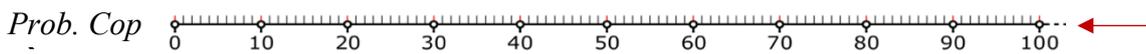


Figura 7.6 Screenshot esemplificativo del file Excel da compilare

Tali informazioni vengono successivamente confrontate con quelle inserite dal Ground truth da parte del revisore, il quale andrà ad assegnare degli *score* che sono suddivisi in: score di selezione, di interpretazione e di valutazione.

Score di selezione- Score di accuratezza della selezione

$$Score\ acc. = \frac{\sum_1^N \frac{\#Ann_{GT} - |\#Ann_{GT} - \#Ann_u|}{\#Ann_{GT}}}{N}$$

Con questa formula si intende valutare l'accuratezza raggiunta nell'annotazione delle immagini presenti all'interno della documentazione fotografica. Per ogni causa, effetto ed elemento danneggiato si confronta la numerosità delle annotazioni dell'utente coinvolto nell'esperimento rispetto a quella del Ground truth. Dopo che per ogni causa, effetto ed elemento danneggiato sono stati confrontati i due valori (dell'utente e del Ground truth) viene effettuata una media per calcolare uno score medio di accuratezza. Il denominatore di tale formula fa riferimento alla somma di cause, effetti ed elementi danneggiati, individuati dal Ground truth.

Score di interpretazione - Score CEEU (Causa-Effetto-Elemento danneggiato-Ubicazione)

$$Score\ CEEU = \frac{\sum_1^1 c_n * y}{n_c} * w_c + \frac{\sum_1^1 e_n * y}{n_e} * w_e + \frac{\sum_1^1 ed_n * y}{n_{ed}} * w_{ed} + u_n * w_u$$

Con questa formula si intende valutare l'accuratezza raggiunta nell'interpretazione delle cause, degli effetti, degli elementi danneggiati e dell'ubicazione del danno. I primi tre membri di tale formula rappresentano una media pesata per il peso associato. In particolare, al numeratore della media pesata c'è una sommatoria del numero di cause, effetti e elementi danneggiati, moltiplicata per una variabile y che rappresenta la correttezza dell'interpretazione della/e causa/e, effetto/i, elemento/ danneggiato/i e può assumere i seguenti valori:

- 0 = interpretazione errata
- 0,5 = interpretazione parzialmente corretta (es. interpretazione "rottura tubo" e "non rottura tubo scarico")

- 1 = interpretazione corretta

È stato stabilito di attribuire i seguenti pesi ai vari membri dello score CEEU:

- Causa 40%
- Elemento danneggiato 30%
- Effetto 20%
- Ubicazione 10%

Si è ritenuto di dare un maggiore peso alla causa e all'elemento danneggiato in quanto, la prima rappresenta la principale informazione per valutare la liquidabilità o meno del danno, mentre la seconda influenza notevolmente la tipologia di lavorazioni necessarie che sono la principale discriminante per la definizione dell'importo da liquidare.

Per quanto riguarda gli score di valutazione, ne sono stati definiti due, uno relativo alla copertura assicurativa e uno relativo all'importo da liquidare.

Score di valutazione I - Score copertura

$$\text{Score Cop.} = 1 - 2 * |\text{Cop}_{GT} - \text{Prob. cop}|$$

Il presente score è stato pensato per valutare con maggiore efficacia quanto un partecipante è in grado di identificare correttamente, con un certo grado di confidenza, la sussistenza di una copertura assicurativa del danno rappresentato nella documentazione fotografica peritale.

Il dominio della relazione è compreso tra -1 e 1, dove -1 equivale a sbagliare tutti gli esiti della perizia, viceversa il valore 1; lo zero rappresenta il valore intermedio nel caso in cui si prenda il 50% delle decisioni corrette e il 50% delle decisioni errate. Il 2 nella formula rappresenta una doppia sottrazione al valore 1 e più specificatamente con la seconda sottrazione del termine $|\text{Cop}_{GT} - \text{Prob. cop}|$ viene assegnata una penalità in funzione della distanza della decisione del partecipante da quella del Ground truth. In questo modo il valore atteso dello score da parte del principiante risulta nullo (cosa che non avverrebbe se non si sottraesse una penalità).

Score di valutazione II - Score liquidazione

$$\text{Score liq} = \sqrt{\frac{|\text{Valore}_{GT} - \text{Valore}_u|}{\text{Valore}_{GT}}}$$

Lo score è stato definito come uno scarto quadratico medio. Il quadrato della radice è stato scelto in quanto, se un partecipante fornisce un certo importo, più quell'importo è vicino a quello definito dal Ground Truth, più il partecipante ottiene uno score sempre più alto, viceversa nel caso in cui si allontana.

Tabella 7.2 Legenda degli score

Legenda	Pesi
$e_n = \text{Score effetto } n - \text{esimo}$ $c_n = \text{Score causa } n - \text{esima}$ $ed_n = \text{Score elemento danneggiato } n - \text{esimo}$ $u = \text{Score ubicazione}$	$w_e = 0,2$ $w_c = 0,4$ $w_{ed} = 0,3$ $w_{ed} = 0,1$
$n_{c,e,ed} = \text{numero totale di cause, effetti, elementi danneggiati identificati dal Ground truth}$ $N = n_c + n_e + n_{ed}$ somma di tutte le cause, effetti, elementi danneggiati identificati dal Ground truth $\#Ann_{CT} = \text{Numero di annotazioni per ogni causa, effetto, elemento danneggiato eseguite dal Ground truth}$ $\#Ann_u = \text{umero di annotazioni per ogni causa, effetto, elemento danneggiato dall'utente}$ $Cop_{CT} = \text{Valore binario copertura stabilito dal Ground Truth}$ $Valore_{CT} = \text{Valore stabilito dal Ground Truth}$ $Valore_u = \text{Valore stabilito dall'individuo}$	

7.4 Metodologia generale

La metodologia dell'esperimento che verrà esposta si articola in diverse fasi, ciascuna mirata a valutare l'efficacia e la precisione interpretativa dei partecipanti nell'analisi della documentazione fotografica peritale in assenza e presenza di assistenza tecnologica.

I partecipanti, per l'esecuzione dell'esperimento, vengono posti in una condizione "blind", cioè non hanno visionato in anticipo le immagini peritali, non conoscono da quale network sono state fornite, non hanno la possibilità di consultarsi i report peritali, né tra loro. Ciò serve a garantire che le loro interpretazioni e valutazioni siano indipendenti e prive di influenze o preconcetti derivanti da una visione preventiva della documentazione fotografica.

Lo strumento centrale per la raccolta dei dati è costituito da un *file Excel* appositamente strutturato in cui i partecipanti sono chiamati a inserire le informazioni richieste. Tale file è strutturato in modo da guidare l'inserimento dei dati tramite specifici menù "a tendina", che consentono di selezionare tra diverse opzioni predeterminate le tipologie di causa, gli effetti, gli elementi danneggiati, l'ubicazione del danno, la copertura assicurativa e il valore indennizzabile. L'obiettivo dell'utilizzo di menù a tendina è quello di standardizzare e uniformare la modalità di inserimento e tipologia di informazioni, riducendo al minimo le ambiguità e garantendo una maggiore omogeneità nei dati raccolti da ciascun partecipante. Per ogni perizia analizzata, inoltre, i partecipanti devono cronometrarsi.

