

POLITECNICO DI TORINO

Corso di Laurea Magistrale in Ingegneria Biomedica

Tesi di Laurea Magistrale

**Discriminazione tra soggetti sani e Parkinsoniani attraverso
l'elaborazione del segnale vocale rilevato tramite microfono in
aria e microfono a contatto**



Relatore

Prof. Alessio Carullo

Correlatore

Prof. Alberto Vallan

Candidato

Manuel Maggio

Anno accademico 2018-2019

A mio zio Fernando ed a mio nonno Aldo.

ABSTRACT

Il Parkinson è una malattia neurologica neurodegenerativa, i cui principali effetti sono i disturbi del movimento, quali rigidità, bradicinesia e tremore. Tuttavia, negli ultimi anni si sta intensificando l'analisi della correlazione tra la malattia del Parkinson e l'alterazione di specifici parametri relativi al segnale vocale. L'attività della presente tesi si colloca in questo ambito, ponendosi l'obiettivo di sviluppare algoritmi che permettano di distinguere una voce "sana" da una voce "Parkinsoniana". Gli algoritmi così individuati potrebbero essere applicati in futuro a dispositivi indossabili in grado di monitorare i soggetti Parkinsoniani e stimare l'efficacia delle terapie in atto. Durante la prima fase dell'attività di tesi è stato acquisito il segnale vocale di 22 soggetti Parkinsoniani presso l'ambulatorio Parkinson dell'ospedale Molinette di Torino. Ai soggetti monitorati è stato chiesto di seguire un protocollo della durata di circa cinque minuti che prevedeva la ripetizione di tre vocali sostenute /a/, la ripetizione dei fonemi /pa/ta/ka, la lettura di un brano foneticamente bilanciato ed un eloquio spontaneo. Durante la fase di acquisizione, ad ogni soggetto è stato fatto indossare un microfono a contatto a giro collo ed un microfono in aria a guancia. Durante la seconda fase, i segnali relativi alle vocali /a/ acquisite con il microfono a contatto ed alla lettura del brano acquisiti mediante i due diversi microfoni sono stati elaborati con uno script Matlab appositamente sviluppato. Tale script ha permesso di estrarre parametri nel dominio del tempo, quali *jitter* e *shimmer*, nel dominio della frequenza (*Harmonic-to-Noise-Ratio*: HNR) e nel dominio della *quefreny* (*Cepstral-Peak-Prominence-Smoothed*: CPPS). Una volta ottenuti questi parametri, è stata eseguita la terza fase, che è consistita nella selezione dei parametri più significativi e nell'identificazione di classificatori di voci sane e Parkinsoniane. Utilizzando la matrice di correlazione sono stati inizialmente esclusi i parametri ridondanti individuandone così un primo sottoinsieme, dal quale sono stati estratti i parametri rilevanti attraverso l'applicativo Matlab "*Classification Learner*". Per ciascun tipo di materiale vocale analizzato sono stati costruiti due modelli di classificatore basati sugli algoritmi *Coarse Tree* e regressione logistica. Per la vocale sostenuta acquisita tramite

microfono a contatto si è ottenuta un'accuratezza del 90% con l'algoritmo *Coarse Tree*, mentre l'accuratezza ottenuta con la regressione logistica è del 83%. Per la lettura del brano, attraverso il microfono a contatto l'algoritmo *Coarse Tree* ha permesso di avere un'accuratezza del 95% rispetto alla regressione logistica che ha mostrato un'accuratezza del 76%. Per la lettura del brano registrato attraverso il microfono in aria l'algoritmo *Coarse Tree* ha fornito un'accuratezza del 89%, a differenza della regressione logistica che ha mostrato un'accuratezza del 73%. Come ci si aspettava, avendo acquisito i segnali in un ambiente affetto da rumore il microfono a contatto mostra prestazioni superiori in termini di accuratezza rispetto al microfono in aria poiché poco influenzato dal rumore di fondo. Nella seconda parte di questo lavoro di tesi si è valutata l'efficacia di un percorso logopedico. Per far ciò, si sono acquisiti, con lo stesso protocollo e con la stessa catena di acquisizione utilizzata nella prima parte, un totale di sei soggetti presso il centro "amici Parkinsoniani Piemonte onlus" all'inizio e dopo due mesi di trattamento. Dall'analisi del segnale relativo alle vocali sostenute si è evinto, nella maggior parte dei casi, un miglioramento dei parametri di perturbazione, mentre per i parametri cepstrali si è notata una diminuzione in termini assoluti della kurtosis e della deviazione standard. I risultati relativi all'elaborazione del segnale vocale derivante dalla lettura del brano, sia per il microfono in aria che il microfono a contatto, hanno mostrato nella maggioranza dei soggetti in esame lo stesso tipo di andamento. Tali risultati, oltre che considerare un numero limitato di soggetti, sono evidentemente soggettivi in quanto il fine ultimo del percorso logopedico è stato quello di insegnare ai pazienti degli esercizi da svolgere autonomamente e con costanza in casa. Questo lavoro di tesi si pone quindi alla base di un progetto futuro, che per consolidare i risultati ottenuti richiederà l'estensione del *data set*.

Sommario

ABSTRACT.....	3
INDICE DELLE FIGURE.....	8
1. INTRODUZIONE	11
1.1 Apparato vocale e fonazione.....	11
1.2 Caratteristiche acustiche del segnale vocale	12
1.3 Circuiteria alla base del Parkinson.....	14
1.4 Prognosi e terapie per la malattia di Parkinson.....	16
1.5 Disturbi del linguaggio nella malattia di Parkinson.....	18
2. STATO DELL'ARTE.....	19
2.1 Parametri di perturbazione del segnale vocale	19
2.2 Misure spettrali e cepstrali	20
3. MATERIALI E METODI	23
3.1 Protocollo utilizzato	23
3.2 Catena di acquisizione utilizzata.....	24
3.3 Soggetti acquisiti.....	26
3.3.1 Ripetizione delle tre vocali /a/ sostenute	26
3.3.2 Lettura del brano foneticamente bilanciato registrato con il microfono a contatto.....	27
3.3.3 Lettura del brano foneticamente bilanciato registrato con il microfono in aria	28
3.4 Pre-elaborazione dei dati	28
3.5 Algoritmo di valutazione del frame in esame	29
3.6 Parametri estratti	31
3.6.1 Parametri estratti dalla vocale /a/ sostenuta	31
3.6.2 Parametri estratti dalla lettura del brano	35

3.7 Selezione dei parametri.....	39
3.8 Coarse tree	41
3.9 Regressione logistica	44
4. RISULTATI.....	47
4.1 Vocale /a/ sostenuta	47
4.1.1 Matrice di correlazione	47
4.1.2 Coarse tree	49
4.1.3 Regressione logistica	53
4.2 Lettura del brano (microfono a contatto).....	63
4.2.1 Matrice di correlazione	63
4.2.2 Coarse tree	64
4.2.3 Regressione logistica	68
4.3 Lettura del brano (microfono in aria)	72
4.3.1 Matrice di correlazione	72
4.3.2 Coarse tree	73
4.3.3 Regressione logistica	75
5. VALUTAZIONE CORSO LOGOPEDICO	79
5.1 Analisi della vocale /a/	80
5.1.1 Variazione dei parametri di perturbazione.....	82
5.1.1 Variazione dei parametri cepstrali	86
5.2 Analisi della lettura del brano	90
5.2.1 Microfono a contatto.....	90
5.2.2 Microfono in aria	93
CONCLUSIONI	96
APPENDICE	99
Brano foneticamente bilanciato	99
BIBLIOGRAFIA	100

Testo.....	100
Figure	103
RINGRAZIAMENTI.....	104

INDICE DELLE FIGURE

Figura 1 Segnale glottico e verbale [1]	13
Figura 2 Neuroni dopaminergici [2]	15
Figura 3 Circuiteria alla base del controllo del movimento [3]	15
Figura 4 Jitter e Shimmer.....	19
Figura 5 Cepstrum	21
Figura 6 Cepstral Peak Prominence Smoothed.....	22
Figura 7 Esempio di registrazione	24
Figura 8 Microfono in aria.....	25
Figura 9 Microfono a contatto	25
Figura 10 Catena di acquisizione completa	26
Figura 11 Flow-Chart.....	30
Figura 12 Distribuzione del CPPS per la vocale /a/ (microfono a contatto).....	37
Figura 13 Distribuzione del CPPS per la lettura del brano (microfono a contatto)	38
Figura 14 Distribuzione del CPPS per la lettura del brano (microfono in aria)	38
Figura 15 Esempio di Parallel Coordinates Plot.....	40
Figura 16 Esempio di albero decisionale	42
Figura 17 Funzione sigmoidale.....	44
Figura 18 Curva ROC [4]	46
Figura 19 Albero decisionale relativo alla vocale /a/	49
Figura 20 Flow-chart relativo alla vocale /a/	50
Figura 21 Scelta del possibile cut-off per il modello di regressione logistica ricavato dall'analisi del segnale vocale relativo alle vocali sostenute (microfono a contatto)	54
Figura 22 Curva ROC ricavata dal modello di regressione logistica ricavato dall'analisi del segnale vocale relativo alle vocali sostenute (microfono a contatto)	55
Figura 23 Grafico di separazione relativo al modello di regressione logistica ricavato dall'analisi delle singole vocali /a/.....	56

Figura 24 Scelta del possibile cut-off per il nuovo modello di regressione logistica ricavato dall'analisi del segnale vocale relativo alle vocali sostenute (microfono a contatto)	58
Figura 25 Curva ROC ricavata dal nuovo modello di regressione logistica ricavato dall'analisi del segnale vocale relativo alle vocali sostenute (microfono a contatto)	59
Figura 26 Grafico di separazione relativo al nuovo modello di regressione logistica ricavato dall'analisi delle singole vocali /a/	60
Figura 27 Grafico di separazione, relativo al nuovo modello di regressione logistica, ricavato dall'analisi della media delle vocali	62
Figura 28 Grafico di separazione complessivo relativo al nuovo modello di regressione logistica	62
Figura 29 Albero Decisionale relativo alla lettura del brano (microfono a contatto)	65
Figura 30 Flow-chart relativo alla lettura del brano (microfono a contatto)	66
Figura 31 Scelta del possibile cut-off per il modello di regressione logistica ricavato dall'analisi del segnale vocale relativo alla lettura del brano (microfono a contatto)	69
Figura 32 Curva ROC ricavata dal modello di regressione logistica ricavato dall'analisi del segnale vocale relativo alla lettura del brano (microfono a contatto)	70
Figura 33 Grafico di separazione relativo al modello di regressione logistica ricavato dall'analisi della lettura del brano (microfono a contatto)	71
Figura 34 Albero Decisionale relativo alla lettura del brano (microfono in aria)	73
Figura 35 Flow-chart relativo alla lettura del brano (microfono in aria)	74
Figura 36 Scelta del possibile cut-off per il modello di regressione logistica ricavato dall'analisi del segnale vocale relativo alla lettura del brano (microfono in aria)	76
Figura 37 Curva ROC ricavata dal modello di regressione logistica ricavato dall'analisi del segnale vocale relativo alla lettura del brano (microfono in aria)	77
Figura 38 Grafico di separazione relativo al modello di regressione logistica ricavato dall'analisi della lettura del brano (microfono in aria)	79
Figura 39 Variazione Jitt	83
Figura 40 Variazione PPQ	84
Figura 41 Variazione vF_0	84

Figura 42 Variazione Shim	85
Figura 43 Variazione vAm	85
Figura 44 Distribuzioni relative al CPPS ricavate dall'analisi del segnale vocale delle vocali /a/ (microfono a contatto)	87
Figura 45 Variazione CPPS _{mean}	88
Figura 46 Variazione CPPS _{std.}	88
Figura 47 Variazione CPPS _{skewness}	89
Figura 48 Variazione CPPS _{kurtosis}	89
Figura 49 Distribuzioni relative al CPPS ricavate dall'analisi del segnale vocale derivante dalla lettura del brano (microfono a contatto)	92
Figura 50 Distribuzioni relative al CPPS ricavate dall'analisi del segnale vocale derivante dalla lettura del brano (microfono in aria)	95

1. INTRODUZIONE

1.1 Apparato vocale e fonazione

L'apparato vocale umano include diversi "componenti" come:

- I polmoni
- I bronchi e la trachea
- La laringe, all'interno della quale si trovano le corde vocali
- La cavità orale
- La cavità nasale

Tutti questi componenti lavorano insieme portando alla fonazione. I polmoni fungono da "generatore", fornendo il flusso d'aria che durante la conversazione si sposta attraverso la trachea giungendo alla laringe. Nella parte centrale della laringe, detta glottide, sono presenti le pliche vocali che risultano essere rilassate durante la normale respirazione, tendendo invece ad unirsi tra loro durante la fonazione. Aprendo e chiudendo rapidamente la glottide durante la fonazione, le corde vocali rilasciano l'aria dai polmoni in un flusso vibrante, producendo delle vibrazioni acustiche, i suoni, che sono le materie prime per le parole stesse. Questa prima sezione tra i polmoni e la glottide è definita come "tratto della glottide" e rappresenta la fonte della fonazione. I suoni sono modellati dal resto dell'apparato vocale, il "tratto vocale", che inizia dopo le corde vocali fino alle labbra, comprese le cavità nasali. Questa sezione funge da "risonatore", un filtro complesso che altera i suoni emessi dalla glottide, amplificando alcune frequenze ed attenuandone altre. Il palato molle, la lingua, i denti, le labbra e le altre parti della bocca modulano il suono variando la loro posizione, in modo che possano essere prodotte vocali e consonanti. Si può quindi concludere che le corde vocali svolgano un ruolo centrale nella produzione della voce in quanto costituiscono il componente "vibratore" dell'apparato fonetico. La frequenza di vibrazione delle corde vocali dipende da parametri meccanici. Infatti, essa è direttamente proporzionale alle caratteristiche elastiche delle corde vocali ed inversamente proporzionale alla loro massa e lunghezza; ciò spiega il differente range frequenziale che caratterizza un uomo ed

una donna. La lunghezza delle corde vocali, difatti, in un uomo adulto è di 17-25 mm, mentre in una donna adulta risulta essere di 12-17 mm.

1.2 Caratteristiche acustiche del segnale vocale

La fonazione, come è stato precedentemente spiegato, è la produzione di un segnale acustico da una sorgente rappresentata essenzialmente dalle corde vocali. Il segnale glottale viene visto, dal punto di vista fisico in condizioni fisiologiche e senza filtri, come un segnale quasi periodico. Essendo un segnale complesso, può essere assimilato ad una somma algebrica di una serie di segnali sinusoidali, detti componenti spettrali, caratterizzate da una propria frequenza, intensità e fase. Se il segnale è periodico e complesso con periodo fondamentale T_0 , le componenti, definite anche armoniche, hanno frequenze che sono multiple intere della frequenza fondamentale $F_0 = 1 / T_0$. La frequenza fondamentale, per il segnale laringeo, coincide con la frequenza di apertura e chiusura della glottide, cioè la frequenza di vibrazione delle corde vocali [1]. Il termine "quasi" suggerisce che le caratteristiche di frequenza ed ampiezza del segnale possono cambiare nel tempo. Potrebbero verificarsi perturbazioni a breve termine, osservabili da un periodo a quello successivo; per la frequenza fondamentale, queste variazioni possono essere dell'ordine di ± 25 Hz e sono necessarie per fornire naturalezza al parlato. Tali variazioni di ampiezza e frequenza fondamentale possono essere utilizzate come parametri utili a distinguere una voce patologica da una voce sana. La frequenza fondamentale del segnale vocale ha un *range* che risulta essere dipendente da diversi fattori, come il genere e l'età. L'estensione vocale di bambini, donne adulte e uomini adulti è differente; difatti è generalmente compresa rispettivamente tra: (225-440) Hz, (175-245) Hz e (105-160) Hz [1]. Una volta uscito dalle corde vocali, il segnale della glottide viene sottoposto all'azione filtrante del tratto vocale, che lo trasforma nel segnale vocale. Tuttavia, F_0 , la prima armonica del segnale glottale, rimane la stessa anche nel segnale vocale; viene modificata solo l'ampiezza delle componenti spettrali. La figura 1 consente di confrontare le caratteristiche del segnale glottale e quelle del corrispondente segnale vocale emesso a livello delle labbra. Il primo segnale viene acquisito con un microfono a contatto mentre il secondo segnale è stato acquisito con un microfono in aria. Possiamo inoltre notare

la differenza all'interno dello spettro dovuta all'effetto del filtro che, assume quindi un ruolo di funzione di trasferimento.

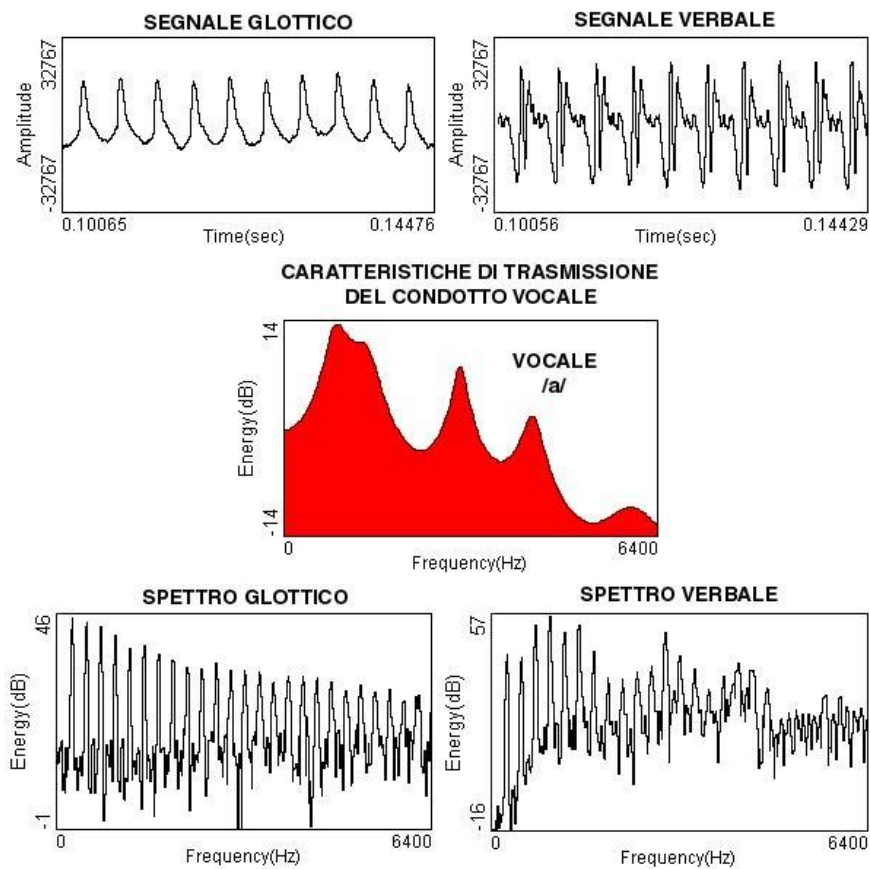


Figura 1 Segnale glottico e verbale [1]

1.3 Circuiteria alla base del Parkinson

Il Parkinson è una malattia neurodegenerativa che presenta una evoluzione molto lenta ma progressiva e fa parte di un gruppo di patologie definite “Disordini del Movimento” poiché coinvolge principalmente il controllo del movimento e dell’equilibrio. Tra i "Disordini del Movimento", la malattia di Parkinson è purtroppo la più frequente. Alla base del controllo del movimento, esiste una circuiteria caratterizzata da diverse aree cerebrali tra cui la “Pars Compacta Substantia Nigra” che è un’area ricca di neuroni serotoninergici. Tali neuroni presentano una lunga innervazione che termina in un’area cerebrale definita “striato” in cui sono presenti neuroni chiamati D1. Tali neuroni sono eccitati quando è rilasciata dopamina dai neuroni serotoninergici, ed una volta eccitati vanno ad attivare una via specifica di comunicazione, detta via diretta, che è una via facilitatoria per il movimento. In un individuo sano, i neuroni dopaminergici scaricano dopamina al livello dello striato attivando i neuroni D1 che metteranno in moto una via diretta terminante sulla corteccia motoria permettendo di ottenere il movimento. Nel caso di soggetti Parkinsoniani invece, non viene rilasciata sufficiente dopamina e di conseguenza compariranno dei disturbi motori come il rallentamento e la bradicinesia. I meccanismi e la circuiteria sopra descritti possono essere visualizzati nelle figure 2 e 3 riportate qui di seguito.

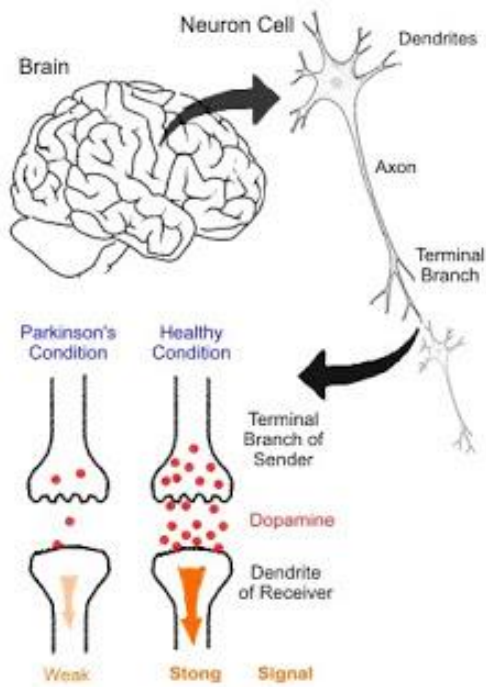


Figura 2 Neuroni dopaminergici [2]

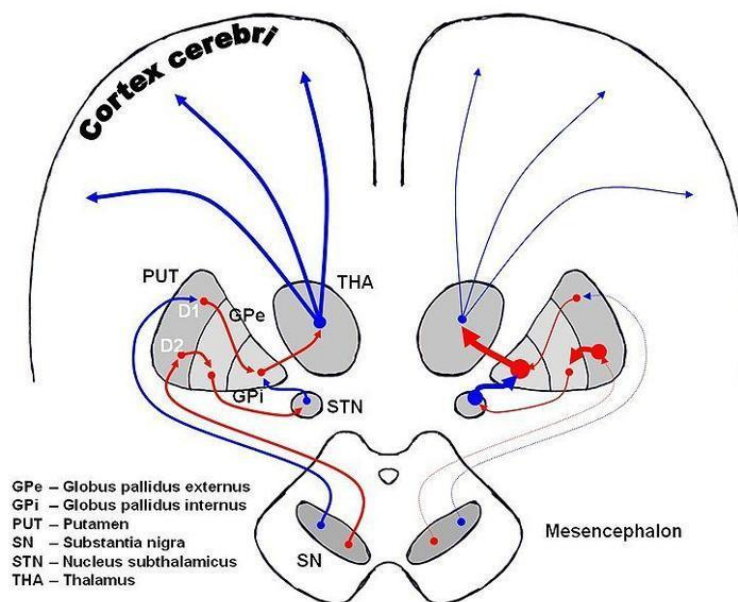


Figura 3 Circuitria alla base del controllo del movimento [3]

1.4 Prognosi e terapie per la malattia di Parkinson

La diminuzione delle prestazioni vocali nei pazienti affetti dalla malattia di Parkinson è stata dimostrata in diversi studi che, oltre ad evidenziare i comuni sintomi della malattia come il tremore a riposo la rigidità muscolare e la bradicinesia, hanno dato importanza ai disturbi relativi al segnale vocale. Generalmente i soggetti caratterizzati dal morbo di Parkinson (PD) presentano un linguaggio inintelligibile e balbuziente. In letteratura tuttavia, non esistono valutazioni della malattia basate solo sul linguaggio ma esistono test clinici che valutano la prognosi generale. Una prima scala di valutazione della malattia di Parkinson è la *Unified Parkinson's Disease Rating Scale* (UPDRS), tale scala è stata originariamente sviluppata nel 1980 [2], ed è diventata la scala di valutazione clinica più utilizzata [3]. Tale scala di valutazione oltre che la più utilizzata è anche la più completa, in quanto valuta gli aspetti cognitivi ed affettivi oltre che le normali attività di vita quotidiana; essa è, infatti, suddivisa in diverse sezioni ognuna delle quali valuta aspetti differenti. Si parte dalla valutazione dello stato mentale del paziente, dal suo comportamento e dal suo umore; successivamente si ha una autovalutazione delle attività quotidiane; un punto fondamentale risiede nell'andare a valutare le abilità motorie, ed infine si vanno a valutare quali siano le complicanze motorie possibili [4]. La sezione uno della parte tre è utilizzata come una misura globale della valutazione della comunicazione. Comprende dizione, volume, intelligibilità ed espressione. Per tale motivo, in questo lavoro di tesi, tale sezione sarà utilizzata come indice della gravità del sistema vocale. Nel 2001, la *Movement Disorder Society* (MDS) ha sponsorizzato una critica all'UPDRS, identificando una serie di ambiguità e debolezze. Si è quindi sviluppata una nuova versione che ha mantenuto i punti di forza della scala originale risolvendone i problemi. Tale nuova versione, denominata MDS-UPDRS, ha superato con successo il test clinimetrico [5]. Anche la MDS-UPDRS è suddivisa in quattro parti: dove la prima parte permette di valutare le esperienze non motorie della vita quotidiana; la seconda parte valuta le esperienze motorie della vita quotidiana, la terza parte effettua una valutazione motoria e la quarta ed ultima parte valuta anche in questo caso le possibili complicanze motorie. Un'altra scala diffusamente utilizzata sempre per valutare la prognosi della patologia è la scala Hoehn Yahr (HY). Tale scala di

valutazione, utilizzata da più di trenta anni, appartiene all'era pre-levodopa e presenta un range di valutazione che va da uno a cinque. Con l'utilizzo di tale sistema di valutazione, si associa al paziente un valore pari ad uno quando si ha un "solo" coinvolgimento unilaterale, un valore pari a due quando si ha un coinvolgimento bilaterale senza la compromissione dell'equilibrio, un valore pari a tre se il paziente inizia ad avere una prima instabilità posturale, un valore pari a quattro quando si perde l'indipendenza fisica ed infine un valore pari a cinque quando il paziente risulta essere allettato. Queste due scale di valutazione sono le più usate ma, essendo il Parkinson una malattia neuro-degenerativa, una classificazione della stessa in base alla sua durata (DD) può essere applicata come una misura della variazione dei sintomi nel tempo. Per quanto riguarda una possibile cura, purtroppo non ne esiste ancora una definitiva, sebbene siano presenti diversi trattamenti in grado di migliorare la qualità della vita del paziente. Tra questi è applicata una terapia farmacologica che prevede la somministrazione di agonisti o antagonisti della dopamina con lo scopo di regolare i livelli del neurotrasmettitore. Oltre all'aiuto farmacologico, negli ultimi anni, stanno nascendo delle nuove tecniche che sono tuttavia ancora in fase di sperimentazione come la *Deep Brain Stimulation (DBS)* [6]. La DBS consiste nel somministrare un segnale elettrico ad alta frequenza attraverso elettrodi impiantati chirurgicamente nel subtalamo; tuttavia, i benefici di questa tecnica devono essere bilanciati con gli effetti avversi, come il dolore cronico e maggiore depressione. Un'altra tecnica che ha visto l'utilizzo anche in soggetti affetti dalla malattia di Parkinson è la *Transcranial Magnetic Stimulation (TMS)*. Tale tecnica prevede l'applicazione di una corrente alternata circolante all'interno di un *coil* posto in prossimità dello scalpo del paziente con l'intento di produrre un campo magnetico indotto che penetrerà lo scalpo del soggetto in esame andando a produrre una corrente indotta nel tessuto cerebrale. Tale corrente indotta presenterà un verso opposto a quella che l'ha generata e ha lo scopo di indurre una corrente trans-membranale. La TMS tuttavia, non è stata introdotta come cura per il morbo di Parkinson ma per far fronte alle complicità della stessa, cercando di ridurre gli effetti collaterali dovuti ad un uso sistemico di farmaci.

1.5 Disturbi del linguaggio nella malattia di Parkinson

Diversi studi hanno indagato sulla correlazione presente tra la malattia di Parkinson e i disturbi associati al linguaggio motorio. Da tali studi è stato evidenziato come fino al 90% dei pazienti affetti dalla patologia di Parkinson sviluppano un disturbo del linguaggio motorio distintivo descritto nel 1969 da Darley et al. ed indicato come disartria ipocinetica (HD) [7]. L'HD si manifesta in diverse aree relative alla produzione del linguaggio umano, in particolare nelle aree di articolazione, fonazione, prosodia e fluidità del linguaggio [8]. “L'HD è caratterizzato da rigidità, bradicinesia e ridotto controllo muscolare della laringe, degli organi articolatori e di altri meccanismi di supporto fisiologico della produzione del linguaggio umano” [9]. Nei pazienti affetti da Parkinson sono stati osservati diversi difetti, quali, l'aumento del rumore acustico [10] inteso come riduzione dell'intensità della voce [11], l'aumento della nasalità vocale [12], l'imprecisione nell'articolazione delle consonanti [13], l'introduzione involontaria di pause [14], ed un'introduzione di improvvise decelerazioni o accelerazioni nel discorso [15]. Sebbene vi siano tutte queste evidenze, l'analisi acustica della voce e della parola non è stata ancora inclusa nei criteri diagnostici per la malattia di Parkinson [16]. Altri studi hanno analizzato il segnale vocale dei pazienti con PD mettendoli a confronto con gruppi di controllo, nella maggioranza dei casi i soggetti Parkinsoniani rispetto ai soggetti sani hanno mostrato una diminuzione delle prestazioni vocali confermando così il risultato atteso. Nelle analisi vocali della produzione di una vocale sostenuta, sono state osservate frequenze fondamentali più elevate rispetto ai soggetti appartenenti a gruppo di controllo [17], come anche maggiori sono risultati essere *jitter* e *shimmer* indicando quindi una maggiore instabilità del sistema fonatorio [18].

2. STATO DELL'ARTE

2.1 Parametri di perturbazione del segnale vocale

Negli ultimi decenni, le misure acustiche hanno assunto un'importanza fondamentale, perché sono non invasive, relativamente a basso costo, *real time* e facili da applicare [19]. I parametri che possono essere estratti dal segnale vocale risultano essere differenti e possono essere riferiti al dominio temporale e/o al dominio della frequenza. Da una vocale sostenuta è possibile calcolare la frequenza fondamentale (f_0), il *jitter*, lo *shimmer* ed il rapporto tra la componente armonica e la componente rumorosa (*harmonic-to-noise ratio* HNR). Il *jitter* rappresenta la perturbazione a breve termine della frequenza fondamentale, mentre lo *shimmer* rileva una irregolarità in termini di ampiezza [20] [21] [22], tali parametri sono illustrati in figura 4. Il *jitter* è influenzato maggiormente dalla mancanza di vibrazione delle corde vocali, perciò pazienti patologici hanno una percentuale più elevata di *jitter*. Le lesioni di massa delle corde vocali invece incidono invece sullo *shimmer* che è correlato con la presenza di rumore e del respiro. [22]

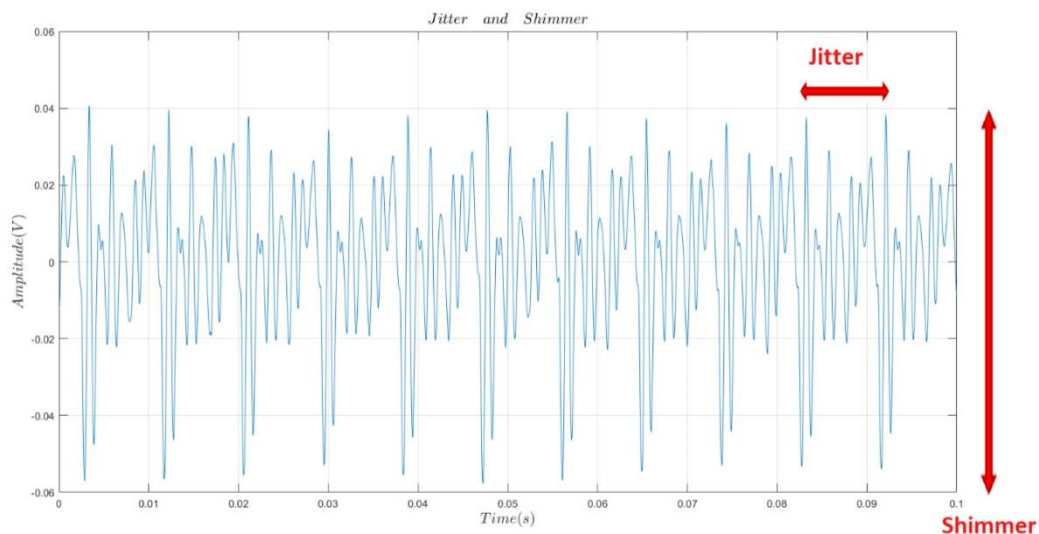


Figura 4 Jitter e Shimmer.

Molti studi hanno evidenziato come un alto grado di disfonia sia correlato ad un alto livello di perturbazione, per tale motivo sono stati implementati dei sistemi efficaci, rapidi ed affidabili, che siano in grado di valutare tali parametri di perturbazione come il *jitter* e lo *shimmer*, ma anche tutta una serie di parametri da essi ricavati. Il parametro definito HNR, precedentemente esposto, può essere stimato sia nel dominio del tempo che nel dominio della frequenza; valori bassi identificano un'alta componente aperiodica all'interno del segnale vocale al contrario di valori elevati che ne sottolineano una bassa presenza. Questo parametro è stato utilizzato, lavorando nel dominio del tempo, da Yumoto et al. [23]; tuttavia Qi e Hillman ne hanno dimostrato un calcolo più semplice nel dominio della frequenza [24]. I parametri sopra menzionati devono però essere estratti da vocali continue prodotte con passo e volume stabili, questo perché ogni cambiamento significativo sarebbe altrimenti letto come un incremento della perturbazione vocale. Tuttavia, questa non è l'unica limitazione che riguarda l'uso delle misure di perturbazione tradizionali. Infatti, il limite maggiore consiste nel fatto che tali parametri dipendono fortemente dalla corretta individuazione della frequenza fondamentale essendo appunto da essa ricavati; piccoli errori nella stima di f_0 affliggono quindi la misura di *jitter*, *shimmer*, HNR e di tutti gli altri parametri da questi derivanti [25].

2.2 Misure spettrali e cepstrali

Per quanto precedentemente esposto, con lo scopo di fornire un'analisi parametrica completa anche in un eloquio libero o attraverso la lettura di un brano, diversi studi si sono orientati verso approcci basati sul dominio della frequenza e della *quefreny*. In particolare, le misure spettrali e cepstrali non richiedono l'individuazione dello pseudo-periodo, quindi sono più accurate per le voci Parkinsoniane o disfoniche in generale. Il metodo di stima dello spettro di potenza consente di ottenere informazioni importanti sulla voce, ma la ricerca si sta muovendo verso un nuovo approccio importante basato sui parametri cepstrali. Nel 1963 Borget et al. [25] definì il concetto di cepstrum come lo spettro di uno spettro, in particolare come "lo spettro di potenza logaritmico dello spettro logaritmico del segnale vocale". Il *cepstrum* viene quindi definito nel dominio del tempo, per tal motivo i termini

spettro e frequenza possono essere sostituiti dalle corrispettive parole: *cepstrum* e *quefrequency*. In particolare, il *cepstrum* mostra il picco di energia di ogni componente di frequenza, ed i segnali periodici mostrano una struttura armonica definita con un picco visibile corrispondente al periodo fondamentale, come mostrato nella figura 5. L'ampiezza del picco, definisce il parametro CP (*Cepstral Peak*) e si trova approssimativamente ad una frequenza compresa tra 60 Hz e 300 Hz, in un intervallo temporale compreso quindi tra 3 ms e 16 ms. Questo parametro non è influenzato solo dall'organizzazione armonica ma anche dall'energia complessiva del segnale e conseguentemente un fattore molto importante è attribuito al rumore. Dal *cepstrum* è possibile andare a ricavare due parametri: la prominente del picco *cepstrale* (*Cepstral Peak Prominence*) CPP e la sua versione smussata CPPS (*Cepstral Peak Prominence Smoothed*). Il CPP è la misura in dB dell'ampiezza del picco *cepstrale* rispetto al valore della regressione lineare alla stessa frequenza, mentre per calcolare il CPPS sono necessarie due precedenti operazioni di smoothing [26]. Il CPPS non è calcolato per ogni singolo cepstrum, ma una finestra mobile si sposta lungo il segnale temporale mantenendo fisso il suo *overlap*, in questo modo si definisce un CPPS per ogni singolo frame e l'output dell'algoritmo risulta essere composto da diversi valori di CPPS che permettono di creare una distribuzione. Lo scopo di tale distribuzione consiste nel poter estrapolare successivamente degli indici statistici di tendenza come la media, la mediana e la moda, ed altri indici di dispersione oltre che di forma.

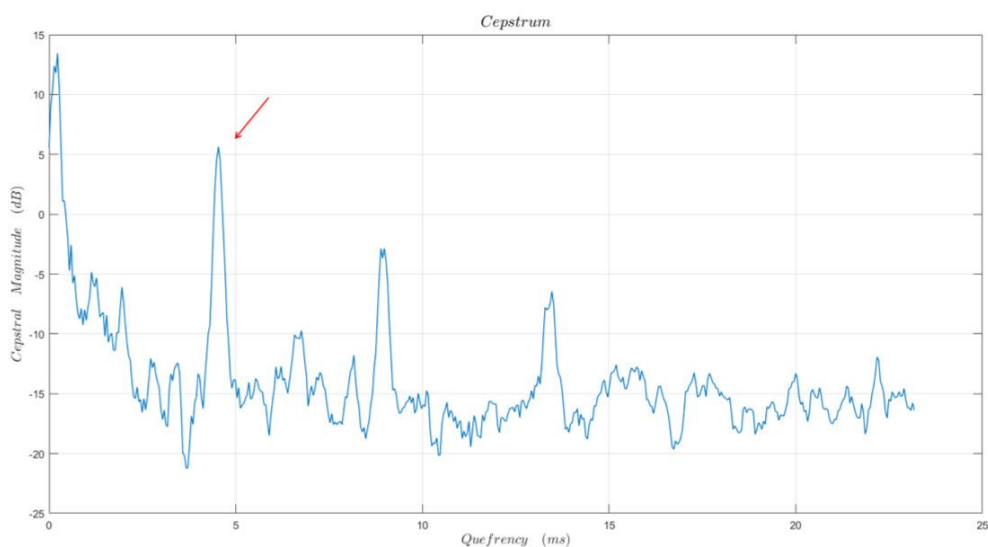


Figura 5 Cepstrum

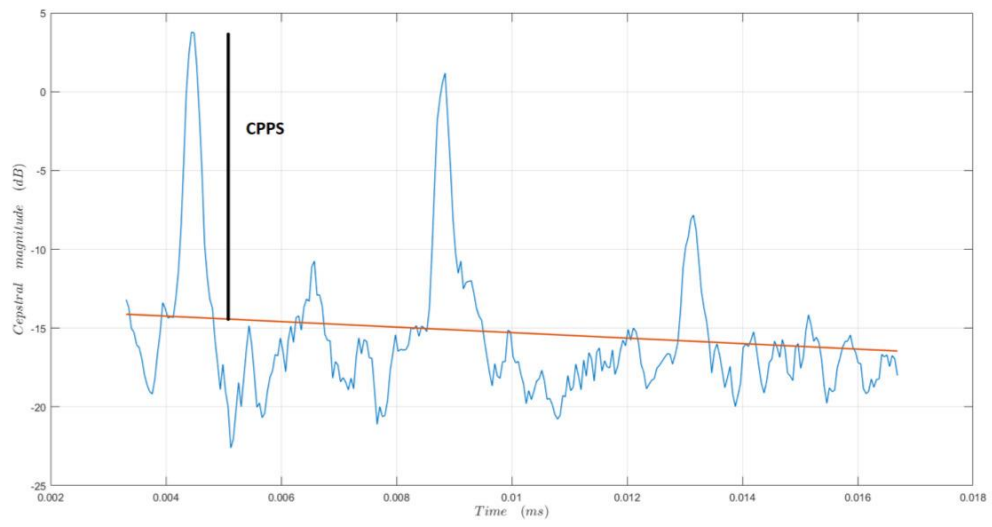


Figura 6 Cepstral Peak Prominence Smoothed

3. MATERIALI E METODI

3.1 Protocollo utilizzato

Per la prima parte di questo lavoro di tesi le registrazioni del segnale vocale sono state effettuate presso il centro Parkinson Città della Salute dell'ospedale Molinette di Torino. Durante la visita neurologica di controllo che i pazienti dovevano effettuare ed ai quali è stato assegnato un valore di UPDRS e di HY da parte del neurologo, sono state eseguite le registrazioni del segnale vocale seguendo un protocollo ben definito. Tale protocollo è suddiviso in quattro parti:

- La prima parte è basata sulla ripetizione della vocale /a/ in maniera sostenuta (per circa dieci secondi) su un tono e volume confortevole, per un totale di tre volte, intervallate dalla fisiologica inspirazione;
- La seconda parte ha previsto la ripetizione dei fonemi /pa/ta/ka, sempre su un tono e volume confortevole per un numero di volte pari a dieci ad una velocità che sia la massima possibile per il paziente in esame;
- La terza parte è consistita nella lettura di un brano foneticamente bilanciato senza interruzione dall'inizio alla fine per la durata di circa un minuto;
- L'ultima parte è invece basata su di un eloquio spontaneo che doveva avere durata di circa un minuto; in questo caso il paziente poteva parlare liberamente senza avere quindi dei vincoli riguardanti l'argomento;

3.2 Catena di acquisizione utilizzata

Durante la fase di registrazione per ogni fase del protocollo, i pazienti hanno indossato due microfoni diversi: un microfono in aria e un microfono a contatto. Il primo registra il segnale vocale dopo il suo passaggio attraverso il tratto vocale, che funge da filtro; di conseguenza, il segnale ottenuto è più complesso; tuttavia, questo microfono acquisisce anche il rumore esterno oltre che il segnale utile. Il microfono a contatto invece rileva le vibrazioni delle corde vocali, fornendo in uscita un segnale più simile a quello della glottide influenzato dal rumore ambientale in maniera trascurabile; tuttavia, la registrazione è meno chiara all'orecchio di quella ottenuta con il microfono in aria.



Figura 7 Esempio di registrazione

Il microfono in aria è un microfono a guancia (MIPRO MU 55-HR), con una banda compresa tra 40 Hz e 20 kHz ed una piattezza di banda di ± 3 dB. Tale microfono è stato posizionato ad una distanza di circa 2,5 cm dalle labbra del paziente (figure 7 e 8).



Figura 8 Microfono in aria

Il microfono a contatto è un microfono piezoelettrico, HX-505-1-1, HKKK, 406, PLant 1, Jiadind Science Park, Dalang, Longhua New Dist., Shenzhen, Guangdong, China. (Figura 9)



Figura 9 Microfono a contatto

Entrambi i microfoni sono stati collegati all'ingresso stereo di un registratore portatile (Edirol Roland R-09HR) tramite un cavo divisibile a Y presentate due entrate, in maniera tale da ottenere due segnali sincronizzati. (Figura 10)



Figura 10 Catena di acquisizione completa

Le registrazioni sono state eseguite ad una frequenza di campionamento pari a 44100Hz. La risoluzione è stata per tutti i casi fissata a 16 bit.

3.3 Soggetti acquisiti

Le registrazioni effettuate presso l'ambulatorio Parkinson sono state eseguite su di un numero di ventidue pazienti. Tuttavia, a causa di differenti condizioni personali o a causa di eventuali disagi riscontrati nell'indossare i microfoni, specie quello a contatto, non tutti i soggetti sono stati in grado di eseguire ogni singola parte del protocollo. Per tale ragione, è stata effettuata una sotto-classificazione del numero di segnali vocali acquisiti in base alla specifica fase del protocollo in esame ed al tipo di microfono utilizzato.

3.3.1 Ripetizione delle tre vocali /a/ sostenute

Per quanto riguarda la prima fase del protocollo, che ha previsto tre ripetizioni della vocale /a/ sostenuta, sono stati archiviati un totale di quarantadue segnali vocali. La metà derivanti dai pazienti Parkinsoniani acquisiti all'ambulatorio precedentemente descritto ed i rimanenti invece appartenenti ad un gruppo di controllo registrato in anni precedenti. I quarantadue soggetti, come era facile aspettarsi, non mostrano la stessa età e sono anche differenti in termini di genere. La tabella sottostante

permette di evidenziarne le differenze in termini di età media, deviazione tipo e genere.

	Parkinsoniani	Sani
Numero totale	21	21
Numero donne	7	8
Numero uomini	14	13
Età media	67	23
Dev. Std.	13	15

Tabella 1 Numero, età e genere dei soggetti che hanno sostenuto le tre vocali /a/ (microfono a contatto)

3.3.2 Lettura del brano foneticamente bilanciato registrato con il microfono a contatto

Per quanto riguarda la terza fase del protocollo, che ha previsto la lettura di un brano foneticamente bilanciato riportato nell'appendice, sono stati acquisiti un totale di trentotto segnali vocali derivanti rispettivamente da diciannove soggetti sani ed altrettanti pazienti Parkinsoniani. Le specifiche dei due gruppi, in termini di età media, genere e deviazione tipo sono riportate nella tabella sottostante.

	Parkinsoniani	Sani
Numero totale	19	19
Numero donne	7	8
Numero uomini	12	11
Età media	65	32
Dev. Std.	13	15

Tabella 2 Numero, età e genere dei soggetti che hanno letto il brano (microfono a contatto)

3.3.3 Lettura del brano foneticamente bilanciato registrato con il microfono in aria

L'analisi della terza fase del protocollo è proseguita con l'utilizzo del microfono in aria. In questo caso, è stato possibile andare ad acquisire, grazie alla minore "invasività" del microfono stesso rispetto a quello a contatto, tutti e ventidue i pazienti. Perciò, il numero dei segnali vocali registrati, dato dalla somma dei segnali riguardanti i pazienti sani ed i pazienti Parkinsoniani, è risultato essere pari a quarantaquattro. Anche per questa ultima analisi, sono riportate le specifiche dei due gruppi in termini di età media, genere e deviazione tipo.

	Parkinsoniani	Sani
Numero totale	22	22
Numero donne	8	8
Numero uomini	14	14
Età media	68	41
Dev. Std.	13	14

Tabella 3 Numero, età e genere dei soggetti che hanno letto il brano (microfono in aria)

3.4 Pre-elaborazione dei dati

Dopo la registrazione del segnale vocale, i dati sono stati scaricati e salvati in un Personal Computer come file audio ".wav", tramite schede SD. Successivamente, ogni registrazione è stata rinominata in base al codice di identificazione del paziente e suddivisa in diversi file audio utilizzando il software Audacity 2.2.2:

- Vocali /a/ sostenute: sono tre per ciascuna registrazione (A1, A2 e A3) della durata di circa 10 secondi, ad eccezione di alcuni soggetti che non sono stati in grado di mantenere la vocale per un tempo così lungo. Da tali file ".wav" si sono ottenuti, tagliando la parte iniziale e finale, i quattro secondi centrali.
- Lettura del brano: ogni registrazione è stata tagliata partendo dall'inizio fino ad arrivare alla parola "sanguisuga", in questo modo si sono analizzate solo

le nove prime nove frasi del brano permettendo di ridurre il tempo di elaborazione del segnale stesso.

Una volta che tutti i file audio sono stati selezionati, raccolti ed organizzati in base al tipo di microfono ed in base alla specifica fase del protocollo, la fase successiva è stata l'elaborazione del segnale vocale di ciascun paziente per ciascuna fase del protocollo.

3.5 Algoritmo di valutazione del frame in esame

Tutti i parametri, ad eccezione di CPPS, dipendono dalla stima della frequenza fondamentale e dell'HNR [27], perciò una buona parte del lavoro di tesi ha riguardato l'implementazione di un algoritmo che fosse in grado di stimare proprio la frequenza fondamentale, misurata con il metodo dell'autocorrelazione, e l'HNR. In particolare, per il significato assunto dal parametro HNR, quando il contenuto di rumore risulta essere maggiore del contenuto relativo alla parte armonica del segnale si ha un valore di HNR minore di zero decibel (0 dB) ed in questo caso la stima della frequenza fondamentale non risulta essere consistente. È noto inoltre che la produzione di consonanti sorde, come / s / e / f /, non provoca la vibrazione delle corde vocali, perché il suono è prodotto dall'articolazione del flusso d'aria della bocca perciò misurare l'HNR di tali consonanti darà un valore negativo e i parametri estratti possono essere poco significativi; la stessa situazione può verificarsi con le vocali prodotte da voci patologiche incluse quelle appartenenti quelli ai soggetti Parkinsoniani. Inoltre, è necessario fissare una soglia massima nello spostamento di frequenza per considerare valido il frame misurato e per ridurre gli errori causati dai salti di ottava. Il diagramma di flusso in figura 11 descrive il funzionamento dell'algoritmo implementato per valutare quando un frame in esame è stato ritenuto valido o meno.