Fase 1

Il *Ground truth*, utilizzando il tool dell'ontologia, analizza le immagini presenti nella documentazione fotografica peritale, procede alla classificazione delle immagini e alla selezione di quelle più rilevanti, segmentando le porzioni di immagini che mostrano danni o interventi specifici. Successivamente compie un lavoro di interpretazione, identificando la tipologia del danno, la causa e gli elementi danneggiati, e riporta le relative informazioni compilando un apposito file Excel. Inoltre, nello stesso file esprime un giudizio di copertura assicurativa in forma binaria (SI/NO) per ciascuna perizia oggetto dell'esperimento e stabilisce l'importo da liquidare.

La formalizzazione di tali valutazioni rappresenta la baseline di riferimento per confrontare i risultati delle successive analisi svolte dagli altri partecipanti.

Fase 2

I *principianti* svolgono un lavoro di interpretazione delle immagini presenti all'interno della documentazione fotografica peritale, verificando la correttezza dei suggerimenti proposti dal tool IA e scegliendo se seguirli o modificarli in base alle proprie deduzioni. Procedono, compilando

il file Excel, esprimono una probabilità di copertura assicurativa su una scala graduata da 0 a 100 e stimano un importo liquidabile.

Fase 3

Gli *esperti* analizzano le immagini presenti nella documentazione fotografica peritale, procedono con un'interpretazione dettagliata dei danni riportati, esprimendo un giudizio sulla copertura assicurativa in forma binaria (SI/NO) e stabiliscono l'importo da liquidare. Questa procedura riflette il metodo tradizionale della valutazione e liquidazione dei sinistri.

A valle di tale processo gli esperti sono invitati a formalizzare le proprie deduzioni compilando il file Excel, al fine di consentire un confronto diretto con i risultati dei principianti e del Ground truth.

Fase 4

- Fase 4.1: Gli esperti hanno la possibilità di rivedere quanto svolto nella fase 3 questa volta avvalendosi dell'ausilio del tool che evidenzia le immagini più rilevanti, le porzioni delle immagini di maggior interesse, una probabilità di copertura assicurativa (su una scala graduata da 0 a 100) nonché una stima dell'importo liquidabile. Essi verificano o meno la correttezza dei suggerimenti del tool e possono modificare le proprie valutazioni effettuate in precedenza.

In questa fase il tool non fornisce alcuna spiegazione in merito alle decisioni prese.

- Fase 4.2: Gli esperti hanno la possibilità di rivedere quanto svolto nella fase 4.1, tuttavia adesso il tool IA fornisce spiegazioni dettagliate delle decisioni (Explainable AI). Ciò serve a valutare l'influenza di specifiche e chiare spiegazioni nel processo decisionale.

Fase 5

Quando tutti i partecipanti hanno completato il proprio task, il revisore confronta gli output ottenuti secondo lo schema riportato di sotto e assegna i punteggi nel modo specificato nel capitolo precedente.

Per completezza si riporta una tabella riassuntiva dei confronti di seguito.

<i>Confronto</i>	<i>Score acc.</i>	<i>Score CEEU</i>	<i>Score Cop.</i>	<i>Score liq</i>
Principiante vs Ground truth				
Esperto (no tool) vs Ground truth				
Esperto (con tool; no expAI) vs Ground truth				
Esperto (con tool; si expAI) vs Ground truth				

Grazie ai punteggi calcolati è possibile, altresì, confrontare indirettamente le performance del principiante con quelle dell'esperto, permettendo di valutare in che misura un utente privo di esperienza di dominio, possa avvicinarsi alle competenze di un esperto. Inoltre, tali punteggi rendono possibile misurare anche l'incremento delle performance dell'utente esperto in funzione delle caratteristiche del supporto fornito.

7.5 Metodologia dell'esperimento pilota

Sulla base di quanto precedentemente illustrato, è stato svolto un *esperimento pilota* al fine di valutare il miglioramento dei processi decisionali attraverso il solo utilizzo del tool dell'ontologia.

In questo contesto, si vuole sottolineare l'importanza del ruolo svolto dall'ontologia in qualità di rappresentazione strutturata di concetti, ideata per definire in modo rigoroso e formale le entità, le proprietà, le funzioni e le relazioni esistenti in un ambiente specifico. È uno strumento fondamentale per la raccolta e l'organizzazione della conoscenza collettiva e permette di identificare e comprendere quali siano gli elementi rilevanti da analizzare e valutare, consentendo di uniformare il processo di interpretazione e giudizio. L'essere umano possiede una razionalità di base che lo orienta nell'esprimere giudizi, ma senza un sistema di riferimento esplicito, come l'ontologia, la razionalità può risultare dispersa o inconsistente.

L'esperimento pilota, dunque, è finalizzato a valutare l'efficacia di uno strumento basato sull'ontologia nel supportare utenti principianti, migliorandone le prestazioni fino a un livello comparabile a quello degli utenti esperti, e a verificare se tale strumento possa apportare benefici anche all'attività degli utenti già esperti.

Tale esperimento, inoltre, si pone l'obiettivo di suggerire anche eventuali modifiche e miglioramenti di natura operativa funzionali all'esperimento generale.

Nello specifico, le risorse coinvolte in questo esperimento pilota sono state il Ground truth, che ha le caratteristiche menzionate in precedenza, un Perito esperto e 2 principianti.

Anche in questo caso, per l'esecuzione dell'esperimento, i partecipanti sono stati posti in condizione "blind", senza aver preso visione in anticipo delle immagini peritali, senza conoscere da quale network sono state fornite, senza possibilità di consultazione tra di loro o con i report peritali. In analogia al caso precedente, la raccolta dei dati è avvenuta tramite un file Excel strutturato con menù "a tendina" per facilitare l'inserimento standardizzato di informazioni come tipo di causa, effetti, elementi danneggiati, ubicazione del danno, copertura assicurativa e valore indennizzabile, minimizzando le ambiguità e assicurando uniformità nei dati raccolti. Per ogni perizia analizzata, inoltre, i partecipanti hanno cronometrato il tempo.

Fase 1

Il *Ground truth*, utilizzando il tool dell'ontologia, ha preso in esame 65 perizie.

Per ogni perizia, ha analizzato le immagini presenti nella documentazione fotografica, proceduto alla classificazione delle immagini e alla selezione di quelle più rilevanti, segmentando le porzioni di immagini che mostrano danni o interventi specifici. Parallelamente il Perito esperto ha svolto un lavoro di interpretazione, identificando la tipologia del danno, la causa e gli elementi danneggiati, e ha riportato tali informazioni compilando l'apposito file Excel. Nello stesso file ha espresso un giudizio di copertura assicurativa in forma binaria (SI/NO). Come già precisato, la formalizzazione di tali valutazioni rappresenta la baseline di riferimento per confrontare i risultati delle successive analisi svolte dagli altri partecipanti.

Fase 2

• *Fase 2.1 (Selezione)*

I 2 *principianti*, utilizzando l'apposito tool dell'ontologia, hanno preso in esame 65 perizie ciascuno.

Per ogni perizia, hanno analizzato le immagini presenti nella documentazione fotografica peritale, compiendo un lavoro di selezione, al fine di determinare le immagini rilevanti, e segmentando le porzioni interessate da danni o da lavorazioni, associando delle etichette.

- *Fase 2.2 (Interpretazione)*

I partecipanti, una volta selezionate e annotate le immagini, hanno compiuto un lavoro di interpretazione, identificando la tipologia di danno, la relativa causa e gli elementi danneggiati e hanno riportato queste informazioni compilando l'apposito file Excel.

- *Fase 2.3 (Valutazione)*

Nello stesso file Excel hanno espresso una probabilità di copertura assicurativa (su una scala graduata da 0 a 100)

Fase 3

Il revisore ha confrontato i risultati dei principianti con quelli del Ground truth e del Perito esperto. In particolare, lo score acc è stato calcolato rispetto al Ground truth, per capire effettivamente il grado di accuratezza con cui i principianti sono riusciti a selezionare e annotare le immagini presenti in perizia, mentre gli score CEEU e Cop. rispetto al Perito esperto.

<i>Confronto</i>	<i>Score acc</i>	
Principiante vs Ground truth		
	<i>Score CEEU.</i>	<i>Score Cop.</i>
Principiante vs Perito esperto		

Lo *score acc.* è stato adattato alla struttura dei dati attualmente in possesso; dunque, viene riportata la relativa descrizione di seguito

$$Score\ acc. = w_1 * \frac{\#Classi_{GT} - |\#Classi_{GT} - \#Classi_u|}{\#Classi_{GT}} + w_2 * \frac{\#ImgAnn_{GT} - |\#ImgAnn_{GT} - \#ImgAnn_u|}{\#ImgAnn_{GT}}$$

Come si nota dalla formula, lo score acc. è costituito da due membri: il primo membro tiene in considerazione l'accuratezza con cui le annotazioni identificate dall'utente rispecchiano quelle stabilite dal Ground truth, mentre il secondo si riferisce all'accuratezza nella selezione delle immagini rilevanti (in cui si inserisce almeno un'annotazione). Si è ritenuto di dare un peso maggiore al primo membro (0,6) in quanto più influente nella valutazione dell'accuratezza.

Pesi
$w_1 = 0,6$
$w_2 = 0,4$
Legenda
$\#Classi_{GT}$ = numero di classi a cui appartengono le annotazioni identificate dal Ground truth
$\#Classi_u$ = numero di classi a cui appartengono le annotazioni identificate dall'utente
$\#ImgAnn_{GT}$ = numero di immagini in cui il Ground truth ha inserito almeno un'annotazione
$\#ImgAnn_u$ = numero di immagini in cui l'utente ha inserito almeno un'annotazione

7.6 Risultati dell'esperimento pilota

<i>Confronto</i>	<i>Score acc</i>	
Principiante 1 vs Ground truth	67,56%	
Principiante 2 vs Perito esperto	67,72%	
	<i>Score CEEU.</i>	<i>Score Cop.</i>
Principiante 1 vs Perito esperto	58,20%	24,30%
Principiante 2 vs Perito esperto	55,28%	20,00%

Dalla tabella emerge che i risultati ottenuti dai due partecipanti risultano essere complessivamente allineati nei tre score di valutazione considerati. Sebbene il coinvolgimento di soli due partecipanti non consenta di trarre conclusioni statisticamente significative in quanto sarebbe necessario un campione più ampio, si può comunque osservare come l'impiego di uno strumento strutturato, quale l'ontologia, consenta di ridurre il divario tra i risultati ottenuti da soggetti privi di esperienza specifica del dominio e quelli del Perito esperto e del Ground truth, il cui output, come precisato, è considerato affidabile e veritiero.

Score di selezione

Per quanto riguarda lo *score di accuratezza*, i risultati ottenuti dai due partecipanti sono pressoché simili, con un lieve vantaggio per il Principiante 2 (+0,16%). Questo dato è indicativo di una capacità moderatamente soddisfacente nel selezionare correttamente i dati disponibili grazie al supporto dell'ontologia, riconoscendo e classificando correttamente circa i due terzi delle informazioni. Tuttavia, un punteggio vicino al 70% evidenzia ancora un margine significativo di errore, attribuibile sia alla mancanza di esperienza dei partecipanti nel dominio peritale, sia alla complessità intrinseca del compito che richiede la precisa selezione degli elementi più rappresentativi.

Score di interpretazione

Nella fase di interpretazione, i partecipanti hanno ottenuto uno *Score CEEU* relativamente basso, evidenziando una difficoltà significativa in questa fase del processo riconducibile, con molta probabilità, alla presenza di un gap di informazioni nella documentazione fotografica fornita.

Un esame dettagliato delle componenti dello *Score CEEU* mostra che una delle criticità principali risiede nell'identificazione delle cause del danno, un compito reso complesso dalla frequente assenza o difficile individuazione di tali cause all'interno delle perizie. Le cause del danno sono state infatti identificate correttamente solo nel 53% dei casi.

Nonostante il basso punteggio legato alle cause, le performance dei partecipanti risultano migliori nell'identificazione degli effetti del danno, con uno score che si assesta al 58% circa. Questo dato evidenzia che, sebbene il riconoscimento delle cause possa rappresentare un limite, la determinazione degli effetti non ne risulta necessariamente compromessa.

L'individuazione degli elementi danneggiati, invece, mostra un'accuratezza più elevata, raggiungendo una percentuale del 64%, un risultato che segnala una discreta capacità dei partecipanti di identificare gli elementi oggetto del danno.