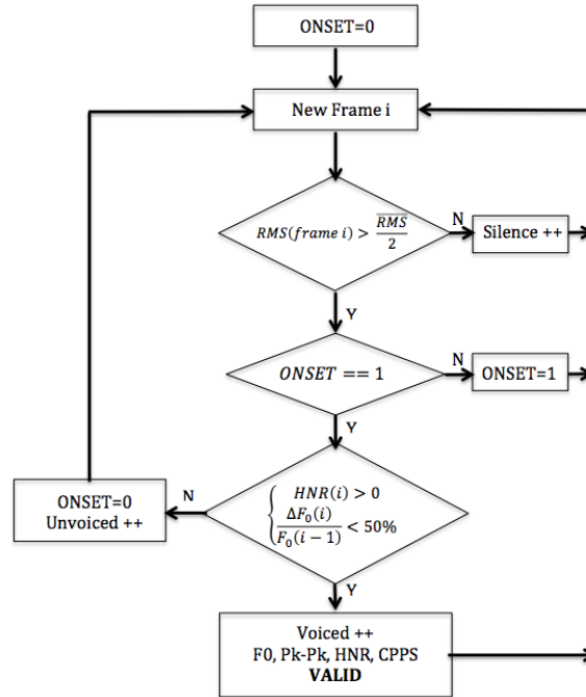


Figura 11 Flow-Chart

Inizialmente la variabile “onset”, che è utilizzata come discriminazione transitoria tra un frame valido ed uno non valido, è stata inizializzata a zero. Come primo passo, si è verificato che il *root mean square* (*RMS*) del frame in esame (*i*) fosse maggiore alla metà dell’*RMS* dell’intero segnale analizzato. Se ciò è stato verificato si è impostata la variabile transitoria “onset” ad uno altrimenti si è incrementata la variabile “*Silence*”. Soddisfatta la condizione “onset”, si è effettuato un secondo test relativo all’*HNR*, che doveva essere maggiore di zero, e relativo alla variazione frequenziale percentuale tra il frame attuale (*i*) e quello precedente che non doveva essere maggiore del 50% dell’ottava considerata (il salto assoluto frequenziale è differente se lo shift di frequenza è verso l’alto o verso il basso). Solo se tutte le condizioni sopra descritte sono state verificate e soddisfatte, il frame è stato considerato valido e si sono calcolati e successivamente memorizzati i parametri da esso ricavati. In caso contrario, la variabile “onset” è stata impostata a zero e la variabile “*Unvoiced*” invece è stata incrementata. Una volta effettuata questa prima analisi, che ha permesso di scartare i frame che non hanno soddisfatto i prerequisiti sopra esposti, si è passati all’estrazione dei parametri.

3.6 Parametri estratti

3.6.1 Parametri estratti dalla vocale /a/ sostenuta

3.6.1.1 Parametri di perturbazione ed HNR

Nel capitolo 2 è stato spiegato come il segnale vocale appartenente ai soggetti Parkinsoniani rispetto al segnale vocale appartenente ai soggetti sani sia caratterizzato da perturbazioni non fisiologiche della frequenza fondamentale F_0 , e quindi del periodo fondamentale T_0 ; analogamente sono presenti delle perturbazioni riguardanti l'ampiezza. Per tale motivo uno dei metodi in grado di discriminare una voce appartenente ad un soggetto Parkinsoniano rispetto ad una voca appartenente ad un soggetto sano consiste nell'estrarre parametri che siano in grado di quantificare il grado di aperiodicità del segnale vocale stesso. Tuttavia, tali parametri risultano avere dei limiti. Il primo limite, già esposto nel paragrafo 3.5, risiede nella forte dipendenza dalla corretta individuazione della frequenza fondamentale F_0 . Il secondo limite, risiede nel fatto che questi stessi parametri possono essere estratti soltanto quando si elaborano delle vocali sostenute e non quindi per la lettura o per l'eloquio spontaneo. Una buona parte di questo lavoro di tesi è consistito nello sviluppo di algoritmi in grado di stimare tali parametri di perturbazione dalle tre vocali /a/ registrate per ogni soggetto. I parametri che sono stati estratti dall'analisi del segnale vocale delle vocali /a/ possono essere suddivisi in due sottogruppi: quelli relativi alle perturbazioni del periodo fondamentale e quelli relativi alle perturbazioni di ampiezza. Di seguito sono riportate le definizioni con le rispettive formule.

Tra i parametri che misurano le perturbazioni di T_0 sono stati implementati:

- **Jita (μs).** È il *jitter* assoluto e descrive la variazione media assoluta del periodo fondamentale T_0 tra un periodo ed il successivo [28]:

$$Jita = \frac{1}{N-1} \sum_{i=1}^{N-1} |T_0^{(i)} - T_0^{(i+1)}|$$

Dove $T_0^{(i)}$, con $i = 1, 2, \dots, N$, sono i periodi estratti dal segnale vocale ed N è il numero di periodi.

- **Jitt (%)**. È il *jitter* locale e descrive la variazione media relativa del periodo fondamentale T_0 tra un periodo ed il successivo:

$$Jitt = \frac{\frac{1}{N-1} \sum_{i=1}^{N-1} |T_0^{(i)} - T_0^{(i+1)}|}{\frac{1}{N} \sum_{i=1}^N T_0^{(i)}}$$

Dove $T_0^{(i)}$, con $i = 1, 2, \dots, N$, sono i periodi estratti dal segnale vocale ed N è il numero di periodi. La formula è simile a quella del *Jita*, difatti differisce solo per il denominatore.

- **RAP (%)**. È la perturbazione media relativa di tre periodi consecutivi con un passo di un periodo alla volta. La formula è:

$$RAP = \frac{\frac{1}{N-2} \sum_{i=2}^{N-1} \left| \frac{T_0^{(i-1)} + T_0^{(i)} + T_0^{(i+1)}}{3} - T_0^{(i)} \right|}{\frac{1}{N} \sum_{i=1}^N T_0^{(i)}}$$

Dove $T_0^{(i)}$, con $i = 1, 2, \dots, N$, sono i periodi estratti dal segnale vocale ed N è il numero di periodi. Il *RAP* è simile al *Jitt*, ma in questo caso invece di andare a calcolare la differenza tra un periodo ed il successivo, viene calcolata la media su tre periodi consecutivi, successivamente viene sottratto il valore del periodo centrale.

- **PPQ (%)**. Definito come *Pitch Period Perturbation Quotient* definisce la perturbazione media relativa di cinque periodi consecutivi. La formula è simile a quella del *RAP* con l'unica eccezione di considerare cinque periodi consecutivi:

$$PPQ = \frac{\frac{1}{N-4} \sum_{i=1}^{N-4} \left| \frac{1}{5} \sum_{r=0}^4 T_0^{(i+r)} - T_0^{(i+2)} \right|}{\frac{1}{N} \sum_{i=1}^N T_0^{(i)}}$$

- **vF₀ (%)**. È la variazione della frequenza fondamentale. Indica la variazione relativa della deviazione standard di F₀ rispetto alla frequenza fondamentale media calcolata:

$$vF_0 = \frac{\sigma}{F_0} \times 100 = \frac{\sqrt{\frac{1}{N} \sum_{i=1}^N (F_0 - F_0^{(i)})^2}}{\frac{1}{N} \sum_{i=1}^N f_0^{(i)}} \times 100$$

Dove F₀ è la media della frequenza fondamentale, σ è la deviazione di f₀, e f₀⁽ⁱ⁾ è il singolo valore di f₀ estratto.

Tra i parametri che misurano le perturbazioni d'ampiezza sono stati implementati:

- **ShdB (dB)**. È lo *shimmer* assoluto e descrive la variabilità media assoluta tra due periodi consecutivi dell'ampiezza picco-picco. La formula è:

$$ShdB = \frac{1}{N-1} \sum_{i=1}^{N-1} \left| 20 \log \frac{A^{(i+1)}}{A^{(i)}} \right|$$

Dove A⁽ⁱ⁾, con i che va da uno ad N a passi di uno, è l'ampiezza picco-picco mentre N è il numero di impulsi estratti. Tale parametro indica una perturbazione di ampiezza a breve termine.

- **Shim (%)**. È lo *shimmer* locale e descrive la variabilità relativa dell'ampiezza picco-picco tra due periodi consecutivi. La formula è la seguente:

$$Shim = \frac{\frac{1}{N-1} \sum_{i=1}^{N-1} |A^{(i)} - A^{(i+1)}|}{\frac{1}{N} \sum_{i=1}^N A^{(i)}}$$

Dove $A^{(i)}$, con i che va da uno ad N a passi di uno, è l'ampiezza picco-picco mentre N è il numero di impulsi estratti. Tale parametro indica una perturbazione di ampiezza a breve termine. Sia lo *Shim* che l'*ShdB* misurano la stessa perturbazione di ampiezza, ma mentre lo *shimmer* locale viene espresso in percentuale, lo *shimmer* assoluto viene espresso in Decibel.

- **APQ (%)**. Viene definito come *Amplitude Perturbation Quotient* e descrive la perturbazione di ampiezza picco-picco relativa tra undici periodi consecutivi con un passo di un periodo alla volta. La formula è la seguente:

$$APQ = \frac{\frac{1}{N-10} \sum_{i=1}^{N-10} \left| \frac{1}{11} \sum_{r=0}^{10} A^{(i+r)} - A^{(i+5)} \right|}{\frac{1}{N} \sum_{i=1}^N A^{(i)}}$$

Dove $A^{(i)}$, con i che va da uno ad N a passi di uno, è l'ampiezza picco-picco mentre N è il numero di impulsi estratti. Tale parametro è meno sensibile all'errore relativo all'estrazione del picco rispetto allo Shim, ma è comunque un parametro che fornisce una forte indicazione sulla variabilità d'ampiezza a breve termine della voce.

- **vAm (%)**. Viene definita come *Peak Amplitude Variation*. Fornisce la variabilità relativa delle variazioni di ampiezza picco-picco (a breve-lungo termine) all'interno del campione vocale analizzato:

$$vAm = \frac{\sigma}{A_0} \times 100 = \frac{\sqrt{\frac{1}{N} \sum_{i=1}^N (A_0 - A^{(i)})^2}}{\frac{1}{N} \sum_{i=1}^N A^{(i)}} \times 100$$

Dove $A^{(i)}$, con i che va da uno ad N a passi di uno, è l'ampiezza picco-picco, N è il numero di impulsi estratti ed A_0 è la media dei valori di ampiezza picco-picco estratti.

Un altro parametro estratto, già precedentemente nominato nel paragrafo 3.4, è rappresentato dall'HNR. Tale parametro è calcolato come il rapporto fra le componenti di energia spettrale armonica e le componenti di energia spettrale disarmonica, misurando quindi la quantità di rumore additivo presente nel segnale vocale. La formula relativa al calcolo dell'HNR è la seguente:

$$HNR = 10 \times \log_{10} \frac{AC_V(0)}{AC_V(0) + AC_V(T)}$$

Dove $AC_V(0)$ è il valore della funzione di autocorrelazione calcolato per un ritardo nullo (*frame power*) mentre $AC_V(T)$ è il valore della funzione di autocorrelazione per un ritardo pari a T. L'HNR quindi, per quanto appena esposto, assume un significato simile di quello espresso dal rapporto segnale-rumore (*Signal-Noise Ratio*: SNR).

3.6.1.2 Parametri Cepstrali

Oltre a tutti i parametri estratti nel dominio del tempo, ad eccezione dell'HNR che è stato estratto nel dominio della frequenza si sono estratti parametri relativi al CPPS calcolati quindi nel dominio della *quefreny*. Per quanto riguarda la descrizione dettagliata di questi parametri, sarà effettuata nel paragrafo relativo alla lettura del brano.

3.6.2 Parametri estratti dalla lettura del brano

Come specificato in precedenza, per l'analisi del segnale vocale relativo alla lettura del brano, non ha senso calcolare i parametri di perturbazione. Al contrario, i parametri relativi al CPPS possono e sono stati estratti sia per la prima parte del protocollo (ripetizione delle tre /a/ sostenute) che per la lettura del brano foneticamente bilanciato. In particolare, per l'analisi di perturbazione relativa alle vocali /a/ il segnale è stato sovra-campionato di 8 volte per aumentare la risoluzione temporale e di ampiezza. Mentre, per l'analisi del CPPS relativo alla lettura del brano ed alle vocali /a/, il segnale è stato sotto-campionato di un fattore 2. Il calcolo

del CPPS è stato effettuato sempre e solo per i frame che sono stati ritenuti validi dall'algoritmo descritto nel paragrafo 3.4. Come già esposto all'interno del capitolo 2, il cepstrum è definito come lo spettro logaritmico di uno spettro logaritmico [29]. Dato un generico segnale vocale x , è possibile definire:

$$xFFT = 20 \times \log | FFT(x) |$$

$$xFFT2 = 20 \times \log | FFT(xFFT) |$$

Dove $xFFT$ indica lo spettro del segnale vocale (x), mentre $xFFT2$ è il *cepstrum*.

L'analisi è stata eseguita dividendo il segnale vocale in finestre di 1024 campioni con un *overlap* di 2ms. Di ogni finestra, moltiplicata per una finestra di Hamming, si è calcolato lo spettro utilizzando la funzione Matlab FFT. A questo punto per calcolare il cepstrum, è stata effettuata una seconda FFT sullo spettro precedentemente calcolato secondo la formula riportata ad inizio paragrafo. Tale operazione è ripetuta per ogni frame andando ad ottenere un certo numero di "cepstra". I "cepstra" sono stati mediati prima nel tempo ogni sette frame (durata di 14 ms) e successivamente nel dominio della *quefreny* ogni sette bin. Il CPPS (*Cepstral Peak Prominence Smoothed*) è la misura della distanza in decibel tra l'ampiezza del picco cepstrale ed il valore della regressione lineare a quella stessa frequenza. Per tale motivo per ogni singolo frame è stata calcolata la regressione lineare e solo dopo si è potuta calcolare tale differenza. A questo punto, poiché è noto a priori il campo di frequenza del segnale vocale, il massimo del cepstrum è stato ricercato in un intervallo temporale compreso tra 3,3 ms e 16,7ms. Come output si è quindi ottenuta una distribuzione di CPPS, uno per ogni singolo frame valido, ciò ha quindi permesso di andare ad effettuare una successiva analisi estraendo indici statistici descrittivi ad esso collegati come:

- La media ($CPPS_{mean}$), la mediana ($CPPS_{median}$), la moda ($CPPS_{mode}$), il 5° percentile ($CPPS_{5prc}$) ed il 95° percentile ($CPPS_{95prc}$) che hanno permesso di verificare quale fosse la misura di localizzazione della distribuzione;

- La deviazione tipo ($CPPS_{std}$) ed il *range* ($CPPS_{range}$) che hanno permesso di quantificare la dispersione della distribuzione;
- La *skewness* ($CPPS_{skewness}$) e la *kurtosis* ($CPPS_{kurtosis}$) che hanno invece permesso di andare a caratterizzare la forma della distribuzione.

Nella figura 12 sono riportate le di distribuzioni relative al CPPS per la vocale /a/ sostenuta, registrata mediante il microfono a contatto piezoelettrico sia per un soggetto sano che per un soggetto Parkinsoniano. Nelle figure successive, sono invece riportate le distribuzioni del CPPS relative alla lettura del brano sia per un paziente Parkinsoniano che per un soggetto sano. Avendo ottenuto le distribuzioni per lo stesso soggetto è evidenziata la differenza tra la distribuzione ottenuta con il microfono a contatto (figura 13) e la distribuzione ottenuta con il microfono in aria (figura 14).

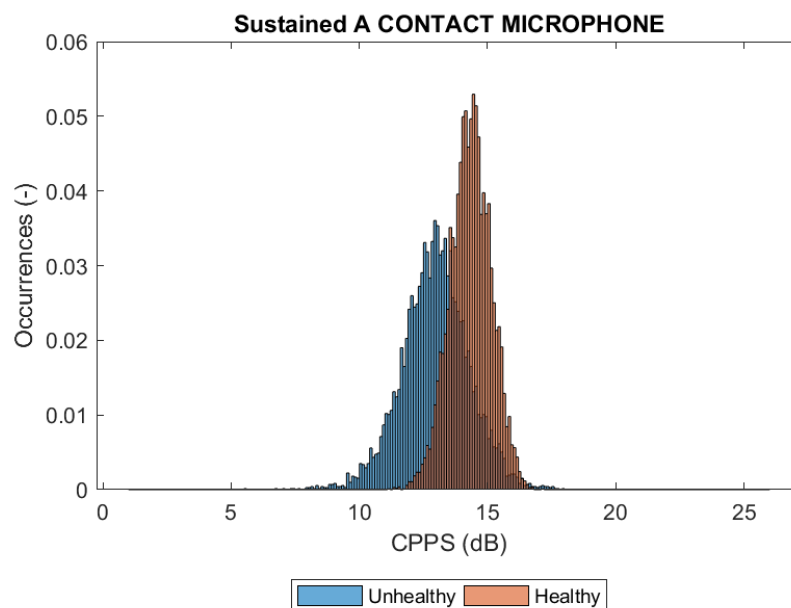


Figura 12 Distribuzione del CPPS per la vocale /a/ (microfono a contatto)

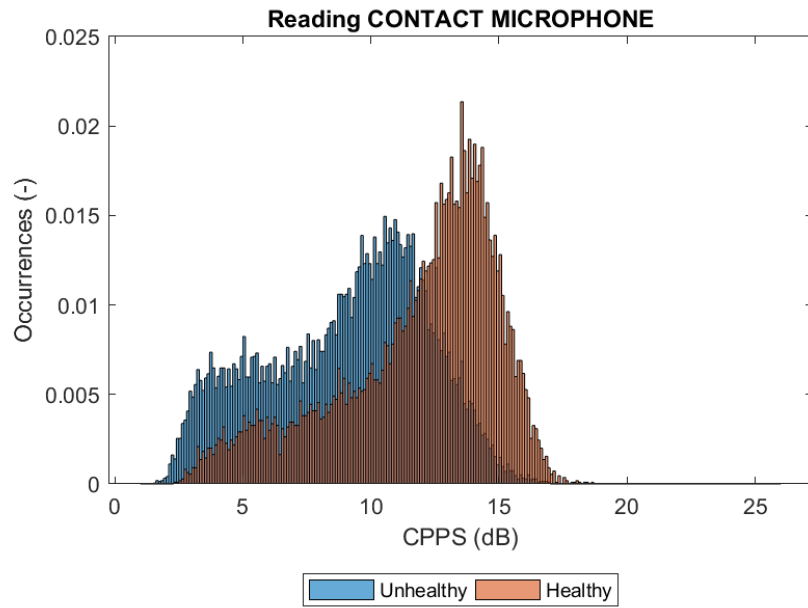


Figura 13 Distribuzione del CPPS per la lettura del brano (microfono a contatto)

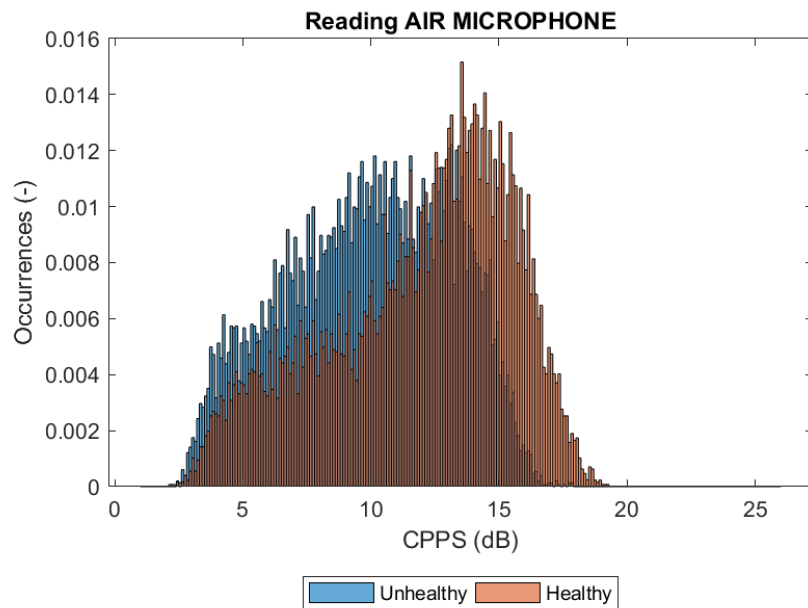


Figura 14 Distribuzione del CPPS per la lettura del brano (microfono in aria)

Oltre ad i parametri precedentemente illustrati, si è introdotto un nuovo parametro che è definito *Voiced-Unvoiced ratio* (V/U_v (%)). Tale parametro quantifica la quantità totale di frame sonori rispetto a quelli non sonori. La sua misura è stata

effettuata considerando le porzioni di frame con HNR positivo (fonemi sonori) rispetto alla lunghezza delle parti non-armoniche (fonemi silenti).

3.7 Selezione dei parametri

Poiché non tutti i parametri esposti nel paragrafo precedente risultano essere rilevanti al fine di classificare un soggetto come sano, al quale è stata associato un valore relativo alla classe di appartenenza pari ad uno, o come Parkinsoniano al quale invece è stato assegnato un valore relativo alla classe di appartenenza pari a zero, si sono adoperati dei metodi che hanno permesso di andare ad identificare le sole variabili rilevanti. Tali metodi, sono stati adoperati per ogni fase del protocollo elaborata per ogni microfono utilizzato e per ogni algoritmo implementato. La prima fase della selezione dei parametri (*features selection*), è consistita nell'andare ad identificare le variabili ridondanti. Tali variabili, essendo fra loro correlate, portano la stessa informazione. Al fine quindi di effettuare una prima selezione si è calcolata la matrice di correlazione tramite la funzione "corr" di Matlab.

```
[R,p]=corr(Parametri,'Type','Pearson');
```

Tale funzione, come mostrato nella riga antecedente riportante il codice Matlab, è stata calcolata dando come input la matrice contenente tutti i parametri precedentemente calcolati. La stessa funzione ha restituito come output due matrici quadrate R e p. La prima matrice è la matrice di correlazione, la seconda è la matrice contenente i p-value. La matrice di correlazione R, contiene i coefficienti di Pearson ricavati tramite la seguente formula:

$$\rho(X,Y) = \frac{cov(X,Y)}{\sigma_x \sigma_y}$$

Dove il numeratore è la covarianza tra X e Y, mentre σ_x e σ_y sono le deviazioni standard di X e di Y. Il risultato è un valore compreso tra +1 e -1, dove un valore pari a +1 sta ad indicare una totale correlazione lineare e positiva, un valore pari a -1 sta ad indicare una totale correlazione lineare ma negativa, al contrario di un valore pari a zero che indica una totale assenza di correlazione. In questo caso si è

scelto di considerare due variabili come correlate se il valore del coefficiente di correlazione di Pearson presentava un valore maggiore o uguale a 0,8. Oltre alla matrice di correlazione contenente i valori dei coefficienti di correlazione di Pearson la funzione Matlab ha restituito la matrice “p”, che ha permesso di andare a non considerare gli elementi $R(i, j)$ qualora il corrispettivo valore $p(i, j)$ risultava essere maggiore o uguale a 0,05. Una volta evidenziate le variabili tra loro correlate, scartando quindi le ridondanze, ci si è concentrati sull’identificare le variabili rilevanti a discapito di quelle irrilevanti. Per far ciò si è adoperata la *tool* di Matlab *Classification Learner*. In tale *tool* è stata inserita la matrice contenente tutti i parametri di interesse precedentemente calcolati e la rispettiva classe di appartenenza. La stessa *tool* ha dato la possibilità di andare a selezionare in maniera distinta le varie variabili permettendo così di considerare esclusivamente quelle tra loro non correlate, identificate precedentemente con la matrice di correlazione. Il metodo finale, su cui è stata basata la scelta finale delle variabili, è ricaduto sul grafico *Parallel Coordinates Plot* presente nella stessa tool (figura 15). Tale grafico, è un grafico 2D che mostra le variabili da includere o escludere in base alla potenzialità delle stesse nell’andare a discriminare meglio il segnale vocale dei soggetti in esame.