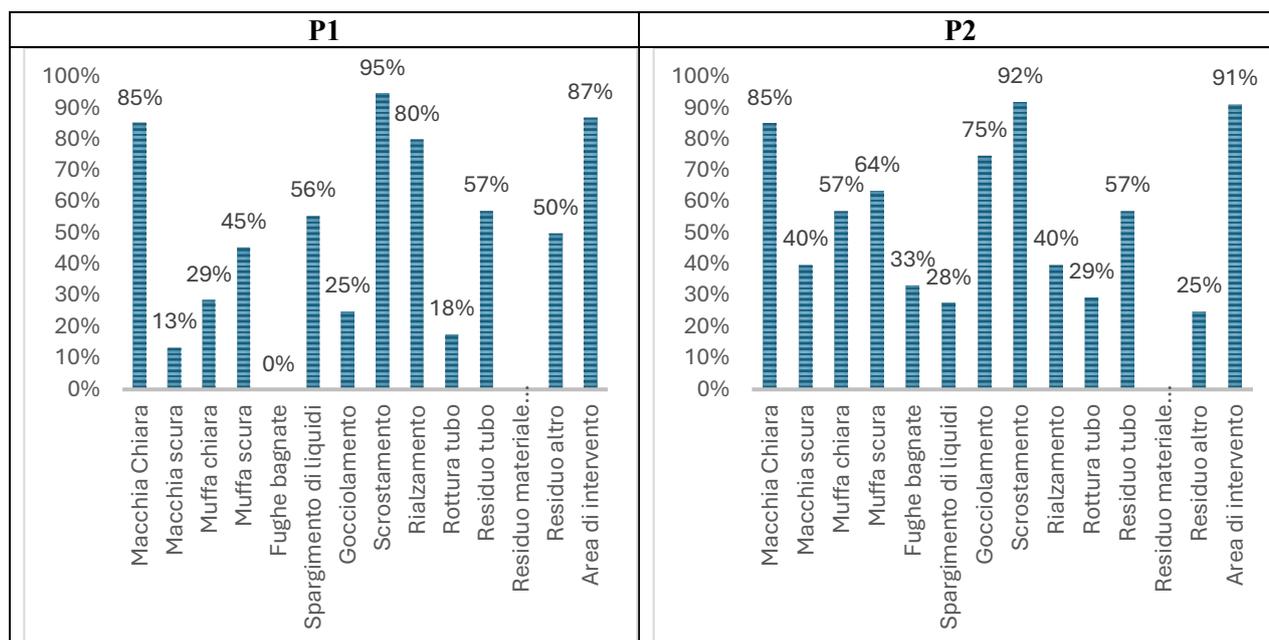
Un ulteriore aspetto critico emerso riguarda la determinazione dell'ubicazione della causa del danno, che risulta problematica per le seguenti ragioni: da un lato, può sussistere una correlazione con l'individuazione della causa del danno che condiziona la determinazione dell'ubicazione, dall'altro il mancato accesso ai report peritali ha impedito ai partecipanti di ricavare con precisione

il luogo specifico del danno (ad esempio, un danno localizzato in un salone è stato genericamente classificato dai partecipanti sotto la categoria “stanza”).

Score di valutazione

Infine, relativamente allo *Score copertura*, i partecipanti hanno ottenuto valori compresi tra il 20% e il 24% circa e ciò indica una capacità piuttosto limitata nel determinare la presenza o meno di una copertura assicurativa. Questo risultato può essere attribuito principalmente al fatto che la parte mancante di informazione che non si riesce ad evincere dalla documentazione fotografica è contenuta esclusivamente nel report peritale.

Occorrenza delle tipologie di errore riscontrate nell'esecuzione dell'esperimento



Il seguente grafico rappresenta la frequenza delle annotazioni per tipologia correttamente individuate dai partecipanti rispetto al Ground truth. Nello specifico le percentuali rappresentate sono state calcolate come rapporto tra le volte in cui sia il principiante che il Ground truth hanno correttamente selezionato una tipologia di annotazione e il totale delle volte in cui il Ground truth ha selezionato una tipologia di annotazione.

Le percentuali risultano essere abbastanza allineate tra i due partecipanti eccetto che nelle seguenti voci: il partecipante 1 presenta percentuali maggiori rispetto all'altro nello “spargimento liquidi” e “rialzamento”, mentre il partecipante 2 nelle voci “fughe bagnate” e “gocciolamento”.

Risulta evidente che le voci individuate con maggior accuratezza sono “macchia chiara”, “scrostamento” e “area di intervento”, mentre i partecipanti hanno riscontrato maggiori difficoltà nell'individuazione delle voci “macchia scura” e “rottura tubo”.

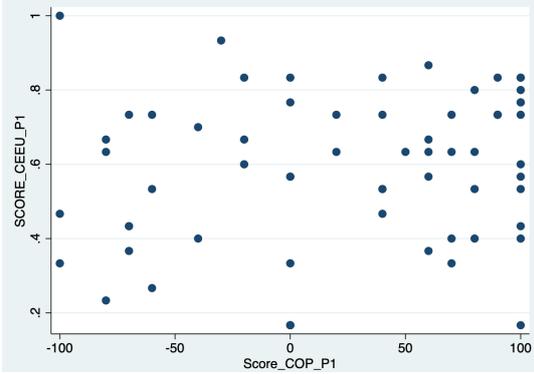
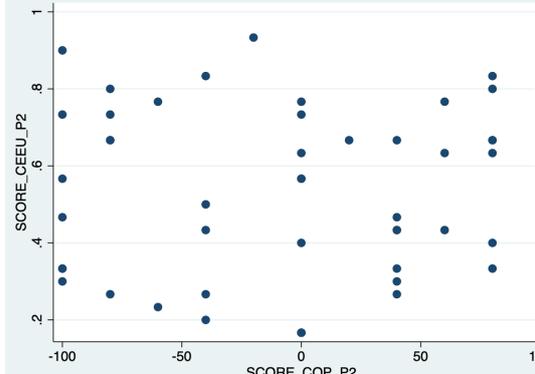
Correlazioni tra le variabili

Sulla base degli score per perizia ottenuti dai 2 partecipanti, attraverso il software STATA si è cercato di capire se vi fossero le seguenti correlazioni

- 1) Score CEEU, Score acc. e Score Cop.
- 2) Score Ceeu e Score Cop.
- 3) Score Ceeu e foto annotate
- 4) Score Cop e foto annotate
- 5) Score acc. e foto annotate

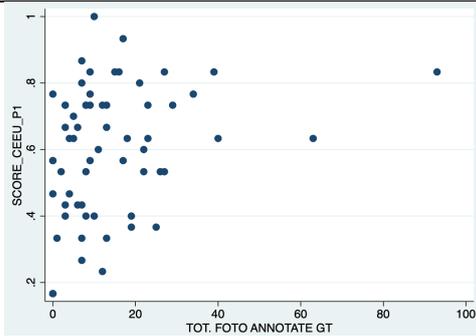
P1					P2				
		SCORE_~1	Score_~1	SCOREA~1			SCO~U_P2	SCO~P_P2	SCO~C_P2
SCORE_CEEU~1		1.0000			SCORE_CEEU~2		1.0000		
Score_COP_P1		0.1775	1.0000		SCORE_COP_P2		0.1774	1.0000	
SCOREACCP1		-0.0519	0.0070	1.0000	SCORE_ACC_P2		0.0196	0.1056	1.0000

Linear regression Number of obs = 65 F(1, 63) = 1.93 Prob > F = 0.1696 R-squared = 0.0315 Root MSE = .20797					Linear regression Number of obs = 65 F(1, 63) = 2.15 Prob > F = 0.1477 R-squared = 0.0315 Root MSE = .21775						
SCORE_CEEU~1	Coef.	Robust Std. Err.	t	P> t	[95% Conf. Interval]	SCORE_CEEU~2	Coef.	Robust Std. Err.	t	P> t	[95% Conf. Interval]
Score_COP_P1	.0005695	.0004098	1.39	0.170	-.0002495 .0013885	SCORE_COP_P2	.0005434	.0003707	1.47	0.148	-.0001974 .0012842
_cons	.5682085	.0295759	19.21	0.000	.5091058 .6273113	_cons	.5419527	.0294636	18.39	0.000	.4830743 .6008311