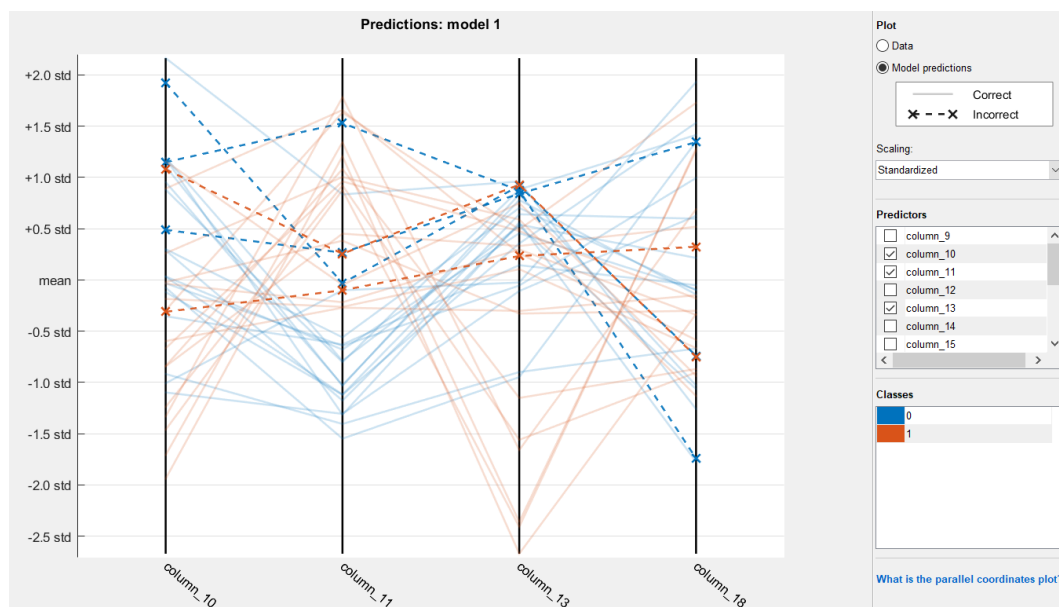


Figura 15 Esempio di Parallel Coordinates Plot

Le variabili rilevanti identificate, come sarà esposto nel capitolo relativo ai risultati, sono risultate essere differenti in base non solo alla particolare fase del protocollo analizzata, ma anche in base al microfono ed all'algoritmo utilizzato per la costruzione del modello. Infatti, per ogni fase del protocollo e per ogni microfono utilizzato, sono stati implementati due differenti algoritmi che hanno permesso di ottenere dei modelli predittivi. Tali modelli, sono stati caratterizzati da delle specifiche in termini di accuratezza, sensibilità e specificità. Gli algoritmi in esame, che hanno permesso di ottenere dei modelli in grado di classificare un soggetto come sano o come Parkinsoniano, sono stati il *coarse tree* e la regressione logistica.

3.8 Coarse tree

L'algoritmo "*coarse tree*" è un particolare algoritmo che fa parte di tutti quelli algoritmi che vanno sotto il nome di "*decision tree*" (alberi decisionali). Gli alberi decisionali sono considerati uno degli approcci più popolari per rappresentare i classificatori e sono espressi come partizione ricorsiva dello spazio degli stati. L'albero decisionale è costituito da dei nodi che formano un albero radicato, il primo nodo è chiamato "root" e non presenta entrate, al contrario, tutti gli altri ne hanno esattamente una. Se un nodo presenta delle diramazioni è chiamato nodo interno o nodo di test, se invece non presenta diramazioni viene definito foglia o nodo terminale. In un albero decisionale, ciascun nodo interno divide lo spazio degli stati in due o più spazi secondari in base a una determinata funzione discreta dei valori di input. L'albero decisionale deve essere letto partendo dalla radice fino ad arrivare ad una foglia, percorrendo il percorso rappresentato dallo stesso [30]. Nella figura 16 è presente un esempio ottenuto mediante la *tool* di Matlab "*classification learner - coarse tree*".

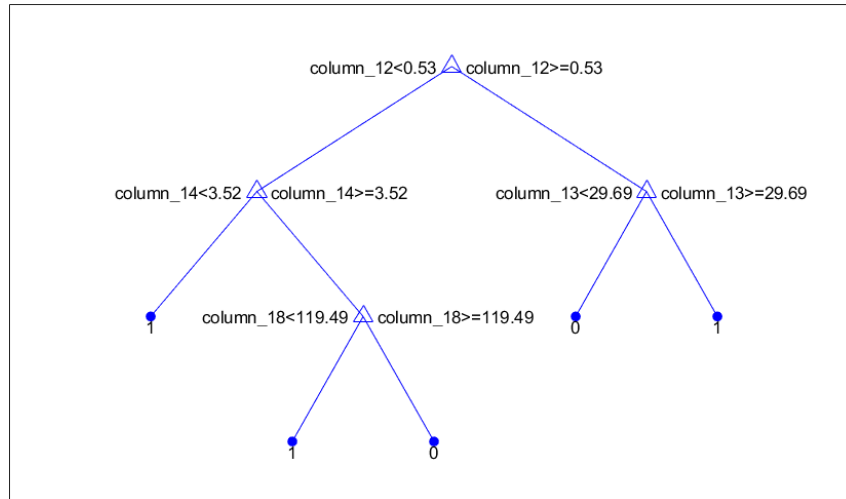


Figura 16 Esempio di albero decisionale

Come si può notare dalla figura, ogni nodo è relativo ad una variabile e da ogni stesso nodo partirà una diramazione fino ad arrivare alle foglie che permettono, in base al percorso effettuato per arrivarci, di andare ad identificare un soggetto come sano (classe 1) o come Parkinsoniano (classe 0). Ciò non risulta essere molto leggibile in quanto non è possibile avere un riscontro immediato su quale sia la variabile in gioco. Per tale motivo, nei risultati si sono costruiti dei diagrammi di flusso per ogni albero decisionale ottenuto. Naturalmente, sono da preferire degli alberi decisionali meno complessi essendo appunto considerati maggiormente comprensibili. Inoltre, secondo Breiman et al. [31] la complessità dell'albero ha un effetto cruciale sulla sua accuratezza. In genere, la complessità dell'albero viene misurata da una delle seguenti metriche: il numero totale di nodi, il numero totale di foglie, la profondità dell'albero e il numero di attributi utilizzati. Inoltre, l'induzione dell'albero decisionale è strettamente correlata all'induzione della regola. Infatti, ogni percorso dalla radice di un albero decisionale a una delle sue foglie, può essere trasformato in una regola. Questa è stata solo una delle motivazioni che sta alla base dell'utilizzo di tale metodo in questo lavoro di tesi. Una volta ottenuto l'albero decisionale utilizzando la tool di Matlab *Classification Learner*, si è esportato e salvato il modello. Successivamente mediante un apposito script Matlab si è verificata quale fosse la classe predetta dal modello precedentemente creato utilizzando utilizzata la seguente riga Matlab:

```
yfit=trainedModel.predictFcn(X);
```

Dove X è la matrice che contiene un numero di righe pari al numero di segnali vocali in analisi ed un numero di colonne pari al numero di parametri di interesse calcolati per ogni singolo segnale vocale. “yfit”, è invece l’output generato dal modello che presenterà quindi valori pari a zero se il modello ha associato il segnale vocale come appartenente ad un soggetto Parkinsoniano o pari ad uno se il modello ha associato il segnale vocale come appartenente ad un soggetto sano. A questo punto, ogni modello è stato descritto in termini di *accuracy*, *sensitivity* e *specificity*. L’ *accuracy*, infatti, non risulta essere una misura sufficiente per andare a valutare un modello, soprattutto quando si ha una distribuzione squilibrata, la cui sola stima può fuorviare la valutazione del modello creato. La definizione di *accuracy* segue la seguente formula:

$$Accuracy = \frac{true_{positive} + true_{negative}}{positive + negative}$$

Per poterla calcolare si è creato uno script Matlab che ha permesso di confrontare la classe reale con la classe predetta, in maniera tale da ottenere una matrice di confusione (*confusion matrix*). Per i motivi sovra esposti, non ci si è fermati alla sola accuratezza ma sono state introdotte le misure di *sensitivity* e *specificity*. La *sensitivity* sta ad indicare quanto bene il classificatore riesce ad identificare campioni positivi ed è definita come segue:

$$Sensitivity = \frac{true_{positive}}{positive}$$

Dove il numeratore indica il numero di soggetti che vengono diagnosticati dal classificatore come positivi e che sono realmente appartenenti a tale classe, al contrario del denominatore che è dato dal numero totale di soggetti che appartengono alla classe dei positivi e che quindi includono i veri positivi ed i falsi negativi. La *specificity* è invece la misura complementare della *sensitivity* e rileva quanto bene il classificatore riesce a discriminare i soggetti appartenenti alla classe negativa. Tale parametro viene definito come segue:

$$Specificity = \frac{true_{negative}}{negative}$$

Il numeratore indica il numero di veri negativi, mentre il denominatore è dato dalla somma dei veri negativi e dei falsi positivi. Questi parametri, come vedremo nel prossimo capitolo, hanno permesso di andare ad effettuare dei confronti tra i diversi microfoni tra le diverse fasi del protocollo ed anche tra i due diversi modelli ottenuti.

3.9 Regressione logistica

Utilizzando sempre la *tool* di Matlab *classification learner*, si è utilizzato il modello di regressione logistica, ciò è concesso poiché siamo di fronte ad un output dicotomico. Nel modello di regressione logistica la curva che rappresenta la relazione tra la variabile indipendente X e la variabile dipendente Y non risulta essere lineare, ma al contrario ha un andamento sigmoidale come quello rappresentato in figura 17.

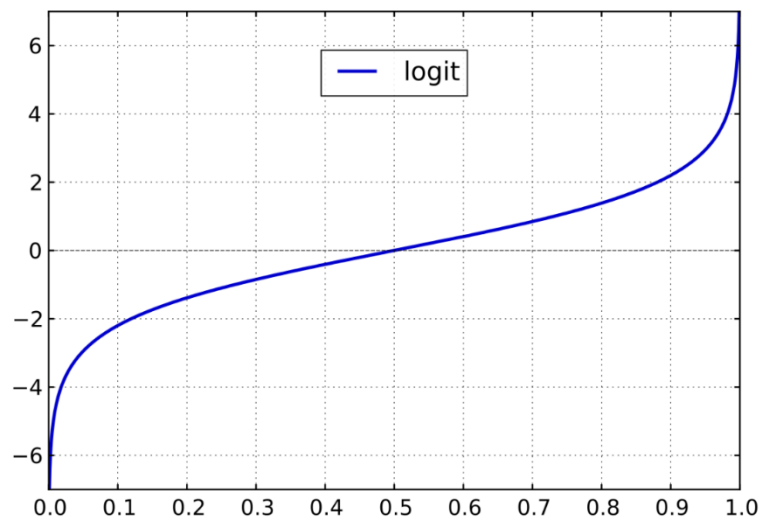


Figura 17 Funzione sigmoidale

Una volta estratto e salvato il modello, la stessa *tool* di Matlab ha permesso di ricavare i parametri relativi all'intercetta e ad ogni *slope* associata a ciascun parametro precedentemente scelto nella fase di *features selection*. Ogni parametro esposto viene affiancato ad un valore di p-value. In questo caso specifico si è scelto di considerare validi i soli parametri che presentavano un tale valore minore di 0,05.

Oltre a questi parametri, la stessa *tool* ha restituito un vettore contenente la probabilità associata al segnale vocale in esame i-esimo di appartenenza alla classe dei sani (classe 1). Tale probabilità, in accordo con la teoria alla base della regressione logistica, è stata ricavata mediante la funzione esposta nella formula sottostante:

$$P(Healthy = 1) = \frac{e^{(Intercept + Slope_1 * Parameter_1 + \dots + Slope_n * Parameter_n)}}{1 + e^{(Intercept + Slope_1 * Parameter_1 + \dots + Slope_n * Parameter_n)}}$$

Dove $P(Healthy=1)$ sta ad indicare la probabilità di avere una voce sana con un *range* che varia da zero ad uno. A questo punto, definito il vettore di probabilità è stato necessario associare ogni valore di probabilità i-esimo ad una delle due possibili classi di output. Per far ciò, si è scelto il valore di *cut-off* (in termini di probabilità). Tale scelta è stata basata su di un metodo grafico. Si è graficata la variazione simultanea della *sensitivity* e della *specificity* in funzione del possibile *cut-off* espresso in termini di probabilità. Al crescere della soglia di *cut-off*, che indica la soglia oltre la quale un soggetto viene fatto appartenere alla classe dei sani (classe 1), si ha un aumento della *sensitivity* in quanto si riduce il numero falsi negativi ed una riduzione della *specificity* in quanto si aumenta il numero dei falsi positivi. Il passo successivo è consistito nell'identificare il punto di incontro tra le due curve ricavando il relativo valore di *cut-off*. Tuttavia, poiché si è indirizzati a massimizzare la *sensitivity* rispetto alla *specificity* non si è sempre scelto come valore di *cut-off* il punto ottenuto dall'incontro delle due curve ma ci si è spostati all'interno di un *range* ristretto con l'intento di ottenere le migliori prestazioni. A questo punto, fissata la soglia, si è calcolata la classe predetta dal modello per ogni singolo segnale vocale. Per far ciò, si è confrontato il valore di probabilità ottenuto con la soglia scelta. In particolare, se la probabilità è risultata essere maggiore della soglia fissata il soggetto è stato associato alla classe dei sani (classe 1), se è risultata essere minore o uguale è stato associato alla classe dei Parkinsoniani. In questo modo si è ottenuto un vettore denominato “classe predetta” formato da soli zeri ed uni. Tale vettore, ha permesso di confrontare la classe predetta i-esima con la rispettiva classe reale di appartenenza ottenendo così una matrice di confusione. Anche in questo caso, si sono valutate le specifiche in termini di *accuracy*,

sensitivity e *specificity*. Inoltre, è stato anche possibile andare a costruire la curva ROC indicante la relazione tra *sensitivity* ed $1 - \text{specificity}$, permettendo di ottenere un altro parametro significativo in termini di prestazioni che è rappresentato dall'*area under curve* (AUC). Difatti, una AUC vicina a 1 suggerisce una forte capacità del modello di separare soggetti con disturbi vocali (Parkinsoniani) da quelli con voci normali (sani); diversamente, un AUC vicina a 0,5 identifica una bassa capacità di distinzione tra i due gruppi, poiché la probabilità di classificare un soggetto come sano o come Parkinsoniano è la stessa. Nella scelta del modello migliore avere un alto valore dell'AUC è decisivo.

In accordo con diversi test statistici, il valore di AUC può essere interpretato come segue:

- $AUC=0,5$ test fallimentare;
- $0,5 < AUC \leq 0,7$ test con bassa accuratezza;
- $0,7 < AUC \leq 0,9$ test con accuratezza moderata;
- $0,9 < AUC < 1$ test con accuratezza elevata;
- $AUC=1$ test perfetto;

Convenzionalmente un test diagnostico può essere considerato rilevante se presenta valori di AUC maggiori o uguali a 0,80 [32]

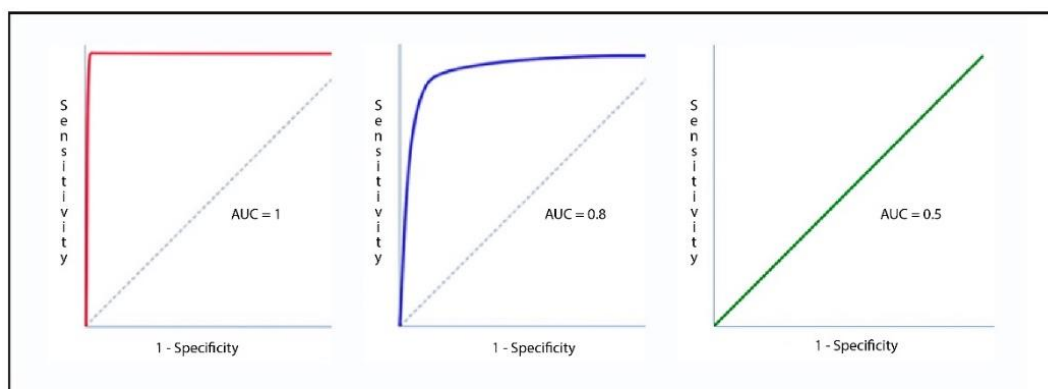


Figura 18 Curva ROC [4]

4. RISULTATI

4.1 Vocale /a/ sostenuta

In questa sezione si analizzeranno i risultati ottenuti dall'elaborazione del segnale vocale derivante dalla registrazione mediante microfono a contatto della vocale /a/ sostenuta. Tale analisi, come è stato precedentemente specificato nel capitolo precedente, verrà effettuata mediante i due diversi algoritmi *coarse tree* e regressione logistica, entrambi basati su di una prima analisi di correlazione ottenuta mediante la matrice di correlazione.

4.1.1 Matrice di correlazione

La tabella 4 riporta i parametri di perturbazione, la frequenza fondamentale media F_0 , l'HNR e tutti i parametri statistici ricavati dal CPPS (evidenziati in grigio) come la media, la mediana, la moda, la deviazione tipo, il range, il 5^o percentile, il 95^o percentile, la *skewness* e la *kurtosis*. Essendo la matrice di correlazione una matrice simmetrica, per rendere la stessa maggiormente leggibile, se ne è riportata solo la metà superiore. Di questa metà, non sono stati tenuti in considerazione gli elementi che presentavano un corrispettivo valore di p-value maggiore di 0,05. Tra i parametri relativi alle perturbazioni frequenziali e di ampiezza, il *jitter* e l'*APQ* risultano essere i parametri che presentano un numero di correlazioni maggiori con i restanti parametri di perturbazione. Un numero di correlazioni rilevanti lo si nota anche per l'HNR. Passando ai parametri *Cepstrali* la media è il parametro che ha un numero di correlazioni maggiori con i rimanenti parametri del dominio *cepstrale*; la deviazione tipo ed il *range* invece sono correlati solo tra loro, al contrario della *skewness* e *kurtosis* che oltre ad essere tra di loro non correlate non presentano alcuna correlazione.

	Jitt	Jita	RAP	PPQ	vF0	Shim	ShdB	APQ	vAm	F0	HNR	CPPS						skewness	kurtosis
	1	0.87	0.98	0.51	0.64	0.87	0.87	0.86	0.54		-0.81	mean	median	mode	sid	range	5prc	95prc	
Jita		1	0.83	0.49	0.47	0.78	0.78	0.80	0.41	-0.40	-0.75	-0.36	-0.32	-0.31	0.53	0.56	-0.49		
RAP			1	0.47	0.59	0.85	0.85	0.83	0.55		-0.81	-0.25	-0.21	-0.22	0.39	0.47	-0.37		
PPQ				1	0.45	0.40	0.39	0.43	0.30		-0.29	-0.41	-0.36	-0.35	0.49	0.52	-0.52	-0.18	
vF0					1	0.54	0.54	0.51	0.52	0.28	-0.43	-0.22	-0.19		0.53	0.56	-0.34		-0.21
Shim						1	1.00	0.98	0.56		-0.91	-0.37	-0.34	-0.31	0.46	0.53	-0.49		
ShdB							1	0.98	0.58		-0.91	-0.37	-0.34	-0.31	0.45	0.52	-0.49		
APQ								1	0.56		-0.89	-0.35	-0.30	-0.28	0.43	0.51	-0.46		
vAm									1	1	-0.51	-0.36	-0.32	-0.29	0.40	0.47	-0.44		
F0										1					0.20				-0.19
HNR											1	0.48	0.46	0.43	-0.41	-0.48	0.56	0.27	
mean												1	0.99	0.93	-0.26	-0.34	0.94	0.88	
median													1	0.96	-0.24	-0.32	0.91	0.87	
mode														1	-0.27	-0.32	0.86	0.81	
range															1	0.88	-0.50		
5prc																1	-0.55		0.18
95prc																	1	0.71	
skewness																		1	-0.19
kurtosis																			1

Tabella 4 Matrice di correlazione relativa ai parametri calcolati dalle vocali /a/ sostenute acquisite con il microfono a contatto

4.1.2 Coarse tree

Evidenziate le correlazioni presenti, il passo successivo è stato quello di ottenere un modello mediante l'utilizzo dell'algoritmo *coarse tree*. Tale algoritmo, nel caso specifico in esame, ha identificato come variabili rilevanti quelle riportate nella tabella sottostante:

<u>Comparison</u>	<u>Algorithm</u>	<u>Features selected</u>
Parkinsonian-Healthy	Coarse tree	PPQ,Shim,F0,vF0

Tabella 5 Variabili rilevanti per l'algoritmo *coarse tree*

Dove i primi tre parametri sono stati già definiti nel paragrafo 3.6 mentre F_0 sta ad indicare la frequenza fondamentale media. Come si evince dalla matrice di correlazione, tali parametri risultano essere tra di loro non correlati e quindi non stati giustamente utilizzati parametri ridondanti. Inoltre, è opportuno sottolineare come tale algoritmo non ha identificato nessuno parametro *cepstrale* come rilevante, ma si è basto esclusivamente sui parametri di perturbazione temporali e di ampiezza. Come anticipato nel paragrafo 3.8 l'algoritmo *coarse tree* restituisce un albero decisionale dal quale si è ricavato il corrispettivo *flow-chart*.

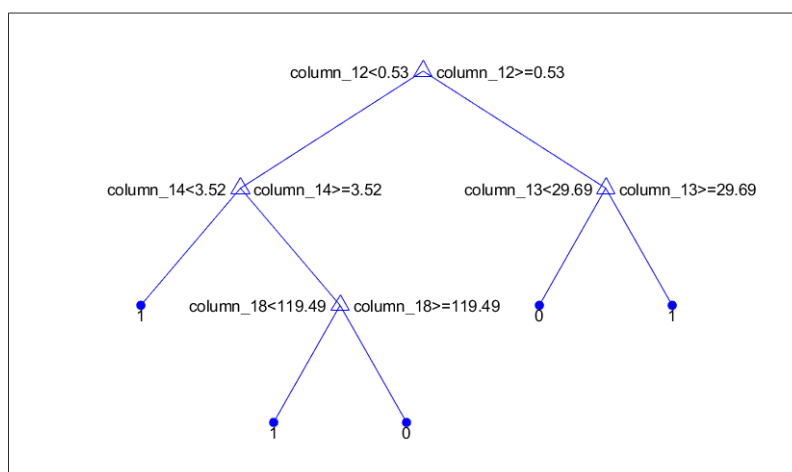


Figura 19 Albero decisionale relativo alla vocale /a/

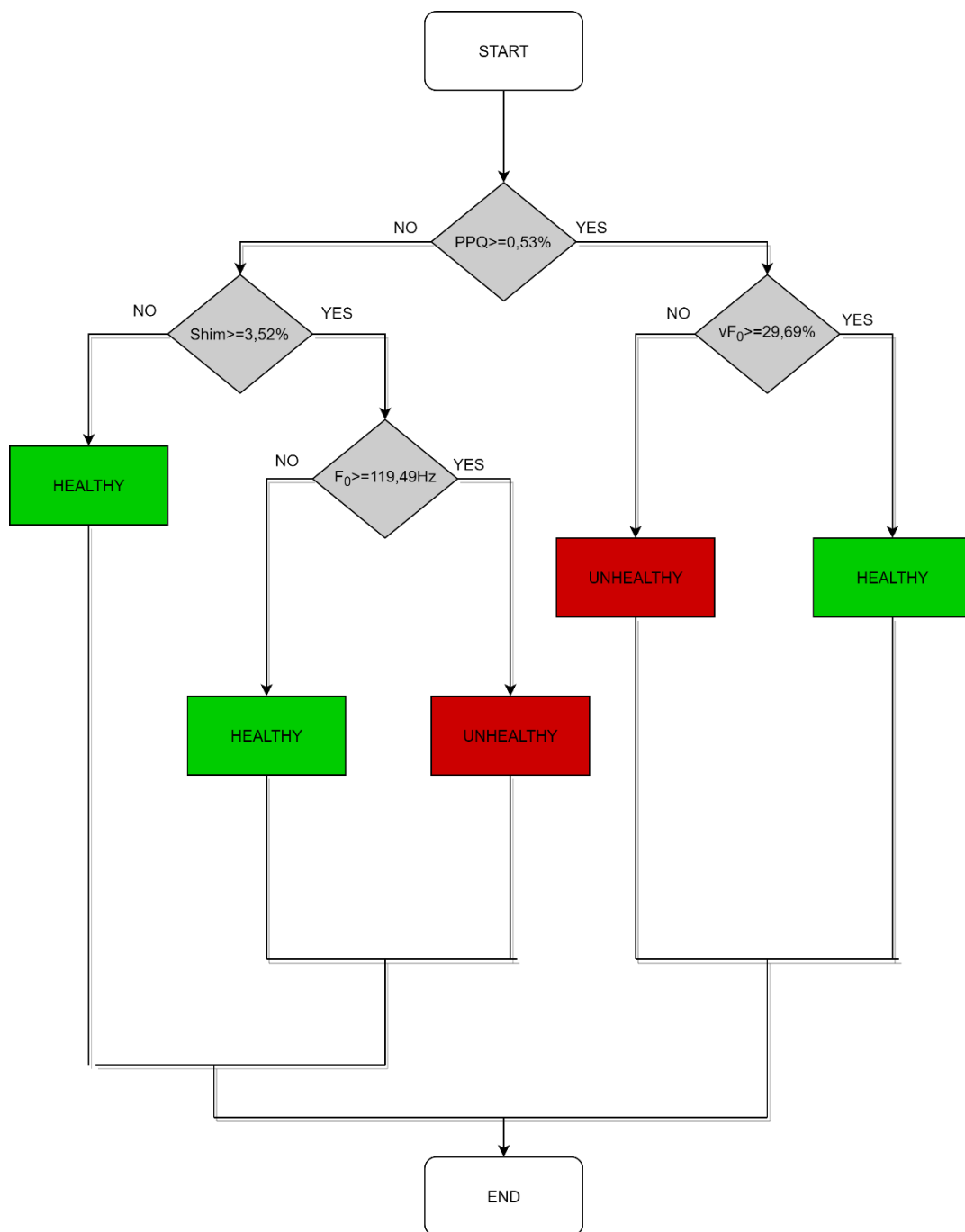


Figura 20 Flow-chart relativo alla vocale /a

Il *flow-chart* riportato in figura 20 parte dal PPQ che mostra un valore di soglia pari a 0,53%; tale parametro, per ogni segnale vocale studiato, può presentare valori maggiori, uguali o inferiori al valore di soglia. In particolare, se assume valori maggiori o uguali dello 0,53% si passa a verificare il vF_0 , che mostra invece una soglia pari al 29,69%. A questo punto, se quest'ultimo parametro è maggiore o

uguale a tale soglia, il soggetto in esame viene diagnosticato come sano, in caso contrario, viene diagnosticato come Parkinsoniano. Se al contrario il *PPQ* calcolato per il segnale vocale del soggetto in esame presenta un valore minore dello 0,53%, si passa a verificare il valore assunto dallo *Shim*. Se lo *Shim*. a sua volta presenta un valore inferiore al 3,52% allora si conclude che il soggetto appartiene alla classe dei sani. Al contrario valori di *Shim*. maggiori del 3,52% portano invece ad una ulteriore verifica basata sulla frequenza fondamentale media F_0 . Quest'ultima, infatti, permette di identificare un soggetto come sano se il suo valore è inferiore a 119,49 Hz o come Parkinsoniano se il suo valore è maggiore o uguale al valore di soglia precedentemente esposto. Seguendo il seguente flow-chart per ogni segnale vocale analizzato, l'algoritmo ha associato ciascun soggetto come appartenente alla classe dei sani o alla classe dei Parkinsoniani. In questo caso specifico, considerando che ogni soggetto ha ripetuto per tre volte la vocale /a/, si è deciso in prima istanza di considerare separatamente ogni vocale. Ottenuta la classe predetta dal modello, e nota a priori la classe reale, mediante un apposito script Matlab si è calcolata la seguente *confusion matrix*:

		ACTUAL VALUE	
		Positive	Negative
CLASSIFIER PREDICTION	Positive	58	8
	Negative	5	55

Tabella 6 Confusion matrix relativa all'analisi di tutte le vocali /a/ mediante l'algoritmo coarse tree

La *confusion matrix* ricavata permette di esprimere le seguenti specifiche:

- Accuracy $\rightarrow 89,68\% \approx 90\%$
- Sensitivity $\rightarrow 92,06\% \approx 92\%$
- Specificity $\rightarrow 87,30\% \approx 87\%$

Il passo successivo, che ha permesso di effettuare una prima “validazione” del modello ricavato, è stato dare in pasto al modello i parametri ottenuti dalla media di quelli relativi alle singole vocali /a/ ripetute dallo stesso soggetto in esame. In tal modo, si è ottenuta una matrice avente un numero di righe pari al numero di soggetti in esame, 21, ed un numero di colonne pari al numero di parametri estratti. Utilizzando il flow-chart precedentemente ricavato dal diagramma ad albero ottenuto con l’algoritmo *coarse tree*, si è implementato un codice Matlab che presa la nuova matrice dei parametri medi andasse a calcolare la classe predetta. Nota la corrispettiva classe reale, per mezzo di un confronto tra le due, si è ottenuta la seguente *confusion matrix*:

		ACTUAL VALUE	
		Positive	Negative
CLASSIFIER PREDICTION	Positive	19	2
	Negative	2	19

Tabella 7 *Confusion matrix* relativa all’analisi media tra le tre vocali /a/ mediante l’algoritmo *coarse tree*

Tale *Confusion matrix* permette di esprimere le seguenti specifiche:

- Accuracy $\rightarrow 90,48\% \approx 90\%$
- Sensitivity $\rightarrow 90,48\% \approx 90\%$
- Specificity $\rightarrow 90,48\% \approx 90\%$

Dai risultati ottenuti, effettuando un confronto con le specifiche precedenti, si nota: un incremento dell’*accuracy* dello 0,8%, una riduzione della *sensitivity* dell’1,58% ed un aumento della *specificity* del 3,18 %.

4.1.3 Regressione logistica

L'analisi del segnale relativo alle vocali sostenute è proseguita utilizzando il modello di regressione logistica. I parametri rilevanti per il caso in esame sono stati identificati mediante l'utilizzo del *parallel coordinates plot* e della matrice di correlazione. Tali metodi hanno evidenziato come rilevanti il *Jitt*, il PPQ, il vAm, il CPPS_{95prc} e la CPPS_{skewness}. Rispetto all'algoritmo *coarse tree* si sono considerati un numero di parametri maggiore. Tuttavia, poiché si è partiti da un totale di venti parametri, una riduzione di un fattore pari a ¼ è stata considerata sufficiente. La tabella sottostante mostra i parametri rilevanti che sono stati utilizzati in questo algoritmo.

<u>Comparison</u>	<u>Algorithm</u>	<u>Features selected</u>
Parkinsonian-Healthy	Logistic regression	Jitt,PPQ,vAm,CPPS95prc,CPPSskewness

Tabella 8 Variabili rilevanti per l'algoritmo regressione logistica

Attraverso la *tool* di Matlab, come esposto nel paragrafo 3.9, si sono ricavati i parametri del modello che hanno permesso di esplicitare la formula indicante la probabilità di far appartenere un soggetto alla classe dei sani. La formula del modello in esame è la seguente:

$$P(\text{Healthy} = 1) = \frac{e^{(14,04+0,60*Jitt-1,20*PPQ-0,13*vAm-0,74*CPPS95prc+0,99*CPPSskewness)}}{1 + e^{(14,04+0,60*Jitt-1,20*PPQ-0,13*vAm-0,74*CPPS95prc+0,99*CPPSskewness)}}$$

A questo punto, si è scelta la soglia che meglio permettesse di discriminare un soggetto sano da uno Parkinsoniano. Per effettuare tale scelta ci si è basati sul grafico riportato in figura 21.

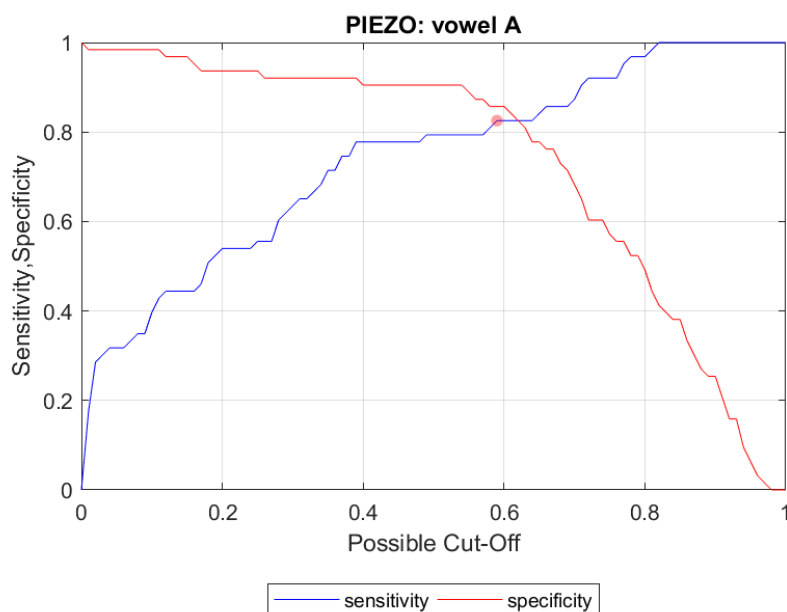


Figura 21 Scelta del possibile cut-off per il modello di regressione logistica ricavato dall'analisi del segnale vocale relativo alle vocali sostenute (microfono a contatto)

Da tale figura si nota come al crescere della soglia di *cut-off*, che indica il valore di probabilità oltre il quale un soggetto viene diagnosticato come sano, si ha un aumento della *sensitivity* ed una riduzione della *specificity*. In questo caso, non si è scelta come soglia il valore di incontro tra le due curve in quanto, spostando il *cut-off* verso valori inferiori la *sensitivity* rimane costante e la *specificity* assume un valore più elevato. Per tale ragione, si è scelto come *cut-off* un valore di probabilità pari a 0,6. La rispettiva curva ROC del modello in esame è riportata in figura 22.

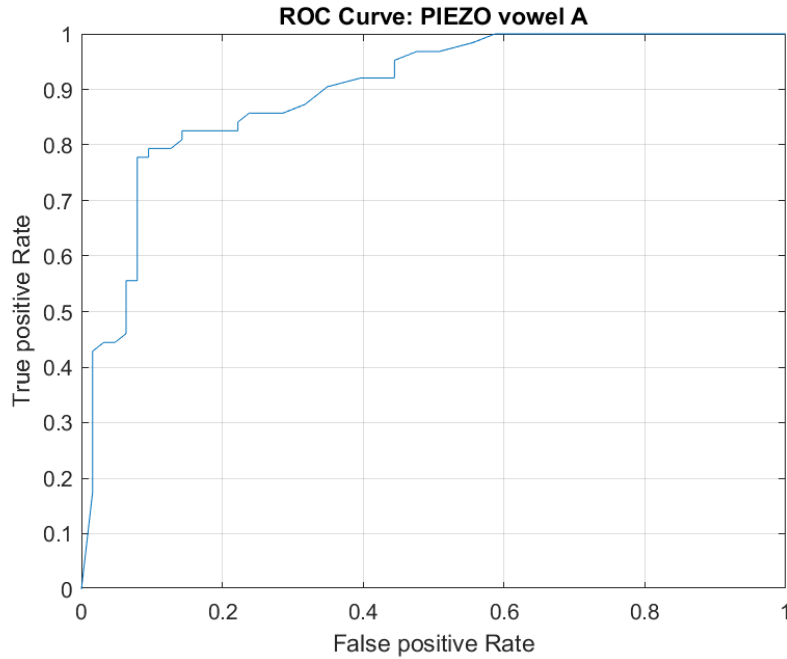


Figura 22 Curva ROC ricavata dal modello di regressione logistica ricavato dall'analisi del segnale vocale relativo alle vocali sostenute (microfono a contatto)

A questo punto, attraverso uno specifico script Matlab appositamente implementato, considerando che ogni soggetto ha ripetuto per tre volte la vocale /a/, e scelto in prima istanza di considerare separatamente ogni vocale, si è andati a confrontare il valore di probabilità ottenuto con la soglia scelta. Se la probabilità è risultata essere maggiore a 0,60 il soggetto è stato classificato come sano (classe 1). Al contrario, se la probabilità è risultata essere inferiore o uguale a 0,60 il soggetto è stato associato alla classe dei soggetti Parkinsoniani (classe 0). Ottenuto il vettore “classe predetta” lo si è confrontato con il vettore contenente la classe reale. Da tale confronto si è ricavata la *confusion matrix* seguente:

		ACTUAL VALUE	
		Positive	Negative
CLASSIFIER PREDICTION	Positive	52	9
	Negative	11	54

Tabella 9 Confusion Matrix relativa all'analisi di tutte le vocali /a/ mediante la regressione logistica

Le specifiche ricavate dalla *confusion matrix* e dalla curva ROC sono le seguenti:

- Accuracy $\rightarrow 84,13\% \approx 84\%$
- Sensitivity $\rightarrow 82,54\% \approx 83\%$
- Specificity $\rightarrow 85,71\% \approx 86\%$
- AUC=0,90.

Dalle specifiche riportate si evince che il modello creato commette degli errori, infatti sono presenti sia dei falsi positivi che dei falsi negativi. Per meglio vedere come si comporta il modello in esame viene riportato un grafico che, in base ai parametri scelti ed in base alla soglia precedentemente fissata, mostra come avviene la separazione tra le due classi.

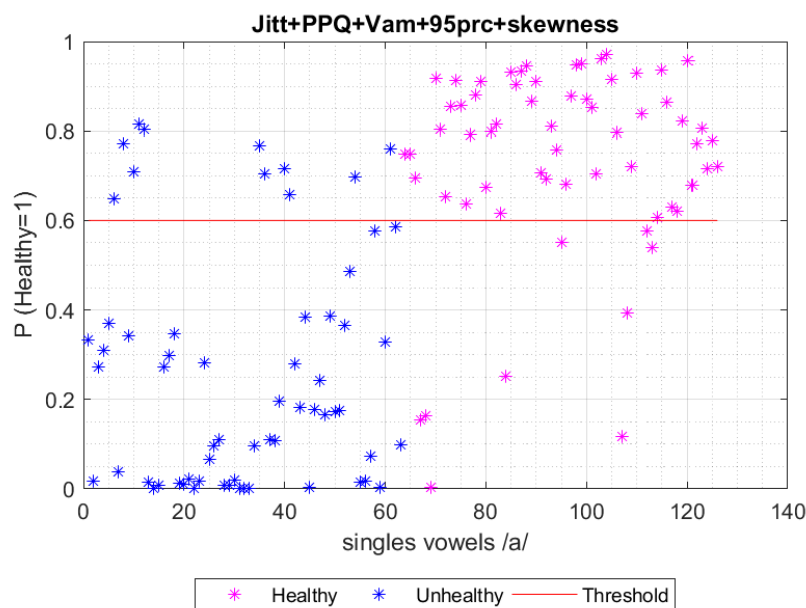


Figura 23 Grafico di separazione relativo al modello di regressione logistica ricavato dall'analisi delle singole vocali /a/

Il modello classifica erroneamente come sani undici vocali /a/ appartenenti a soggetti parkinsoniani (asterischi blu), e come parkinsoniani nove segnali vocali di

soggetti sani (asterischi viola). A questo punto, così come si è fatto con l'algoritmo *coarse tree*, si è effettuata una prima “validazione” considerando la media tra i parametri relativi alle tre vocali /a/ sostenute dallo stesso paziente. Per far ciò, si è ottenuta una matrice presentante un numero di righe pari al numero di soggetti in esame, 21, ed un numero di colonne pari al numero di parametri estratti. Dando questa matrice in pasto alla funzione di probabilità precedentemente trovata e considerata la stessa soglia fissata in precedenza, è stato possibile ottenere la seguente *confusion matrix*:

		ACTUAL VALUE	
		Positive	Negative
CLASSIFIER PREDICTION	Positive	10	1
	Negative	11	20

Tabella 10 *Confusion matrix* relativa all'analisi media tra le 3 vocali /a/ mediante la regressione logistica

Il classificatore, come si nota dalla *confusion matrix*, ha subito un crollo delle prestazioni:

- Accuracy → 71,43% ≈ 71%
- Sensitivity → 47,62% ≈ 48%
- Specificity → 95,24% ≈ 95%

Tali risultati sono indicativi del fatto che il modello precedentemente creato non è sufficientemente robusto. Il motivo di ciò è da associare alla scelta iniziale di considerare tre parametri con p-value poco maggiore di 0,05. Il passo successivo è stato perciò quello di eliminare tali parametri. In questo modo, il modello ricavato è più semplice del precedente, e naturalmente presenta dei valori di intercetta e di *slope* differenti. Difatti il modello ha evidenziato come variabili rilevanti solo il PPQ ed il vAm permettendo di ridurre complessità e costo computazione.

<u>Comparison</u>	<u>Algorithm</u>	<u>Features selected</u>
Parkinsonian-Healthy	Logistic regression	PPQ,vAm

Tabella 11 Variabili rilevanti per l'algoritmo regressione logistica

La funzione di probabilità ricavata per il modello di regressione logistica è la seguente:

$$P(\text{Healthy} = 1) = \frac{e^{(2,37-0,99*PPQ-0,09*vAm)}}{1 + e^{(2,37-0,99*PPQ-0,09*vAm)}}$$

A questo punto, per verificarne le rispettive specifiche si è scelta la soglia di *cut-off* grazie al grafico riportante la *sensitivity* e *specificity* in funzione del valore di probabilità (figura 24).

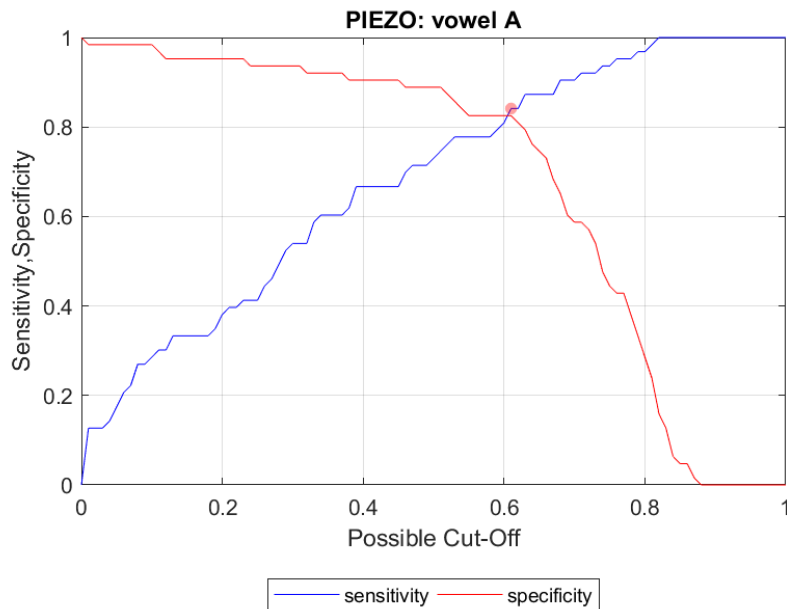


Figura 24 Scelta del possibile cut-off per il nuovo modello di regressione logistica ricavato dall'analisi del segnale vocale relativo alle vocali sostenute (microfono a contatto)

Anche in questo caso, la soglia scelta non corrisponde al punto di incontro esatto tra la *sensitivity* e la *specificity*. Infatti, spostandosi in un piccolo intorno destro del punto di incontro si nota un aumento della *sensitivity*. La soglia, che ha permesso di discriminare soggetti sani da soggetti Parkinsoniani, è stata fissata a 0,61. La curva ROC del modello in esame, che ha permesso di ricavare l'*area under curve* viene riportata nella figura 25.

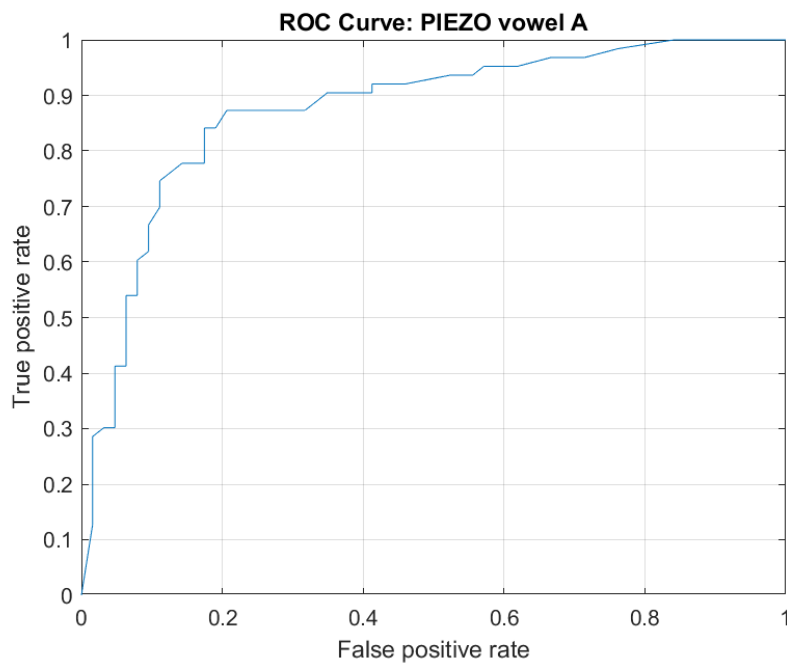


Figura 25 Curva ROC ricavata dal nuovo modello di regressione logistica ricavato dall'analisi del segnale vocale relativo alle vocali sostenute (microfono a contatto)

Servendosi dello stesso script Matlab utilizzato in precedenza, si è ottenuto un vettore relativo alla classe predetta dal modello. Confrontato tale vettore con il vettore della classe reale si è ricavata la *confusion matrix* sottostante:

		ACTUAL VALUE	
		Positive	Negative
CLASSIFIER PREDICTION	Positive	53	11
	Negative	10	52

Tabella 12 Confusion matrix relativa all'analisi di tutte le vocali /a/ mediante la regressione logistica

Tale Confusion matrix, assieme alla curva ROC riportata nella figura 25 ha permesso di esprimere le seguenti specifiche:

- Accuracy $\rightarrow 83,33\% \approx 83\%$
- Sensitivity $\rightarrow 84,13\% \approx 84\%$
- Specificity $\rightarrow 82,54\% \approx 83\%$
- AUC=0,87.

Dalle specifiche si nota che, rispetto al modello precedente, si ha un aumento della *sensitivity*, a discapito dell'*accuracy* dell'AUC e della *specificity*. Tuttavia, considerando la diminuzione non rilevante dei tre parametri sopra esposti, e considerando anche la diminuzione della complessità del modello oltre che all'aumento della *sensitivity*, si è preferito considerare quest'ultimo modello rispetto al precedente. La *confusion matrix* ottenuta può essere meglio esplicitata dalla figura 26 che mostra la classificazione di tutte le vocali /a/ secondo il classificatore appena creato.

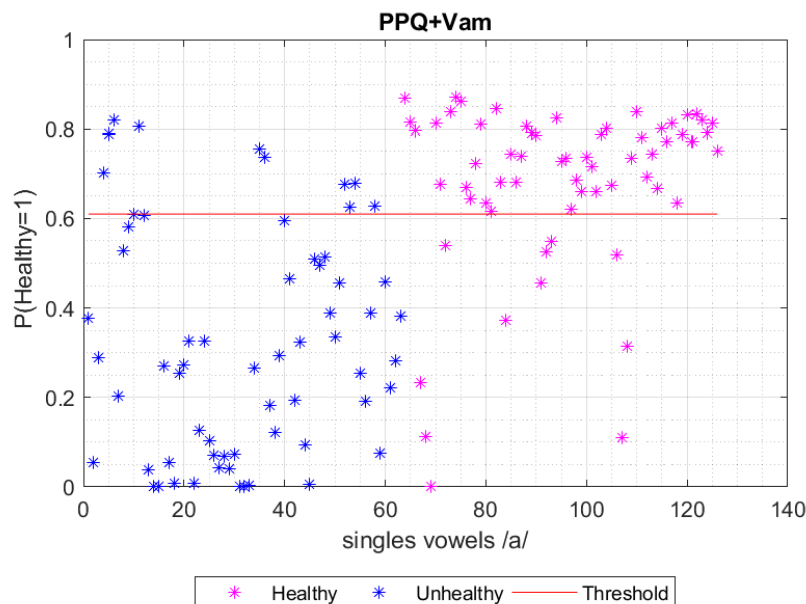


Figura 26 Grafico di separazione relativo al nuovo modello di regressione logistica ricavato dall'analisi delle singole vocali /a/

A questo punto, così come si è fatto per il modello di regressione logistica scartato, si può andare a calcolare la media tra i parametri relativi alle tre vocali /a/ sostenute dallo stesso paziente. In tal modo, si è ottenuta una matrice con un numero di righe pari al numero di soggetti in esame, 21, ed un numero di colonne pari al numero di parametri ricavati. Dando questa matrice in pasto alla funzione di probabilità precedentemente trovata e considerata la soglia scelta in precedenza è stata effettuata una “validazione”. La *confusion matrix* ricavata considerando la media tra i parametri delle tre vocali /a/ viene di seguito riportata.

		ACTUAL VALUE	
		Positive	Negative
CLASSIFIER PREDICTION	Positive	18	3
	Negative	3	18

Tabella 13 *Confusion matrix* relativa all'analisi media tra le 3 vocali /a/ mediante la regressione logistica

Tale *confusion matrix* ricavata ha permesso di esprimere le seguenti specifiche:

- Accuracy → 85,71% $\approx$ 86%
- Sensitivity → 85,71% $\approx$ 86%
- Specificity → 85,71% $\approx$ 86%

Dalle specifiche ricavate si può quindi concludere che, eliminando i parametri aventi un p-value poco maggiore di 0,05, il modello ottenuto è risultato essere consistente. La figura 27 mostra come il modello appena descritto riesca a distinguere un soggetto sano da un soggetto parkinsoniano basandosi sui parametri medi delle tre vocali /a/.