I 2 score non risultano correlati (o correlati con un p-value del 17%)	I 2 score non risultano correlati (o correlati con un p-value del 15% circa)
	

Linear regression Number of obs = 64 F(1, 62) = 12.02 Prob > F = 0.0010 R-squared = 0.0963 Root MSE = .1961					Linear regression Number of obs = 64 F(1, 62) = 5.45 Prob > F = 0.0228 R-squared = 0.0629 Root MSE = .21053						
SCORE_CEEU_P1	Coef.	Robust Std. Err.	t	P> t	[95% Conf. Interval]	SCORE_CEEU_P2	Coef.	Robust Std. Err.	t	P> t	[95% Conf. Interval]
TOTFOTOANNOTATEGT	.0040994	.0011826	3.47	0.001	.0017355 .0064633	TOTFOTOANNOTATE	.0034192	.0014644	2.33	0.023	.0004919 .0063466
_cons	.532175	.033691	15.80	0.000	.4648277 .5995223	_cons	.5119467	.0354564	14.44	0.000	.4410704 .5828231

Lo score CEEU e il totale di foto annotate risultano correlati	Lo score CEEU e il totale di foto annotate risultano correlati
--	--

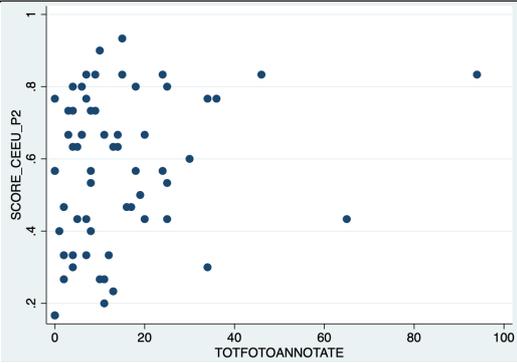


Linear regression

Number of obs	=	65
F(1, 63)	=	16.76
Prob > F	=	0.0001
R-squared	=	0.0816
Root MSE	=	63.132

SCORE_COP_P1	Coef.	Robust Std. Err.	t	P> t	[95% Conf. Interval]
TOTFOTOANNOTATEGT	1.207285	.2949186	4.09	0.000	.6179371 1.796633
_cons	7.962912	9.647762	0.83	0.412	-11.3166 27.24243

I due score sono correlati

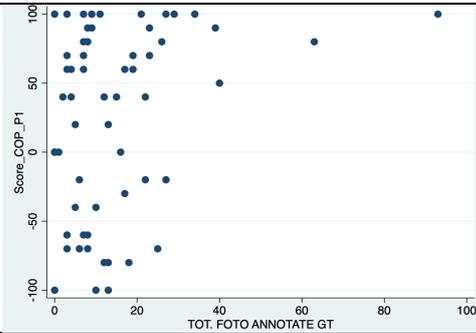


Linear regression

Number of obs	=	64
F(1, 62)	=	9.95
Prob > F	=	0.0025
R-squared	=	0.0565
Root MSE	=	70.033

SCORE_COP_P2	Coef.	Robust Std. Err.	t	P> t	[95% Conf. Interval]
TOTFOTOANNOTATE	1.073875	.3404686	3.15	0.002	.3932885 1.754462
_cons	4.017771	10.36304	0.39	0.700	-16.69766 24.7332

I due score sono correlati

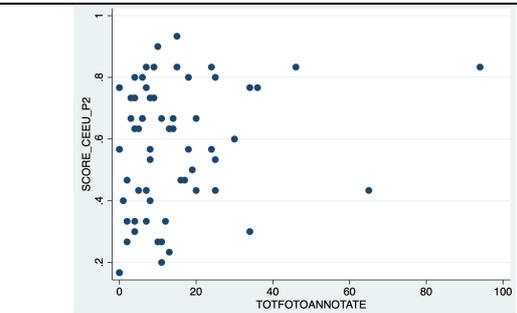


Linear regression

Number of obs	=	65
F(1, 63)	=	3.35
Prob > F	=	0.0719
R-squared	=	0.0464
Root MSE	=	.25298

SCORE_ACC_P1	Coef.	Robust Std. Err.	t	P> t	[95% Conf. Interval]
TOTFOTOANNOTATEGT	-.0035828	.0019575	-1.83	0.072	-.0074947 .0003291
_cons	.7242929	.0457372	15.84	0.000	.6328943 .8156914

Lo score acc. e il totale delle foto annotate risultano correlati

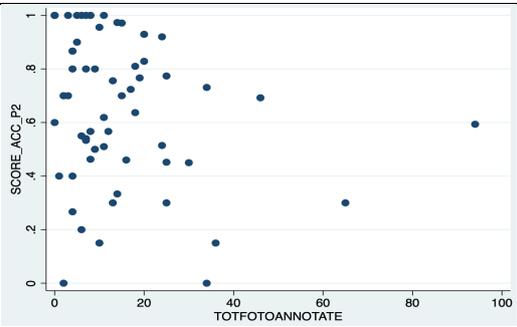
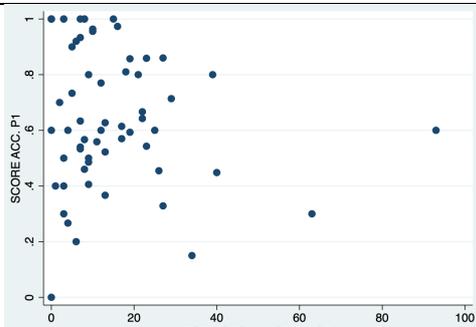