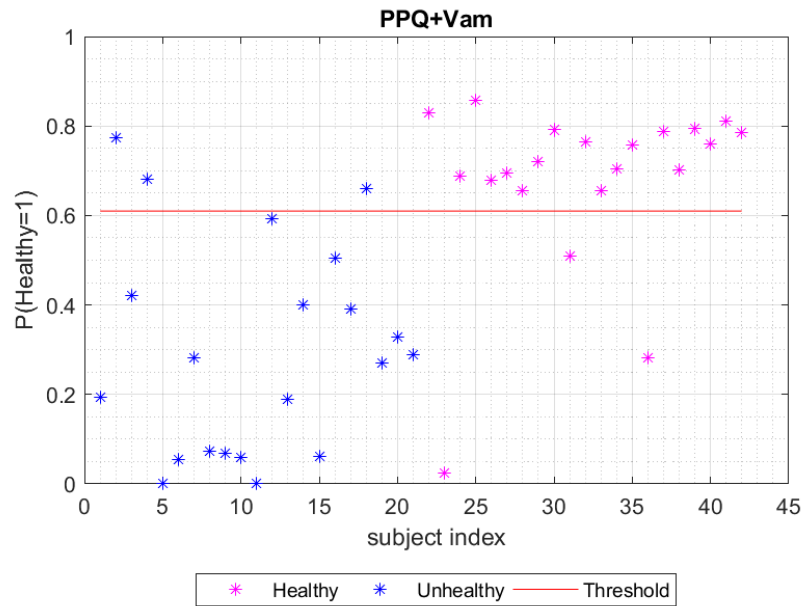


Figura 27 Grafico di separazione, relativo al nuovo modello di regressione logistica, ricavato dall'analisi della media delle vocali

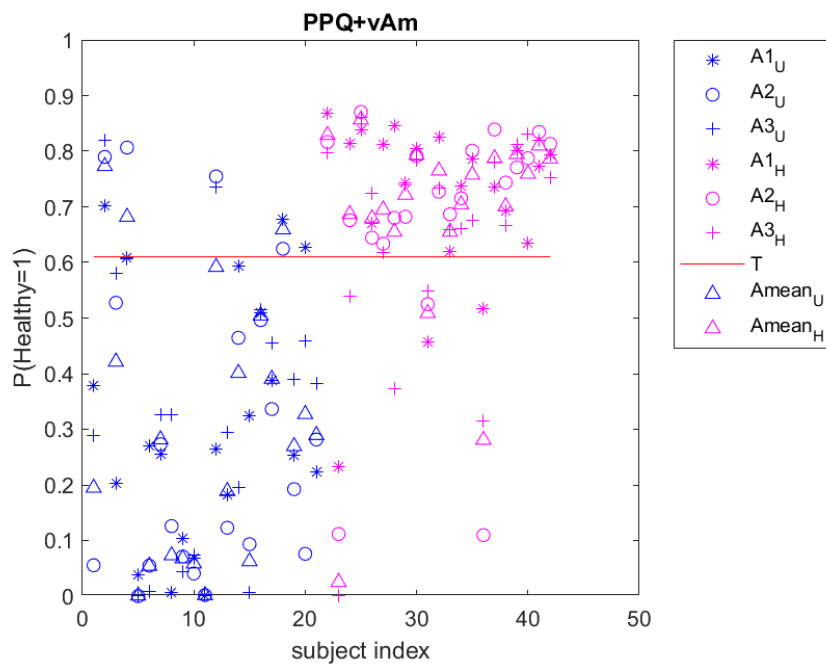


Figura 28 Grafico di separazione complessivo relativo al nuovo modello di regressione logistica

Al fine di verificare una possibile correlazione tra l'errore introdotto dal modello nella classificazione dei soggetti come sani o parkinsoniani in base all'analisi dei parametri relativi alle singole vocali /a/, e l'errore introdotto dal modello nella classificazione dei suddetti soggetti in base all'analisi dei parametri medi tra le

vocali stesse, si è ricavato il grafico mostrato in figura 28. Tale grafico riporta con il colore blu la probabilità ricavata dal modello per i soggetti parkinsoniani, mentre con il colore magenta la probabilità ricavata per i soggetti sani. Sull'ascissa di tale grafico troviamo il numero totale di soggetti in esame (42), sull'ordinata il valore di probabilità. Per ogni valore di ascissa, quindi per ogni soggetto, è riportata la probabilità associata all'analisi delle singole vocali /a/, A1 (asterisco), A2 (cerchietto), A3 (croce), ed anche la probabilità associata all'analisi dei parametri medi delle tre A; tale probabilità viene rappresentata con il simbolo di un triangolo. Da tale grafico si evince che, considerando la media dei parametri, il modello classifica erroneamente i segnali vocali dei soggetti sani solo quando ci sta un errore di classificazione in tutte le vocali /a/ considerate singolarmente. Lo stesso modello, considerando sempre i parametri medi, classifica erroneamente i soggetti parkinsoniani solo quando il modello commette già almeno due errori sulle singole vocali /a/.

4.2 Lettura del brano (microfono a contatto)

In questa sezione si andranno ad analizzare i risultati ottenuti dall'elaborazione del segnale vocale derivante dalla registrazione mediante microfono a contatto della lettura del brano. Tale analisi, come fatto per l'analisi delle vocali, verrà effettuata mediante i due diversi algoritmi *coarse tree* e regressione logistica; entrambi basati su di una prima analisi di correlazione dei parametri ottenuti mediante la matrice di correlazione.

4.2.1 Matrice di correlazione

In questo caso non essendo una vocale sostenuta non sono stati calcolati i parametri di perturbazione. I parametri estratti sono relativi al CPPS (evidenziati in grigio) ad eccezione della F_0 , dell'HNR, e del rapporto *voiced/unvoiced*. Anche in questo caso si è ripotata solo la prima metà superiore della matrice e si sono eliminati gli elementi con un corrispettivo valori di p-value maggiore di 0,05. Tale matrice ha

evidenziato una correlazione tra l'HNR e il parametro *voiced/unvoiced*, ma soprattutto una forte correlazione tra la media e tutti gli altri parametri relativi al CPPS ad eccezione della deviazione tipo, del *range* e della *kurtosis*. In questo caso, a differenza dell'analisi del segnale vocale relativo alle vocali sostenute, la *kurtosis* e la *skewness* sono risultate essere correlate.

	F0	HNR	V/Uv	CPPS								
				mean	median	mode	std	range	5prc	95prc	skewness	kurtosis
F0	1	0,46			0,32		-0,42		0,33		-0,49	0,52
HNR		1	0,86	0,58	0,59	0,58	-0,48		0,57	0,37	-0,75	0,73
V/Uv			1	0,51	0,51	0,48			0,53	0,37	-0,59	0,56
mean				1	0,99	0,92		0,61	0,94	0,90	-0,89	0,78
median					1	0,93		0,59	0,92	0,90	-0,90	0,77
mode						1		0,53	0,80	0,85	-0,88	0,71
std							1	0,35	-0,41		0,41	-0,67
range								1	0,49	0,74	-0,32	
5prc									1	0,78	-0,82	0,84
95prc										1	-0,68	0,50
skewness											1	-0,87
kurtosis												1

Tabella 14 Matrice di correlazione relativa ai parametri calcolati dalla lettura del brano acquisto con il microfono a contatto

4.2.2 Coarse tree

Il passo successivo è stato quello di ottenere un modello mediante l'utilizzo dell'algoritmo *coarse tree*. Tale algoritmo, nel caso specifico in esame, ha permesso di andare ad identificare come variabili rilevanti quelle riportate nella seguente tabella.

<u>Comparison</u>	<u>Algorithm</u>	<u>Features selected</u>
Parkinsonian-Healthy	Coarse tree	CPPSrange, CPPS95prc, CPPSkurtosis

Tabella 15 Variabili rilevanti per l'algoritmo *coarse tree*

Come si nota dalla matrice di correlazione, tali parametri sono tra loro non correlati e quindi non stati giustamente utilizzati parametri ridondanti. Anche in questo caso l'albero decisione ricavato è stato sostituito dal corrispettivo *flow-chart*.

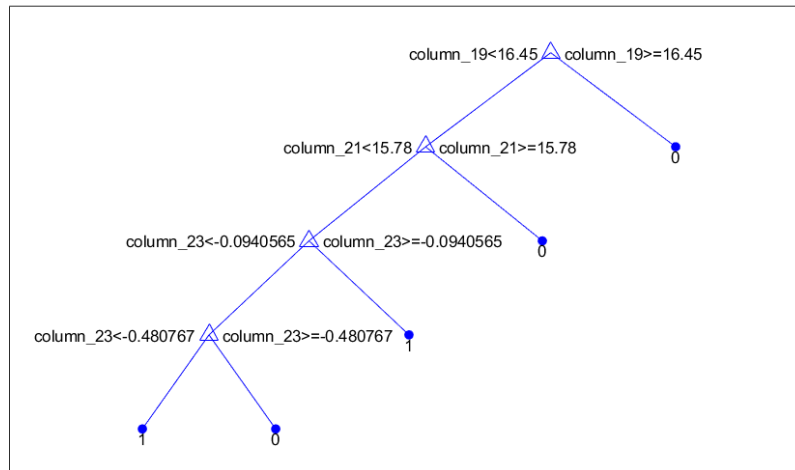


Figura 29 Albero Decisionale relativo alla lettura del brano (microfono a contatto)

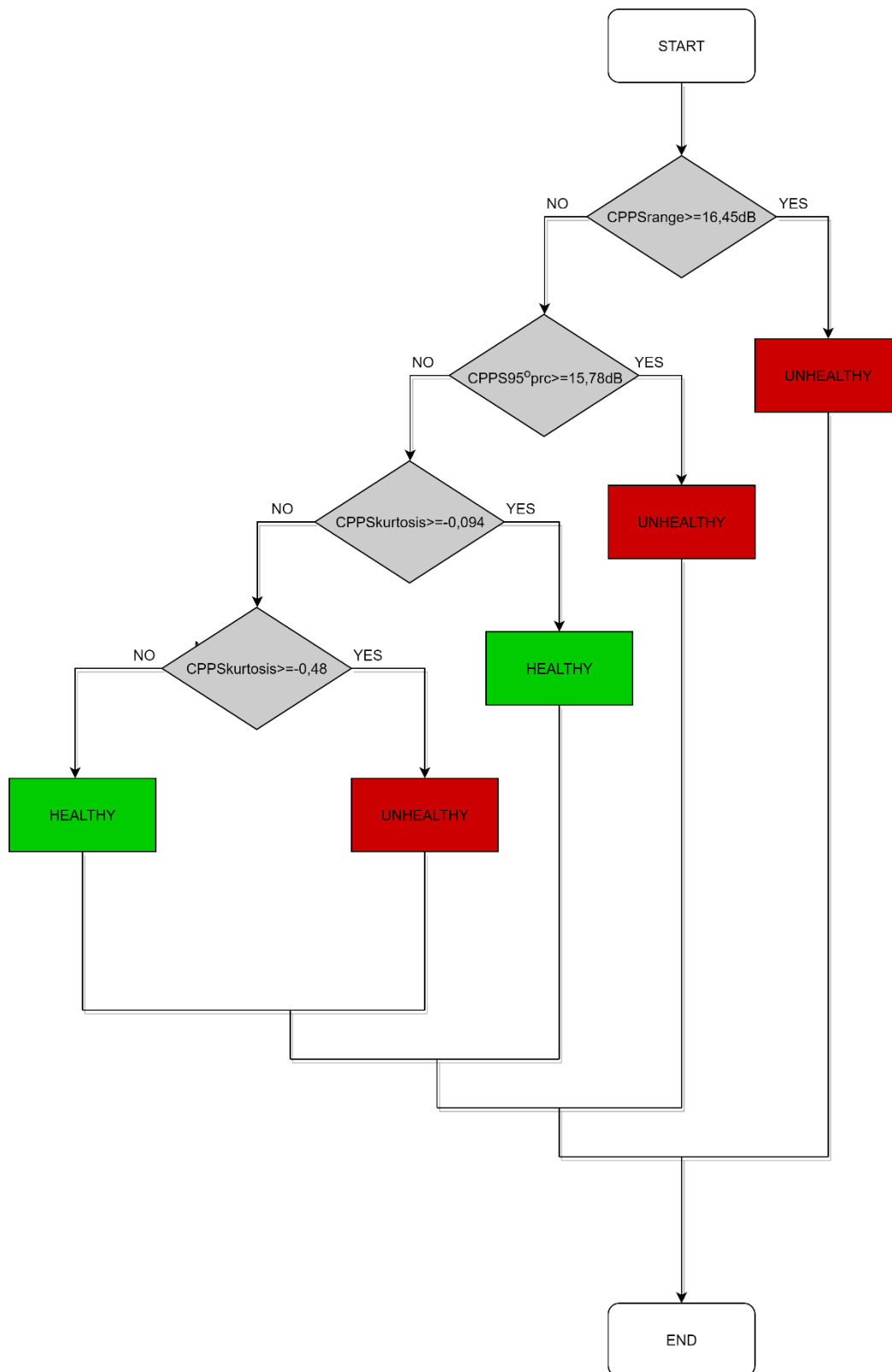


Figura 30 Flow-chart relativo alla lettura del brano (microfono a contatto)

Dal *flow-chart* sopra riportato si evince che la radice è rappresentata dal *range* che mostra un valore di soglia pari a 16,45dB; tale parametro, per ogni segnale vocale

studiato, può presentare valori maggiori, uguali o inferiori al valore di soglia. In particolare, se assume valori maggiori o uguali a 16,45dB il soggetto in esame viene diagnosticato come parkinsoniano. Al contrario, se il *range* calcolato per il segnale vocale del soggetto in esame presenta un valore minore del valore di soglia si passa a verificare il valore assunto dal CPPS_{95prc}. Tale parametro, a sua volta, presenta una soglia pari a 15,78dB. Perciò, valori maggiori permettono di concludere che il soggetto appartiene alla classe dei parkinsoniani, valori inferiori impongono invece una ulteriore verifica basata sulla *kurtosis*. Quest'ultima, infatti, permette di identificare un soggetto come sano se il suo valore è maggiore o uguale a -0,094. Altrimenti, la stessa *kurtosis* discriminerà un soggetto come sano se il segnale vocale ad esso associato presenta un valore di tale parametro inferiore a -0,48 o come Parkinsoniano se il valore di tale parametro è maggiore o uguale a tale soglia. Seguendo il seguente *flow-chart* per ogni segnale vocale analizzato, l'algoritmo ha associato ciascun soggetto come appartenente alla classe dei sani (1) o alla classe dei Parkinsoniani (0). Ottenuta la classe predetta dal modello, e nota a priori la classe reale, mediante un apposito script Matlab si è calcolata la seguente *confusion matrix*:

		ACTUAL VALUE	
		Positive	Negative
CLASSIFIER PREDICTION	Positive	17	0
	Negative	2	19

Tabella 16 Confusion matrix relativa alla lettura del brano ottenuta con l'algoritmo coarse tree (microfono a contatto)

Tale Confusion Matrix ricavata permette di esprimere le seguenti specifiche:

- Accuracy $\rightarrow 94,74\% \approx 95\%$
- Sensitivity $\rightarrow 89,47\% \approx 89\%$
- Specificity $\rightarrow 100\%$

4.2.3 Regressione logistica

I parametri individuati con l'ausilio del *parallel coordinates plot* e della matrice di correlazione hanno permesso di ricavare il modello di regressione logistica per l'analisi del segnale vocale relativo alla lettura del brano. In questo, le variabili rilevanti sono risultate essere il rapporto *voiced/unvoiced* ed il *CPPSrange*.

<u>Comparison</u>	<u>Algorithm</u>	<u>Features selected</u>
Parkinsonian-Healthy	Logistic regression	Voiced/Unvoiced,CPPSrange

Tabella 17 Variabili rilevanti per l'algoritmo regressione logistica

Attraverso la *tool* di Matlab, come esposto nel paragrafo 3.9, si sono ricavati i parametri del modello che hanno permesso di esplicitare la formula indicante la probabilità di far appartenere un soggetto alla classe dei sani. La formula del modello in esame risulta essere la seguente:

$$P(\text{healthy} = 1) = \frac{e^{(18,27-0,07*V/Uv-0,77*CPPSrange)}}{1 + e^{(18,27-0,07*V/Uv-0,77*CPPSrange)}}$$

Anche in questo caso, per poter scegliere la soglia che permetta di convertire il valore di probabilità *i*-esima ad un corrispettivo valori output dicotomico, ci si è basati sul seguente grafico:

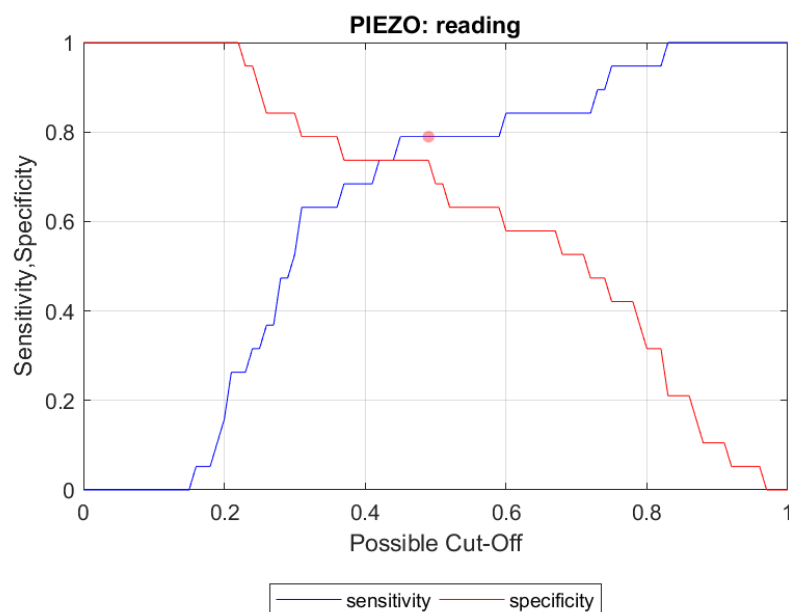


Figura 31 Scelta del possibile cut-off per il modello di regressione logistica ricavato dall'analisi del segnale vocale relativo alla lettura del brano (microfono a contatto)

Non si è scelta come soglia il valore di incontro tra le due curve in quanto, come si può notare dal grafico, spostando il *cut-off* verso valori superiori la *specificity* resta costante e la *sensitivity* assume valori superiori. Come soglia si è quindi scelto un valore di probabilità pari a 0,49 con l'intento di massimizzare le prestazioni. In figura 32 viene riportata la curva ROC che ha permesso di calcolare l'AUC del modello creato.

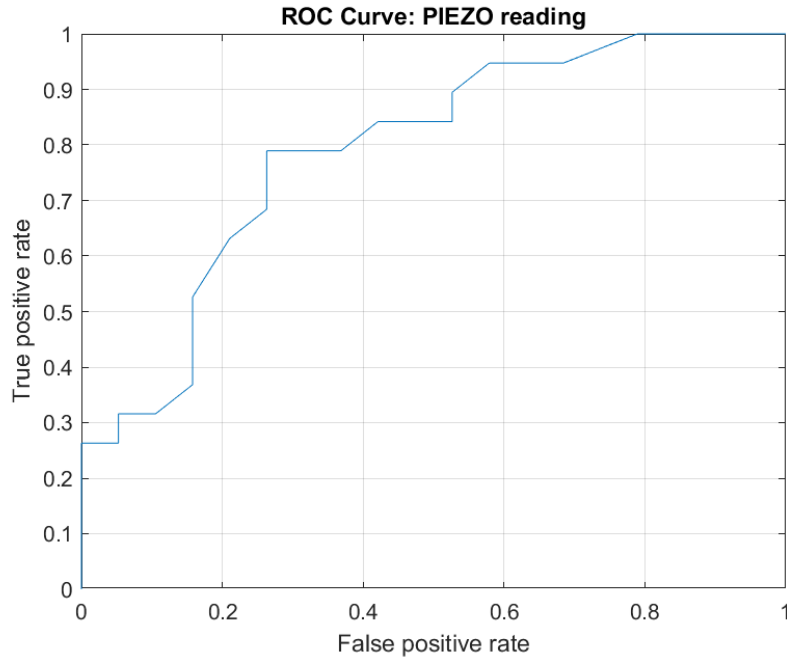


Figura 32 Curva ROC ricavata dal modello di regressione logistica ricavato dall'analisi del segnale vocale relativo alla lettura del brano (microfono a contatto)

Attraverso lo stesso script Matlab utilizzato per i casi precedenti, si è andati a confrontare il valore di probabilità ottenuto con la soglia scelta. Se la probabilità è risultata essere maggiore a 0,49 il soggetto è stato classificato come sano (classe 1), al contrario, se la probabilità è risultata essere inferiore o uguale a 0,49 il soggetto è stato associato alla classe dei soggetti parkinsoniani (classe 0). Ottenuto il vettore contenente la “classe predetta” è stato effettuato il confronto con il vettore contenente la classe reale. Da tale confronto si è costruita la seguente *confusion matrix*:

		ACTUAL VALUE	
		Positive	Negative
CLASSIFIER PREDICTION	Positive	15	5
	Negative	4	14

Tabella 18 Confusion matrix relativa al modello di regressione logistica per la lettura del brano (microfono a contatto)

La *confusion matrix* assieme alla ROC curve ricavata hanno permesso di esplicitare le seguenti specifiche:

- Accuracy=76,32% $\approx$ 76%
- Sensitivity=78,95% $\approx$ 79%
- Specificity=73,68% $\approx$ 74%
- AUC=0,79

Come si nota dalle specifiche il modello creato commette degli errori, infatti sono presenti sia dei falsi positivi che dei falsi negativi. Per meglio vedere come si comporta il modello in esame viene qui di seguito riportato un grafico che, in base ai parametri scelti ed in base alla soglia precedentemente fissata, mostra la separazione tra le due classi.

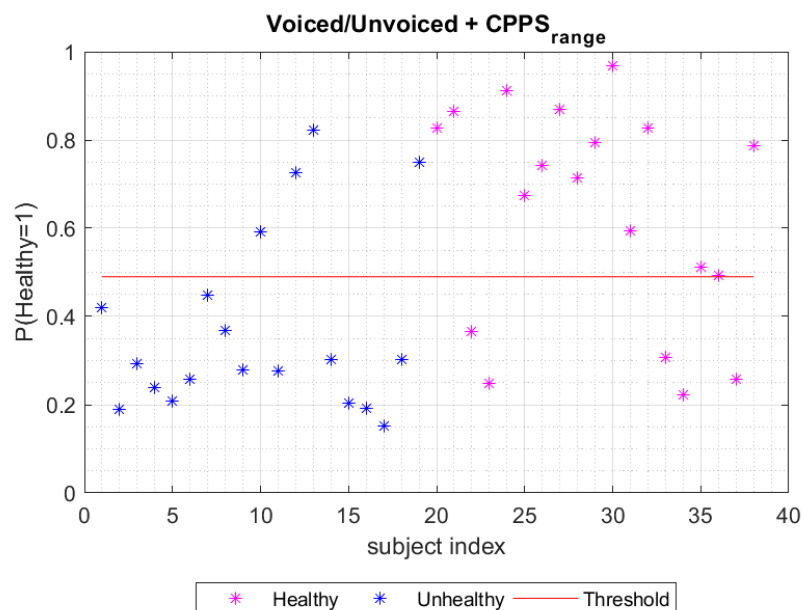


Figura 33 Grafico di separazione relativo al modello di regressione logistica ricavato dall'analisi della lettura del brano (microfono a contatto)

4.3 Lettura del brano (microfono in aria)

In questa sezione si analizzeranno i risultati ottenuti dall'elaborazione del segnale vocale relativo alla lettura del brano acquisito con il microfono in aria. Tale analisi, come fatto sin ora, verrà effettuata mediante i due algoritmi *coarse tree* e regressione logistica, entrambi sempre basati su di una prima analisi di correlazione ottenuta mediante la matrice di correlazione.

4.3.1 Matrice di correlazione

Anche in questo caso, non essendo una vocale sostenuta, non sono stati calcolati i parametri di perturbazione. Tutti parametri, ad eccezione della F_0 dell'HNR e del rapporto *voiced/unvoiced*, sono relativi al dominio *cepstrale*. La matrice di correlazione, riportante solo la metà diagonale superiore dei soli valori presentati p-value inferiori a 0,05, evidenzia una forte correlazione tra la media e la mediana. Entrambi i parametri, oltre che ad essere fortemente correlati tra loro, presentano correlazioni con tutti gli altri parametri relativi al CPPS ad eccezione della moda, della deviazione standard e del range.