Linear regression

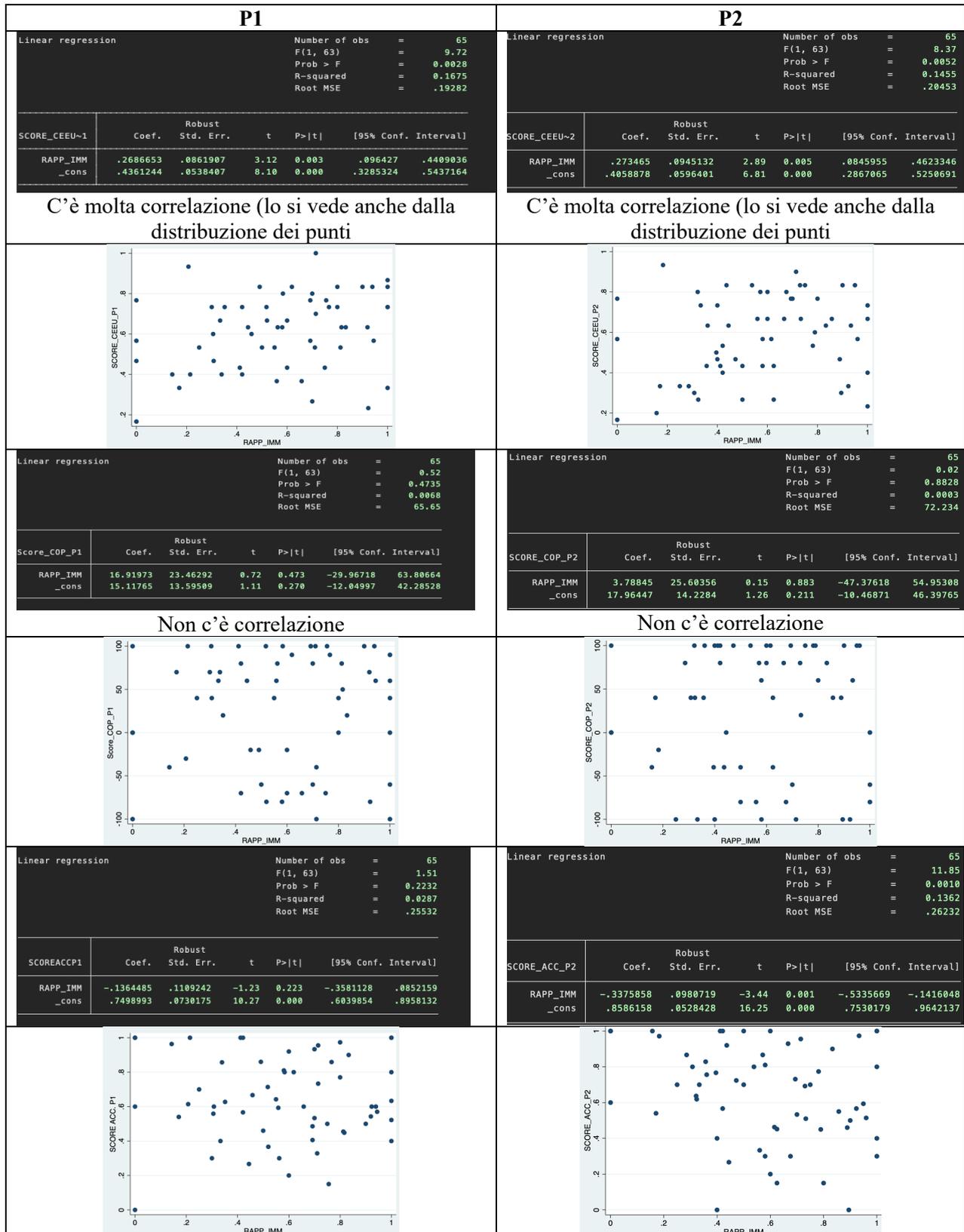
Number of obs	=	64
F(1, 62)	=	4.56
Prob > F	=	0.0367
R-squared	=	0.0740
Root MSE	=	.27089

SCORE_ACC_P2	Coef.	Robust Std. Err.	t	P> t	[95% Conf. Interval]
TOTFOTOANNOTATE	-.0047986	.0022479	-2.13	0.037	-.009292 -.0003052
_cons	.7380187	.0462773	15.95	0.000	.6455118 .8305257

Lo score acc. e il totale delle foto annotate risultano correlati



Si è cercato di capire, inoltre, se vi fosse correlazione tra gli score e il rapporto tra le immagini annotate dai partecipanti e il totale delle immagini all'interno della documentazione fotografica.



7.7 Osservazioni finali e sviluppi futuri

Nel corso dell'esecuzione dell'esperimento, e in particolare della fase 2, sono emerse alcune difficoltà nelle procedure di selezione e di interpretazione delle immagini.

Per quanto riguarda la fase di selezione, i partecipanti sono riusciti agevolmente a discernere le immagini rilevanti da quelle irrilevanti. Tuttavia, a causa dell'incertezza interpretativa sono state riscontrate criticità nell'associazione delle etichette appropriate alle immagini. Ad esempio, nel caso di una macchia visibile su una parete, in alcuni casi è risultato complesso determinare se la stessa fosse da classificare come chiara o scura. A tal fine, si ritiene opportuno, per i prossimi sviluppi, fornire ai partecipanti un *vademecum* contenente una descrizione dettagliata di ciascun elemento selezionabile, al fine di ridurre le incertezze interpretative e garantire maggiore uniformità nell'esecuzione dell'esperimento.

Nella fase di interpretazione, in alcuni casi, non è stato possibile individuare la causa del danno in quanto essa non era rappresentata visivamente nelle immagini presenti nella documentazione peritale. Di conseguenza, si è verificata l'impossibilità di stabilirne l'ubicazione. Pertanto, in tali casi, i partecipanti hanno inserito una probabilità di copertura inferiore. In aggiunta a ciò, in quelle perizie contenenti un numero elevato di fotografie (corrispondenti a casi più complessi) è emersa una maggiore difficoltà nell'identificare la causa del danno e, conseguentemente, nel determinarne l'ubicazione.

È opportuno, inoltre, formulare alcune osservazioni in merito alla compilazione del file Excel. Le tre colonne dedicate agli effetti del danno si sono rivelate, in taluni casi, insufficienti pertanto si raccomanda di prevedere ulteriori colonne, fino ad un totale di cinque. Per quanto concerne le colonne riservate alle cause, tre sono risultate eccessive, in quanto in una singola perizia un danno può essere attribuito al massimo a due cause; si suggerisce dunque di ridurre a due le colonne relative alle cause.

Infine, si consiglia di ridurre il livello di dettaglio dell'elenco relativo agli elementi danneggiati, proponendo di escludere le voci "tinteggiatura" e "intonaco", poiché risultano già incluse nelle voci "parete" e/o "soffitto".

Capitolo 8

Conclusione e open points

La presente tesi ha consentito di analizzare in modo approfondito il ruolo in costante crescita dell'Intelligenza artificiale all'interno dei processi decisionali, mettendone in luce le potenzialità nel migliorare la qualità delle decisioni in ambiti complessi e dinamici, ma al contempo evidenziando le numerose sfide legate ad una sua corretta, consapevole e realmente efficace implementazione.

L'analisi si è focalizzata, in particolare, sull'importanza cruciale rivestita dal *Data Quality Management*, dimostrando come la validità e l'affidabilità dei risultati prodotti dai modelli di IA siano strettamente legate alla qualità dei dati di input. A tal riguardo, è stata sottolineata la necessità imprescindibile di adottare un approccio metodico e rigoroso alla raccolta, alla pulizia e alla successiva validazione dei dati, approccio che è stato seguito nel caso studio analizzato, contribuendo a colmare un gap presente in letteratura.

Tale caso studio ha evidenziato la presenza, all'interno del dataset fornito e funzionale all'implementazione di sistemi di IA, di criticità e disomogeneità che era necessario eliminare o ridurre quanto più possibile e ciò è stato realizzato soprattutto grazie all'importante contributo offerto dagli esperti di dominio senza il quale né lo sviluppo dell'ontologia, né il processo di data cleaning sarebbero stati possibili.

Un ulteriore aspetto di rilievo approfondito riguarda i potenziali impatti che l'IA può generare nel settore esaminato, un tema esplorato attraverso l'esperimento di valutazione comparativa.