	F0	HNR	V/Uv	CPPS								
				mean	median	mode	std	range	5prc	95prc	skewness	kurtosis
F0	1	0,42										
HNR		1	0,70								-0,42	
V/Uv			1									
mean				1	0,99	0,71	-0,49	0,50	0,94	0,92	-0,84	0,82
median					1	0,75	-0,46	0,47	0,90	0,91	-0,88	0,81
mode						1	-0,43	0,33	0,66	0,58	-0,78	0,53
std							1		-0,66		0,50	-0,76
range								1	0,46	0,57		0,33
5prc									1	0,79	-0,75	0,86
95prc										1	-0,68	0,65
skewness											1	-0,76
kurtosis												1

Tabella 19 Matrice di correlazione relativa ai parametri calcolati dalla lettura del brano acquisto con il microfono in aria

4.3.2 Coarse tree

Il primo classificatore è stato ottenuto mediante l'utilizzo dell'algoritmo *coarse tree*. Tale algoritmo nel caso specifico in esame ha permesso di andare ad identificare come variabili rilevanti quelle riportate nella tabella 20.

<u>Comparison</u>	<u>Algorithm</u>	<u>Features selected</u>
Parkinsonian-Healthy	Coarse tree	CPPSstd, CPPS5prc, HNR

Tabella 20 Variabili rilevanti per l'algoritmo coarse tree

L'albero decisionale ottenuto mediante la *toll* di Matlab *Classification learner* per l'analisi del segnale vocale relativo alla lettura del brano acquisito microfono in aria è il seguente:

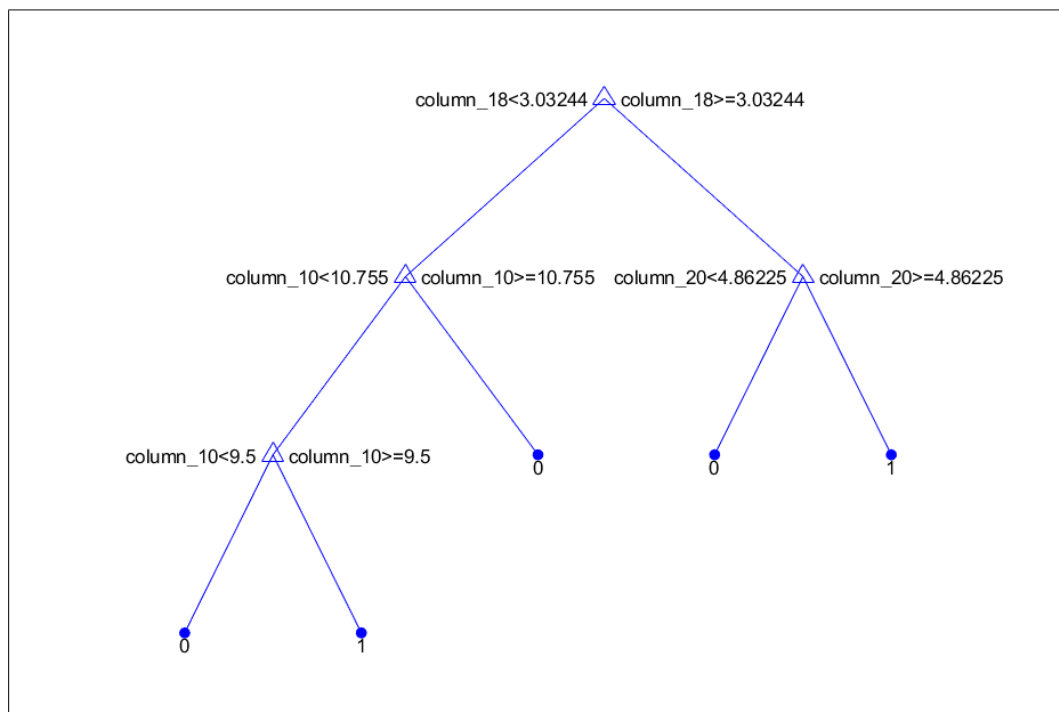


Figura 34 Albero Decisionale relativo alla lettura del brano (microfono in aria)

Dal diagramma ad albero riportato in figura 34 si è passati al corrispettivo diagramma di flusso riportato in figura 35.

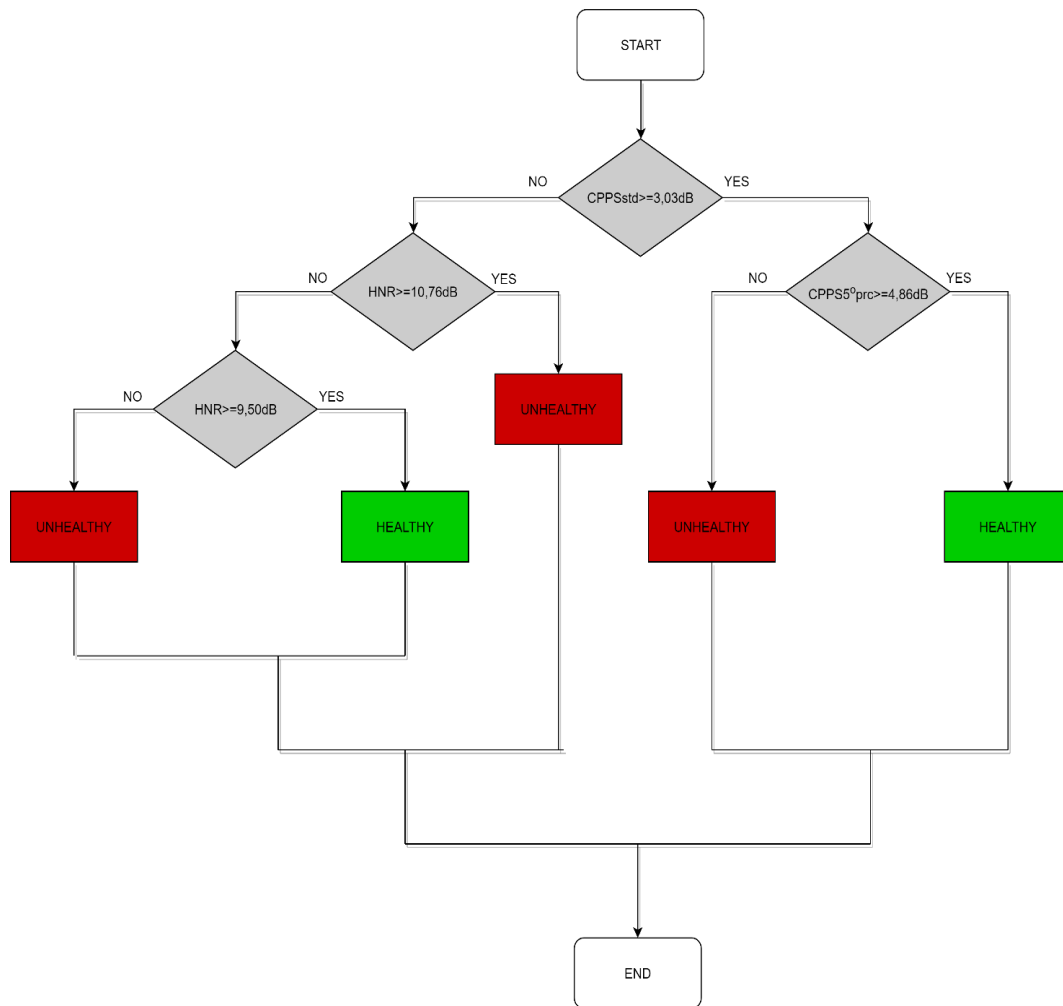


Figura 35 Flow-chart relativo alla lettura del brano (microfono in aria)

In questo caso la radice dell'albero è rappresentata dalla $CPPS_{std}$, e tale radice mostra una soglia pari a 3,03dB. Perciò, se tale parametro per il segnale vocale analizzato risulta essere maggiore o uguale a tale valore di soglia si passa a confrontare il $CPPS_{5prc}$ che permette di diagnosticare un soggetto come sano o parkinsoniano sulla base di un valore di soglia fissato a 4,86dB. In particolari, valori maggiori o uguali a tale valore permettono di classificare un soggetto come sano, al contrario di valori inferiori che classificano il soggetto come parkinsoniano. Se invece, la $CPPS_{std}$ del segnale vocale in esame è inferiore al valore di soglia si effettua una valutazione dell'HNR che, se maggiore-uguale a 10,76dB consente di classificare un soggetto come Parkinsoniano, al contrario di un valore inferiore che

impone una ulteriore verifica sempre sullo stesso parametro. Infatti, il segnale vocale studiato per raggiungere una classificazione come segnale vocale appartenente ad un soggetto sano deve presentare un valore di HNR maggiore di 9,50dB, in caso contrario viene classificato come parkinsoniano. Seguendo tale *flow-chart* per ogni segnale vocale analizzato e confrontando la classe reale con la classe predetta dal modello si è ricavata la seguente *confusion matrix*:

		ACTUAL VALUE	
		Positive	Negative
CLASSIFIER PREDICTION	Positive	21	4
	Negative	1	18

Tabella 21 Confusion matrix relativa alla lettura del brano ottenuta con l'algoritmo coarse tree (microfono in aria)

La *confusion matrix* sopra riportata permette di andare ad esprimere le seguenti specifiche:

- Accuracy $\rightarrow 88,64\% \approx 89\%$
- Sensitivity $\rightarrow 95,45\% \approx 95\%$
- Specificity $\rightarrow 81,82\% \approx 82\%$

4.3.3 Regressione logistica

Grazie ai parametri identificati mediante l'utilizzo del *parallel coordinates plot* e della matrice di correlazione, si è utilizzato il modello di regressione logistica. Per questa specifica fase del protocollo, le variabili rilevanti sono risultate essere l'HNR ed il CPPS_{median}.

<u>Comparison</u>	<u>Algorithm</u>	<u>Features selected</u>
Parkinsonian-Healthy	Logistic regression	HNR,CPPSmedian

Tabella 22 Variabili rilevanti per l'algoritmo regressione logistica

Attraverso la stessa *tool* di Matlab, si sono ricavati i parametri del modello che hanno permesso di esplicitare la formula indicante la probabilità di far appartenere un soggetto alla classe dei sani (classe 1).

La formula del modello in esame risulta essere la seguente:

$$P(\text{Healthy} = 1) = \frac{e^{(-2,06-0,45*HNR+0,57*CPPSmedian)}}{1 + e^{(-2,06-0,45*HNR+0,57*CPPSmedian)}}$$

Anche in quest'ultimo caso per scegliere la soglia si è graficata la variazione della *sensitivity* e *specificity* in funzione del possibile valore di *cut-off*, come riportato dal grafico seguente:

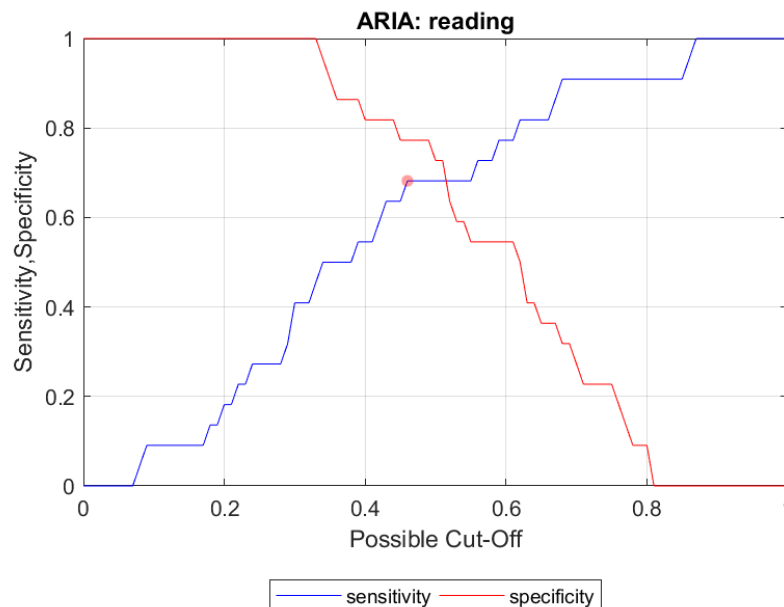


Figura 36 Scelta del possibile cut-off per il modello di regressione logistica ricavato dall'analisi del segnale vocale relativo alla lettura del brano (microfono in aria)

Come per tutti i casi riportati in questo lavoro di tesi, la scelta della soglia non si è basata sul punto di incontro delle due curve. In quanto, come si può notare dalla figura sopra riportata, spostando il *cut-off* verso un intorno inferiore la *sensitivity* rimane costante ad un valore prossimo a 0,7 e la *specificity* assume un valore prossimo a 0,8. Come soglia si è perciò scelto un valore di probabilità pari a 0,46. La curva ROC del modello in esame viene riportata in figura 37.

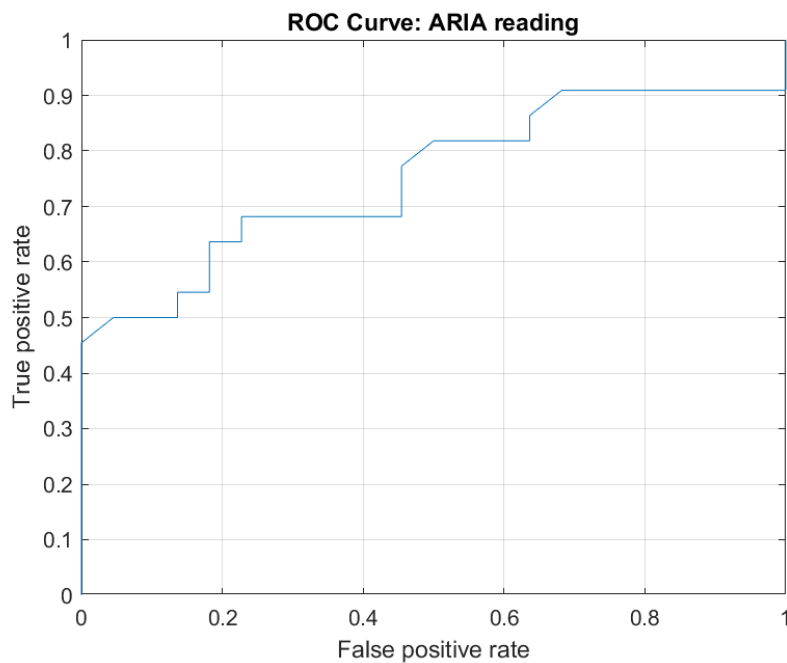


Figura 37 Curva ROC ricavata dal modello di regressione logistica ricavato dall'analisi del segnale vocale relativo alla lettura del brano (microfono in aria)

Attraverso lo script Matlab appositamente implementato, si è confrontato il valore di probabilità ottenuto con la soglia scelta. In questo caso se la probabilità è risultata maggiore a 0,46 il soggetto è stato classificato come sano (classe 1), al contrario se la probabilità è risultata inferiore o uguale a 0,46 il soggetto è stato associato alla classe dei soggetti Parkinsoniani (classe 0). A questo punto, ottenuto un vettore contenente la classe predetta lo si è confrontato con il vettore contenente la classe reale. Tale confronto ha permesso di costruire la *confusion matrix* riportata in tabella 23.

		ACTUAL VALUE	
		Positive	Negative
CLASSIFIER PREDICTION	Positive	15	5
	Negative	7	17

Tabella 23 Confusion matrix relativa al modello di regressione logistica per la lettura del brano (microfono in aria)

La *confusion matrix* assieme alla curva ROC hanno permesso di esprimere le seguenti specifiche:

- Accuracy $\rightarrow 72,73\% \approx 73\%$
- Sensitivity $\rightarrow 68,18\% \approx 68\%$
- Specificity $\rightarrow 77,27\% \approx 77\%$
- AUC=0,75

Come si evince dalle specifiche il modello creato commette degli errori, infatti sono presenti sia dei falsi positivi che dei falsi negativi. Il grafico di figura 38, a conferma della *confusion matrix* riportata, mostra che il modello classifica erroneamente come sani sette soggetti parkinsoniani (falsi negativi), e come parkinsoniani cinque soggetti sani (falsi positivi).

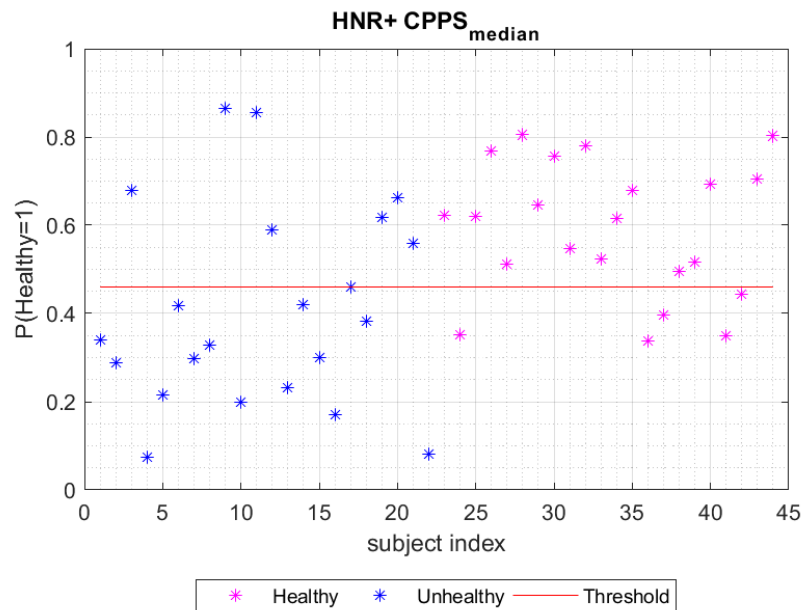


Figura 38 Grafico di separazione relativo al modello di regressione logistica ricavato dall'analisi della lettura del brano (microfono in aria)

5. VALUTAZIONE CORSO LOGOPEDICO

Al fine di valutare la possibile efficacia di un trattamento logopedico nei pazienti affetti dalla malattia di Parkinson, nella seconda parte di questo lavoro di tesi si sono analizzate le caratteristiche temporali, frequenziali e *cepstrali* del segnale vocale all'inizio ed alla fine di un percorso di logopedia. I soggetti in esame sono stati registrati con la stessa catena di acquisizione e seguendo lo stesso protocollo esposto nel Capitolo 3. In questo caso, le registrazioni sono state effettuate presso il centro "associazione amici parkinsoniani Piemonte Onlus" di Torino, dove i soggetti volontari si sono recati settimanalmente, per un periodo complessivo di circa 2 mesi. Per poter evidenziare le possibili differenze in termini di parametri relativi al segnale vocale i sei soggetti (una donna e cinque uomini) sono stati acquisiti all'inizio del trattamento (ottobre 2018) ed alla fine del trattamento (dicembre 2018). Anche in questo caso, l'analisi è stata condotta al fine di

effettuare una distinzione in termini di microfono (contatto, aria) oltre che in termini di specifico segnale vocale registrato. Come per la prima parte del lavoro di tesi, è stato analizzato il segnale vocale relativo alle tre vocali /a/ registrate con il microfono a contatto, ed il segnale vocale relativo alla lettura del brano foneticamente bilanciato utilizzando in questo caso entrambi i microfoni.

5.1 Analisi della vocale /a/

Per l'analisi del segnale vocale relativo alla registrazione delle tre vocali /a/ sostenute, è stato possibile ricavare i parametri di perturbazione oltre che quelli relativi al CPPS. Poiché ogni soggetto anche in questo caso ha ripetuto per tre volte la vocale /a/, si sono ottenuti tre segnali vocali (A1, A2, A3) per ogni paziente. A questo punto, da ogni segnale vocale sono stati calcolati tutti i parametri di perturbazione, utilizzando gli stessi script Matlab adoperati nella prima parte del lavoro di tesi. Al fine di rendere visibile una differenza complessiva fra l'inizio e la fine del trattamento logopedico, si è deciso di calcolare la media tra i parametri delle tre A. In questo modo, è stato possibile riportare in dei *bar diagrams* la differenza dei parametri di interesse pre e post trattamento evidenziandone anche lo scarto tipo. Anche in questo caso, non tutti i parametri calcolati sono risultati tra loro non correlati. Perciò, al fine di riportare esclusivamente le variazioni relative ai parametri non ridondanti, si è calcolata la matrice di correlazione che ha permesso di decidere su quali parametri concentrarsi.

	Jitt	Jita	RAP	PPQ	vF0	Shim	Shdb	APQ	vAm	F0	HNR	CPPS						skewness	kurtosis
	mean	median	mode	std	range	5prc	95prc												
Jitt	1	0.96	0.98	0.83	0.77	0.86	0.87	0.84	0.55		-0.86								
Jita		1	0.95	0.84	0.70	0.83	0.83	0.82	0.47		-0.81								
RAP			1	0.81	0.68	0.85	0.85	0.82	0.55		-0.81								
PPQ				1	0.65	0.68	0.68	0.68	0.41		-0.73								
vF0					1	0.64	0.66	0.70	0.53		-0.74								
Shim						1	1.00	0.98	0.63		-0.89								
ShdB							1	0.98	0.64		-0.90								
APQ								1	0.64		-0.92								
vAm									1		-0.66								
F0										1									
HNR											1								
mean												1	0.98	0.90	0.85	0.87	-0.32		
median													1	0.94	0.79	0.89	-0.39		
mode														1	0.74	0.81	-0.42		
std															1	0.83	-0.59	-0.28	
range																1	-0.40	-0.43	0.37
5prc																	1	-0.29	
95prc																		1	-0.41
skewness																			1
kurtosis																			

Tabella 24 Matrice di correlazione relativa ai parametri calcolati dalle vocali /a/ sostenute acquisite con il microfono a contatto

Da tale matrice è possibile notare che un'analisi esaustiva, per quanto riguarda i parametri di perturbazione, potrebbe essere basata sul *Jitt*, *vF0* e *vAm*. Tuttavia, al

fine di poter avere un confronto dei parametri di perturbazione sia a breve termine che a medio-lungo termine, si è deciso di riportare anche le variazioni riguardanti il PPQ e lo *Shim*. Ciò ha permesso di verificare l'andamento dei parametri di perturbazione ciclo a ciclo pre e post trattamento in termini di frequenza e di ampiezza con il Jitt e lo Shim; ma anche l'andamento dei parametri di perturbazione a medio-lungo termine in frequenza ed ampiezza con rispettivamente il PPQ e vAm. Per i parametri relativi al CPPS (evidenziati in grigio), si è considerata la moda la deviazione tipo, la *skewness* e la *kurtosis*. Dopo aver scelto il sottoinsieme dei parametri, sono stati confrontati i risultati ottenuti prima e dopo il trattamento logopedico, attraverso diversi *bar diagrams*. Ciascuno di essi è associato ad un parametro e per ogni soggetto sono state tracciate due barre. Quella blu rappresenta il valore medio dei parametri corrispondenti alle tre vocali / a / prima del trattamento, quella rossa rappresenta il valore medio dei parametri dopo due mesi di trattamento logopedico. Su ogni barra viene identificata la deviazione tipo. Una bassa deviazione tipo indica che i dati tendono ad essere vicini alla media dei due valori del parametro, mentre una deviazione tipo elevata indica che i dati sono distribuiti su un intervallo più ampio di valori.

5.1.1 Variazione dei parametri di perturbazione

In figura 39 è esposto il confronto relativo al *Jitt* (%). Il primo il quinto il terzo ed il sesto soggetto hanno evidenziato un miglioramento, infatti il valore di *Jitt* è diminuito soprattutto nei primi due. Ciò indica che tali soggetti hanno acquisito un controllo maggiore della stabilità frequenziale vocale dopo le sedute di logopedia. Il soggetto numero due invece, non ha riportato un cambiamento prestazionale, al contrario del numero quattro che ha mostrato un minimo peggioramento. Tuttavia, verificando la variazione di perturbazione frequenziale a medio termine attraverso il PPQ si nota un miglioramento in tutti i soggetti (figura 40). Nella figura 41 è invece descritto il confronto in termini di vF₀ (%). Il parametro vF₀ indica la variazione relativa della deviazione standard di F₀ rispetto alla frequenza fondamentale media calcolata. Anche per questo parametro i soggetti numero uno tre e sei hanno dimostrato un miglioramento delle prestazioni. Al contrario, i soggetti numero due e cinque hanno subito un aumento di vF₀. Il soggetto numero

quattro, che per il *Jitt* ha riportato un aumento, ha mostrato un minimo miglioramento per il parametro vF_0 . Nelle figure 42 e 43 sono evidenziati i parametri relativi alle perturbazioni in ampiezza attraverso lo *Shim* (%) e vAm (%). Per lo *Shim* si nota un miglioramento in tutti i soggetti, miglioramento che si evince anche nell'analisi di perturbazione a medio termine attraverso il vAm , eccetto che per il soggetto numero due. Sebbene i miglioramenti sono differenti da soggetto a soggetto, questi ultimi risultati sono indicativi del fatto che le sedute di logopedia, nella riproduzione delle vocali /a/ sostenute, hanno portato una maggiore stabilità in ampiezza ed in frequenza.