Tale esperimento, infatti, ha fornito evidenze empiriche che hanno attestato l'efficacia, in una prima fase, di uno strumento di classificazione basato su un'ontologia nell'influire positivamente sulla qualità delle decisioni assunte da individui con competenze limitate in un contesto operativo reale quale il settore assicurativo. L'ontologia, integrata nel relativo strumento, ha dimostrato la propria capacità di supportare validamente i processi decisionali, consentendo di conseguire risultati che, sebbene soddisfacenti, rappresentano un promettente preludio a sviluppi futuri e, in particolare, per la realizzazione di strumenti più complessi e articolati quali quelli basati su IA.

Quanto svolto apre la strada ad ulteriori indagini sperimentali mirate al coinvolgimento di soggetti esperti, come nel caso specifico di esperti del dominio peritale, al fine di verificare se sia possibile ottenere un incremento ancora più marcato delle performance decisionali, con conseguente riduzione della soggettività intrinseca ai processi decisionali umani.

Infine, si ritiene opportuno promuovere lo sviluppo di ulteriori studi sperimentali su scala più ampia, nonché approfondire l'analisi del comportamento dei modelli di IA in ulteriori contesti operativi reali, al fine di consolidare e ampliare le evidenze empiriche e teoriche presentate nel corso di questo lavoro.

Elenco delle figure

Figura 2.1 Investimenti nell'IA per settore nel 2023.....	12
Figura 2.2 Complessità del compito, incertezza del compito e Intelligenza Artificiale.....	13
Figura 2.3 Value Chain del settore assicurativo.....	16
Figura 3.1 Dai dati grezzi alle decisioni.....	20
Figura 3.2 Data Quality Management nel settore assicurativo.....	22
Figura 3.3 Processo di addestramento di un modello di classificazione di AI per la Computer Vision (Fonte Urbach N., Roglinger M. 2021).....	26
Figura 3.4 Data preparation con data cleaning.....	27
Figura 3.5 Influenza dell'aggiunta di dati sull'accuratezza del modello.....	29
Figura 3.6 Relazione tra IA e conoscenza (Sanzogni et al., 2017).....	31
Figura 3.7 Andamento del bias e della varianza in funzione della complessità del modello.....	34
Figura 3.8 Andamento del training error e del validation error in funzione della complessità del modello.....	35
Figura 4.1 Framework per il training e test del modello di IA nel progetto.....	43
Figura 4.2 Schema di classificazione delle immagini in perizia.....	48
Figura 4.3 Schermata principale del tool di annotazione.....	49
Figura 4.4 Schermata principale del tool con l'ultimo aggiornamento.....	50
Figura 6.1 Evoluzione dell'F1 Score.....	59
Figura 6.2 Importo liquidato in funzione del numero di foto presenti in perizia.....	60

Elenco delle tabelle

Tabella 3.1 Tipologie di Machine Learning, costruzione del modello sulla base dei dati ed esempi (Vijayaraj et al., 2023).....	24
Tabella 3.2 Sintesi degli elementi che contribuiscono al miglioramento delle performance di un modello di ML.....	36
Tabella 3.3 Principali metriche per la valutazione di modelli di ML di classificazione.....	37
Tabella 4.1 Esempi di elementi contenuti nell'ontologia.....	46
Tabella 5.1 Risultati e statistiche del campione analizzato (200 perizie).....	51
Tabella 5.2 Risultati analitici: confronto tra i due network.....	53
Tabella 5.3 Statistiche analitiche: confronto tra i due network.....	53
Tabella 5.4 Cause di mancata liquidazione per il network 1 e relativa frequenza.....	55
Tabella 5.5 Cause di ambiguità nell'analisi della liquidazione delle perizie per il network 1 e relativa frequenza.....	55
Tabella 5.6 Cause di mancata liquidazione per il network 2 e relativa frequenza.....	55
Tabella 5.7 Cause di ambiguità nell'analisi della liquidazione delle perizie per il network 2 e relativa frequenza.....	56
Tabella 5.8 Risultati a confronto delle quattro risorse.....	57
Tabella 5.9 Statistiche a confronto delle quattro risorse.....	57
Tabella 7.1 Riepilogo dei Gap, delle Research questions e degli obiettivi dell'esperimento...	64

Bibliografia

1. Banks, S., Ocampo, A. C., Marrone, M., Restubog, S. L. D., & Woo, S. E. (2024). A multilevel review of artificial intelligence in organizations: Implications for organizational behavior research and practice. In *Journal of Organizational Behavior* (Vol. 45, Issue 2, pp. 159–182). John Wiley and Sons Ltd. <https://doi.org/10.1002/job.2735>
2. Batini, C., Scannapieca, M., *Data Quality Concepts, Methodologies and Techniques*, Milano, Springer, 2006.
3. Benbya, H. (n.d.). *Artificial Intelligence in Organizations: Current State and Future Opportunities*. <https://aisel.aisnet.org/misqe/vol19/iss4/4>
4. Brynjolfsson, E., Li, D., Raymond, L. R., (2023). Generative AI at work. <http://www.nber.org/papers/w31161>
5. Bullock, J. B. (2019). Artificial Intelligence, Discretion, and Bureaucracy. *American Review of Public Administration*, 49(7), 751–761. <https://doi.org/10.1177/0275074019856123>
6. Cai, L., & Zhu, Y. (2015). The challenges of data quality and data quality assessment in the big data era. *Data Science Journal*, 14. <https://doi.org/10.5334/dsj-2015-002>
7. Caplin, A., Deming, D. J., Li, S., Martin, D. J., Marx, P., Weidmann, B., Jiada Ye, K., Deming, D., & Martin, D. (2024). *The ABC's of Who Benefits from Working with AI: Ability, Beliefs, and Calibration*. <http://www.nber.org/papers/w33021>
8. Carey, M. J., Ceri, S., Board, E., Bernstein, P., Dayal, U., Faloutsos, C., Freytag, J. C., Gardarin, G., Jonker, W., Krishnamurthy, V., Neimat, M.-A., Valduriez, P., Weikum, G., Whang, K.-Y., & Widom, J. (n.d.). *Data-Centric Systems and Applications Series Editors*.
9. Cichy, C., & Rass, S. (2019). An overview of data quality frameworks. *IEEE Access*, 7, 24634–24648. <https://doi.org/10.1109/ACCESS.2019.2899751>
10. Cosma, S., & Rimo, G. (2024). Redefining insurance through technology: Achievements and perspectives in Insurtech. In *Research in International Business and Finance* (Vol. 70). Elsevier Ltd. <https://doi.org/10.1016/j.ribaf.2024.102301>
11. Curado, C. (2006). *THE KNOWLEDGE BASED-VIEW OF THE FIRM: FROM THEORETICAL ORIGINS TO FUTURE IMPLICATIONS*. <https://www.researchgate.net/publication/242091944>
12. D'acquisto, G., *Decisioni algoritmiche, Equità, causalità, trasparenza*, Torino, G. Giappichelli Editore, 2022.
13. *DATA CLEANING*. (2016).
14. *DATA CLEANING Dealing with messy data*. (2016).

15. Dell'Acqua, F., Saran, A., Mcfowland, R. E., Krayner, L., Mollick, E., Candelon, F., Lifshitz-Assaf, H., Lakhani, K. R., & Kellogg, K. C. (2023). *Navigating the Jagged Technological Frontier: Field Experimental Evidence of the Effects of AI on Knowledge Worker Productivity and Quality*. <https://ssrn.com/abstract=4573321>
16. European Union Agency for Fundamental Rights, F. (n.d.). *Data quality and artificial intelligence-mitigating bias and error to protect fundamental rights HELPING TO MAKE FUNDAMENTAL RIGHTS A REALITY FOR EVERYONE IN THE EUROPEAN UNION FRA Focus Contents*. <https://doi.org/10.2811/615718>
17. Fenstermacher, K. D. (2005). The tyranny of tacit knowledge: What artificial intelligence tells us about knowledge representation. *Proceedings of the Annual Hawaii International Conference on System Sciences*, 243. <https://doi.org/10.1109/hicss.2005.620>
18. Fokas, A. S. (2023). Can artificial intelligence reach human thought? In *PNAS Nexus* (Vol. 2, Issue 12). National Academy of Sciences. <https://doi.org/10.1093/pnasnexus/pgad409>
19. Gamble, J. R. (2020). Tacit vs explicit knowledge as antecedents for organizational change. In *Journal of Organizational Change Management* (Vol. 33, Issue 6, pp. 1123–1141). Emerald Group Holdings Ltd. <https://doi.org/10.1108/JOCM-04-2020-0121>
20. *Global Insurance Report 2023: Reimagining life insurance*. (n.d.).
21. Graham, H., Jones, M., Madigan, K., Zheng, S., *Data Quality Management in the P&C Insurance Sector*, Arlington, Casualty Actuarial Society, 2020.
22. Heale, B. (2014). *Data Quality is the Biggest Challenge*. <https://www.moodyanalytics.com/risk-perspectives-magazine/managing-insurance-risk/insurance-regulatory-spotlight/data-quality-is-the-biggest...1/10>
23. Holland, C. P. (n.d.). *Artificial Intelligence (AI) and Digital Transformation in the Insurance Market: A Case Study Analysis of BGL Group*. <https://hdl.handle.net/10125/79890>
24. Jussupow, E., Spohrer, K., Heinzl, A., Gawlitzka, J. (2021). *Augmenting medical diagnosis decisions? An investigation into physicians' decision-making process with artificial intelligence*.
25. Keller, B. (n.d.). *Promoting Responsible Artificial Intelligence in Insurance*. www.genevassociation.org
26. Kumar, N., Srivastava, J. D., & Bisht, H. (n.d.). *Artificial Intelligence in Insurance Sector*.
27. Manmeet Kaur. (2024). A Comprehensive Overview of Artificial Intelligence-Based Classification Techniques. *International Journal of Science and Research Archive*, 11(2), 125–129. <https://doi.org/10.30574/ijsra.2024.11.2.0387>
28. McKinsey & Company, *Global Insurance Report 2023: Reimagining life insurance, 2023*
29. Mohammad Hossein Jarrahia, D. A. A. E. P. S. (2022). *Artificial intelligence and knowledge management: A partnership between human and AI*.
30. Musashi Ai White paper Understanding the Metrics for Better Evaluation of Your AI Models <https://musashiai.com/wp-content/uploads/2024/08/Musashi-AI-White-Paper-Understanding-the-metrics-for-better-evaluation-of-your-AI-Models.pdf>
31. Newman, J., Mintrom, M., & O'Neill, D. (2022). Digital technologies, artificial intelligence, and bureaucratic transformation. *Futures*, 136. <https://doi.org/10.1016/j.futures.2021.102886>

32. OECD (2020), The Impact of Big Data and Artificial Intelligence (AI) in the Insurance Sector, www.oecd.org/finance/Impact-Big-Data-AI-in-the-Insurance-Sector.htm.
33. Rawat, S., Rawat, A., Kumar, D., & Sabitha, A. S. (2021). Application of machine learning and data visualization techniques for decision support in the insurance sector. *International Journal of Information Management Data Insights*, 1(2). <https://doi.org/10.1016/j.jjime.2021.100012>
34. Riikkinen, M., Saarijärvi, H., Sarlin, P., & Lähtenmäki, I. (2018). Using artificial intelligence to create value in insurance. *International Journal of Bank Marketing*, 36(6), 1145–1168. <https://doi.org/10.1108/IJBM-01-2017-0015>
35. Sanzogni, L., Guzman, G., & Busch, P. (2017). Artificial intelligence and knowledge management: questioning the tacit dimension. *Prometheus: Critical Studies in Innovation*, 35(1), 37–56. <https://doi.org/10.1080/08109028.2017.1364547>
36. Sessions, V., & Valtorta, M. (n.d.). *The effects of data quality on machine learning algorithms*.
37. Shi, G., Ma, Z., Feng, J., Zhu, F., Bai, X., & Gui, B. (2020). The impact of knowledge transfer performance on the artificial intelligence industry innovation network: An empirical study of Chinese firms. *PLoS ONE*, 15(5). <https://doi.org/10.1371/journal.pone.0232658>
38. Shrestha, Y. R., Ben-Menahem, S. M., & von Krogh, G. (2019). Organizational Decision-Making Structures in the Age of Artificial Intelligence. *California Management Review*, 61(4), 66–83. <https://doi.org/10.1177/0008125619862257>
39. Stuart Geiger, R., Yu, K., Yang, Y. (2019) Garbage In, Garbage Out? Do Machine Learning Application Papers in Social Computing Report Where Human-Labeled Training Data Comes From? <https://doi.org/10.1145/3351095.3372862>
40. Sushant K, S. (2020). A Commentary on the Application of Artificial Intelligence in the Insurance Industry. *Trends in Artificial Intelligence*, 4(1). <https://doi.org/10.36959/643/305>
41. The essential guide to AI training data-TELUS International The essential guide to AI training data. (n.d.).
42. Tournas, L. N., & Bowman, D. M. (2021). AI Insurance: Risk Management 2.0. *IEEE Technology and Society Magazine*, 40(4), 52–56. <https://doi.org/10.1109/MTS.2021.3123750>
43. Vijayaraj, V., Balamurugan, M., & Oberai, M. (2023). Machine learning approaches to identify the data types in big data environment: An overview. *The Scientific*
44. Wang, L., Huang, N., Hong, Y., Liu, L., Guo, X., & Chen, G. (n.d.). *Voice-based AI in Call Center Customer Service: A Natural Field Experiment*.
45. *Temper*, 14(03), 950–956. <https://doi.org/10.58414/scientifictemper.2023.14.3.60>
46. Zarifis, A., & Cheng, X. (2022). A model of trust in Fintech and trust in Insurtech: How Artificial Intelligence and the context influence it. *Journal of Behavioral and Experimental Finance*, 36. <https://doi.org/10.1016/j.jbef.2022.100739>