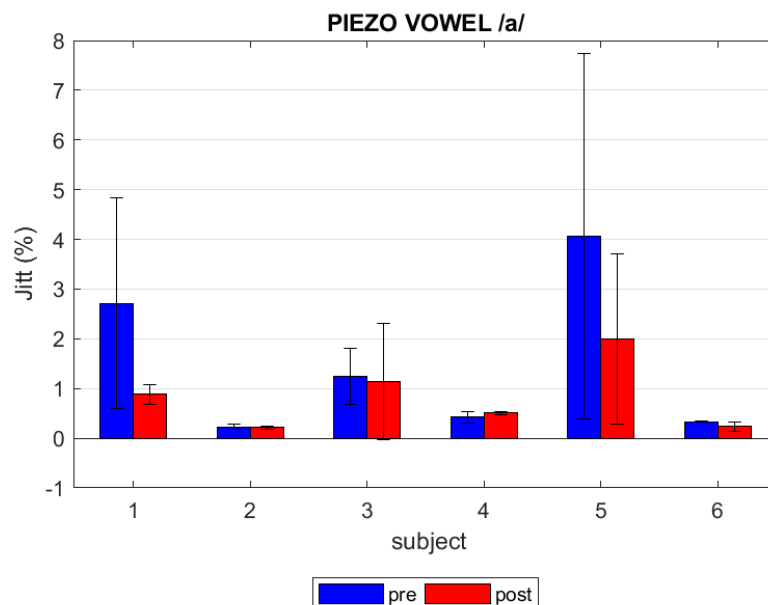


Figura 39 Variazione Jitt

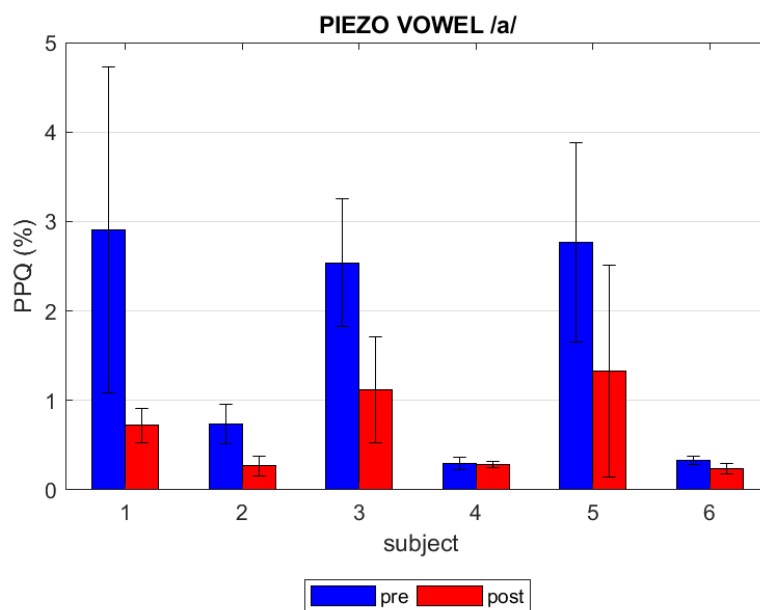


Figura 40 Variazione PPQ

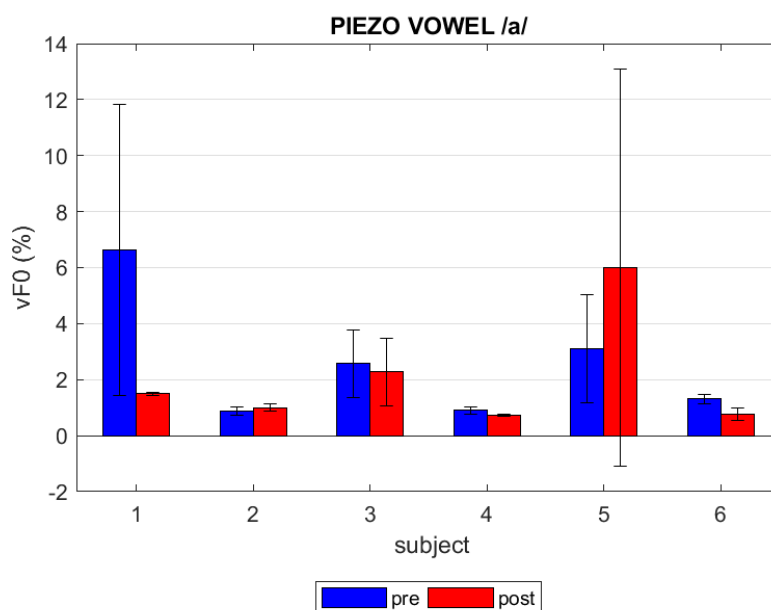


Figura 41 Variazione vF₀

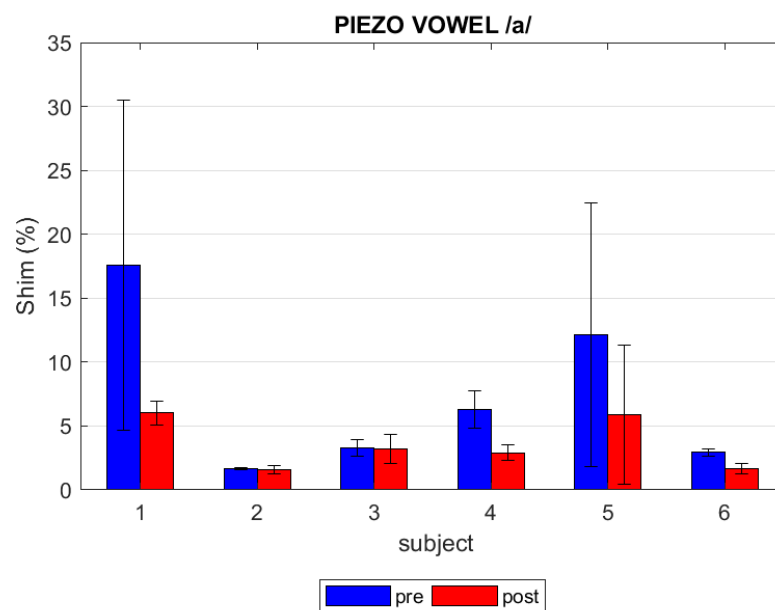


Figura 42 Variazione Shim

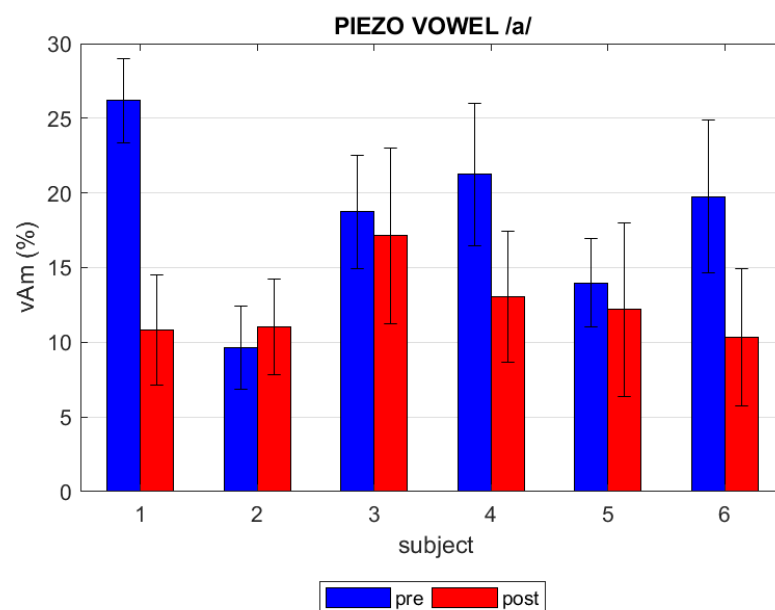


Figura 43 Variazione vAm

5.1.1 Variazione dei parametri cespstrali

Per ogni segnale vocale studiato, sono stati estratti anche i parametri *cepstrali* riportati nella matrice di correlazione. Tali parametri, sono stati ricavati dalle distribuzioni relative al CPPS. Per tal motivo, al fine di rendere maggiormente visibili le differenze pre e post trattamento sono state graficate le distribuzioni relative al CPPS, per ogni soggetto, prima e dopo trattamento logopedico (figura 44). In questo caso, per non riportare tre distribuzioni relative alle tre vocali /a/, si è scelto di riportare una distribuzione complessiva. Gli indici statistici estratti dalle distribuzioni non sono stati riportati tutti nei *bar diagrams* seguenti, ma solo quelli scelti per mezzo della matrice di correlazione. Dalle distribuzioni riportate in figura 44 si nota come, in tutti i casi eccetto per soggetto numero cinque, si ha diminuzione del $CPPS_{mean}$. Il soggetto che ha subito una maggior diminuzione dell'indice $CPPS_{mean}$ è risultato essere il numero due (figura 45). Un altro indice che ha subito una diminuzione è il $CPPS_{std}$, infatti le distribuzioni, eccetto quelle ricavate per il soggetto numero quattro, tendono tutte ad essere maggiormente concentrate intorno al valor medio dopo i due mesi di trattamento. A conferma di ciò, la figura 46 mostra i valori pre e post trattamento del $CPPS_{std}$ con il corrispettivo scarto tipo. Entrambi gli indici appena descritti, hanno permesso di verificare quale fosse la misura di localizzazione della distribuzione. Tuttavia, con lo scopo di rendere l'analisi maggiormente esaustiva, si sono riportati anche la $CPPS_{skewness}$ e la $CPPS_{kurtosis}$. Questi ultimi indici hanno permesso di andare a caratterizzare le forme delle distribuzioni stesse. Per quanto concerne la *skewness*, tutte le distribuzioni, eccetto quella relativa al soggetto numero tre ottenuta post trattamento, presentano una gobba verso destra ed una coda allungata verso sinistra. Tale andamento grafico è confermato dal *bar diagrams* relativo alla $CPPS_{skewness}$ (figura 47). Difatti, tale figura mostra come la *skewness* ha assunto sempre valori negativi ad eccezione del valore positivo assunto per il soggetto numero tre post trattamento. La forma delle distribuzioni viene anche descritta dalla *kurtosis* che assume, valori nulli quando le distribuzioni presentano andamento normale, valori maggiori di zero se presentano un maggiore allungamento, e valori inferiori a zero quando le distribuzioni evidenziano un maggiore appiattimento rispetto alla distribuzione normale. La descrizione quantitativa di tale andamento viene esplicitata nel *bar diagram* relativo

alla $CPPS_{kurtosis}$ (figura 48). Da tali *bar diagrams* si nota che, per le distribuzioni dei soggetti numero due tre e sei, la *kurtosis* ha assunto sempre un valore positivo sia prima che dopo il trattamento logopedico. Le distribuzioni associate ai soggetti numero uno e quattro invece hanno mostrato un'inversione passando da un valore positivo ad un valore negativo, al contrario della distribuzione relativa al soggetto numero cinque che ha subito un'inversione opposta.

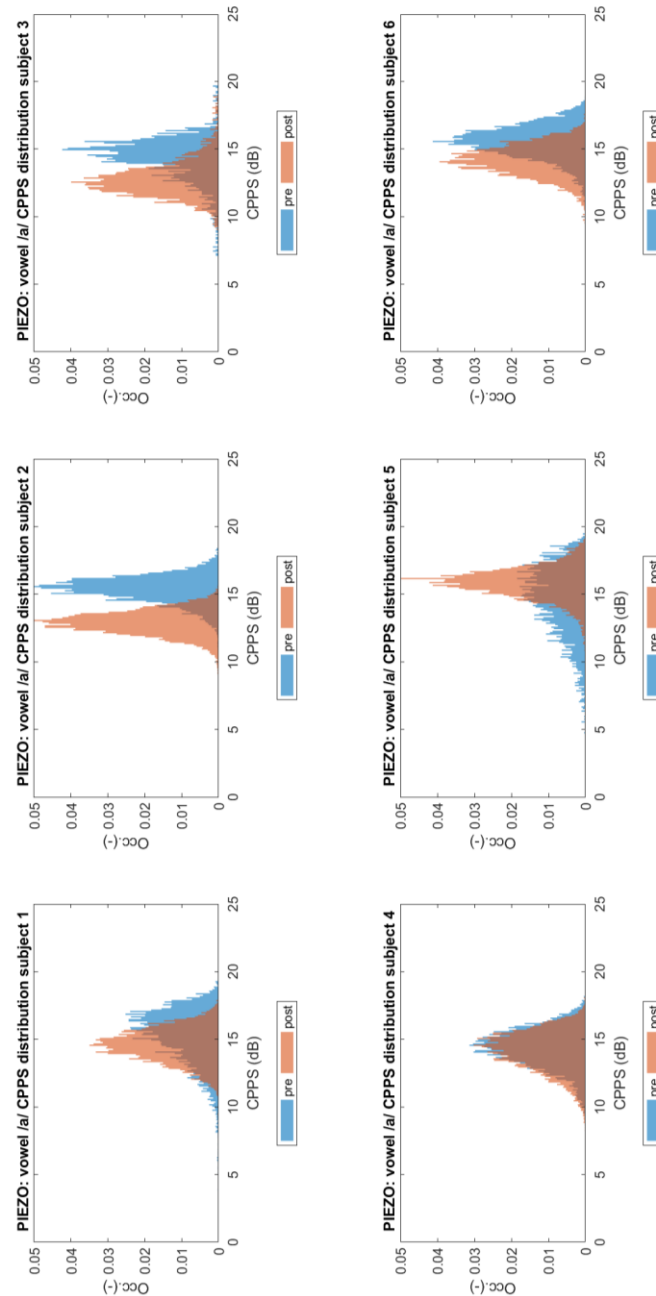


Figura 44 Distribuzioni relative al CPPS ricavate dall'analisi del segnale vocale delle vocali /a/ (microfono a contatto)

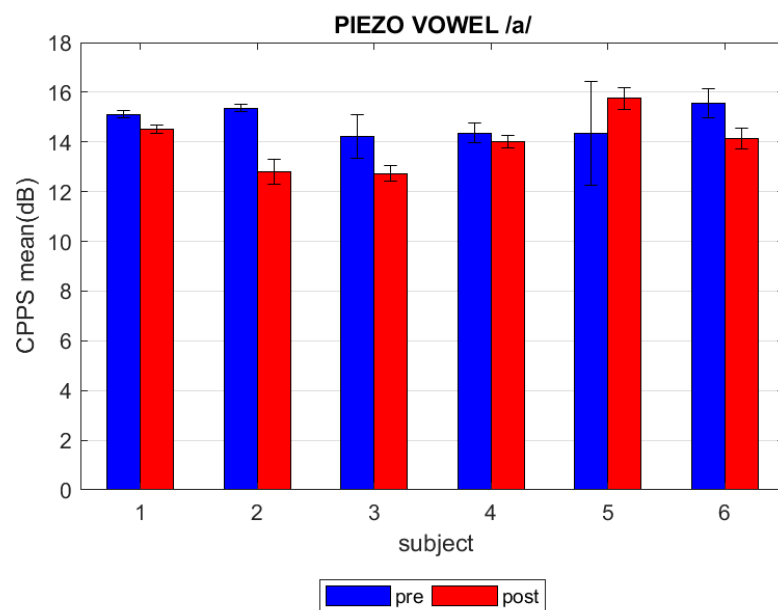


Figura 45 Variazione CPPS_{mean}

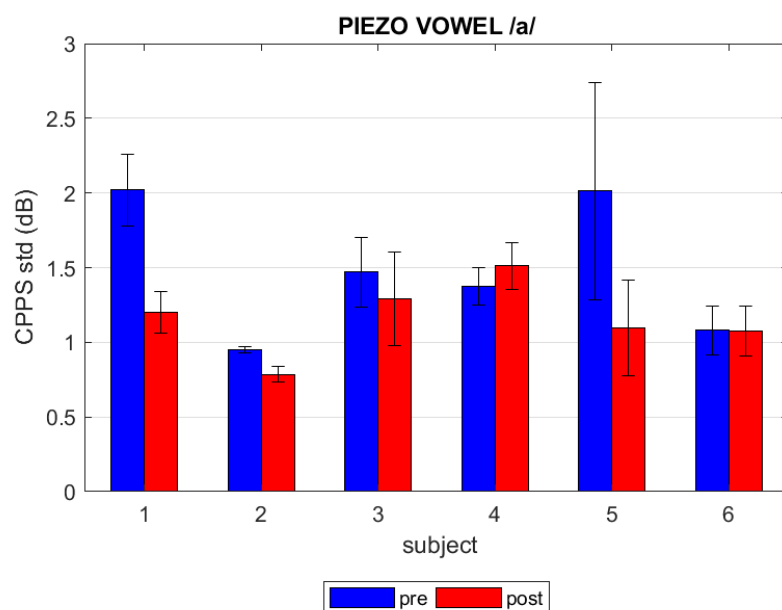


Figura 46 Variazione CPPS_{std.}

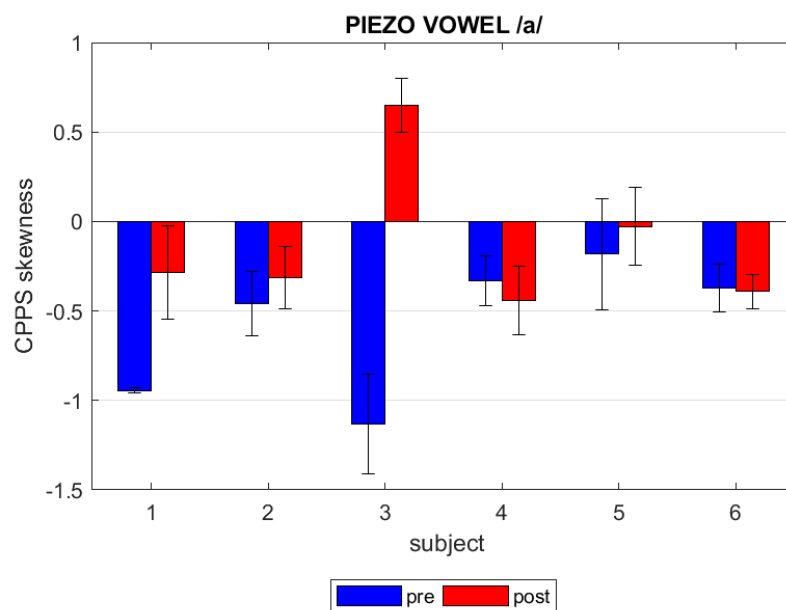


Figura 47 Variazione $CPPS_{skewness}$

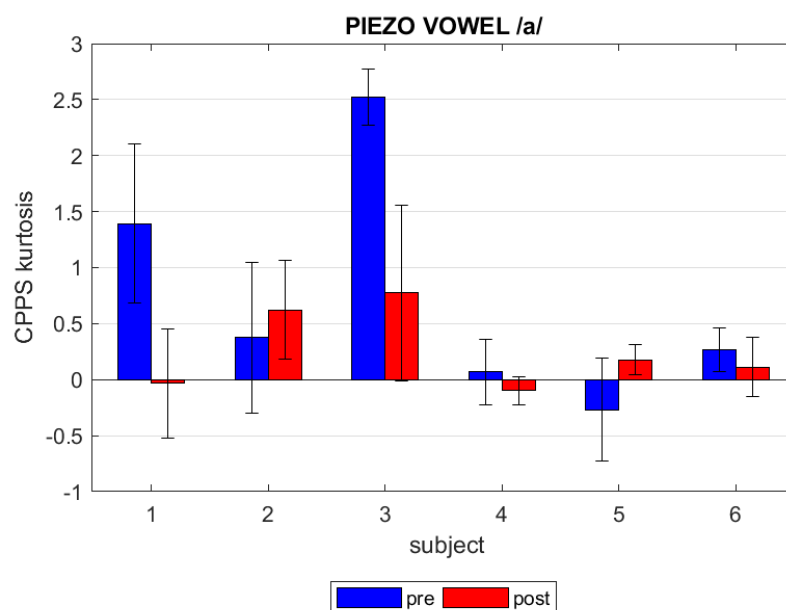


Figura 48 Variazione $CPPS_{kurtosis}$

5.2 Analisi della lettura del brano

Poiché per la lettura del brano non ha senso calcolare i parametri relativi alle perturbazioni, l'analisi effettuata si è basata sui soli parametri relativi al dominio *cepstrale* oltre che sull'HNR e sul rapporto *voiced/unvoiced*. Al fine di poter effettuare un confronto grafico vengono riportate le distribuzioni “pre” e “post” trattamento logopedico per ogni soggetto e per entrambi i microfoni utilizzati (figure 49 e 50).

5.2.1 Microfono a contatto

Con lo scopo di minimizzare il numero di parametri che evidenziano la differenza pre e post trattamento logopedico, è stata calcolata la matrice di correlazione che ha permesso di escludere i parametri ridondanti (tabella 25). Da tale matrice si evince come l'HNR ed il rapporto *voiced/unvoiced* siano tra loro correlati per tal motivo si è scelto di considerare solo uno tra i due in particolare il rapporto *voiced/unvoiced*. Inoltre, dei parametri cepstrali, si è riportata la variazione pre e post trattamento del CPPS_{mean}, CPPS_{std}, CPPS_{range}, e CPPS_{kurtosis} (per facilità riportate in seguito senza il prefisso CPPS). Dalla tabella 26 si evince come il soggetto numero uno ha mostrato un aumento del rapporto *voiced/unvoiced* ed una riduzione dei rimanenti parametri. Per il soggetto numero due non sono state rilevate delle differenze significative pre e post trattamento. Il soggetto numero tre ha mostrato una riduzione di tutti i parametri ad eccezione del rapporto *voiced/unvoiced* che è invece aumentato. Il soggetto numero quattro, all'interno del gruppo, ha riportato i valori più bassi per tutti i parametri *cepstrali*, che sono in questo caso aumentati post trattamento. Il soggetto numero cinque è caratterizzato dal rapporto *voiced/unvoiced* più basso, sebbene si sia evinto un aumento post trattamento. Aumento che ha coinvolto anche la media, la deviazione tipo ed il *range*; a differenza della *kurtosis* che si è avvicinata maggiormente a zero. L'ultimo soggetto ha riportato un andamento circa costante per tutti i parametri ad eccezione di un aumento trascurabile per il rapporto *voiced/unvoiced*.

	F0	HNR	V/Uv	CPPS								
				mean	median	mode	std	range	5prc	95prc	skewness	kurtosis
F0	1											
HNR		1	0,82									
V/Uv			1				-0,56		0,50			
mean				1	0,99	0,84			0,85	0,85	-0,90	
median					1	0,89			0,80	0,87	-0,92	
mode						1		0,52	0,61	0,81	-0,83	
std							1	0,73		0,55		-0,55
range								1		0,76		
5prc									1	0,56	-0,78	0,51
95prc										1	-0,67	
skewness											1	-0,56
kurtosis												1

Tabella 25 Matrice di correlazione relativa ai parametri calcolati lettura del brano acquisita con il microfono a contatto

			Subject					
			1	2	3	4	5	6
CPPS	V/Uv (%)	PRE	90,5	90,8	91,3	95,8	60,1	97,5
		POST	92,2	90,2	98,9	99,0	77,5	98,3
	Mean (dB)	PRE	9,5	12,2	11,3	7,7	9,7	11,9
		POST	8,0	12,1	10,3	8,2	11,2	11,2
	Std (dB)	PRE	3,4	3,3	2,8	2,5	3,4	2,6
		POST	2,9	3,1	2,5	2,6	3,2	2,6
	Range (dB)	PRE	17,0	17,4	15,8	12,6	17,3	16,1
		POST	14,9	15,6	15,2	14,5	18,1	16,0
	Kurtosis	PRE	-1,0	0,0	0,2	-0,7	-0,7	0,1
		POST	-0,8	0,3	0,4	-0,6	-0,4	-0,3

Tabella 26 Parametri pre e post trattamento logopedico ricavati dalla lettura del brano registrata con il microfono a contatto

A conferma di quanto si è descritto grazie alla tabella riportante i valori pre e post trattamento logopedico, in figura 49 vengono riportate le corrispettive distribuzioni.

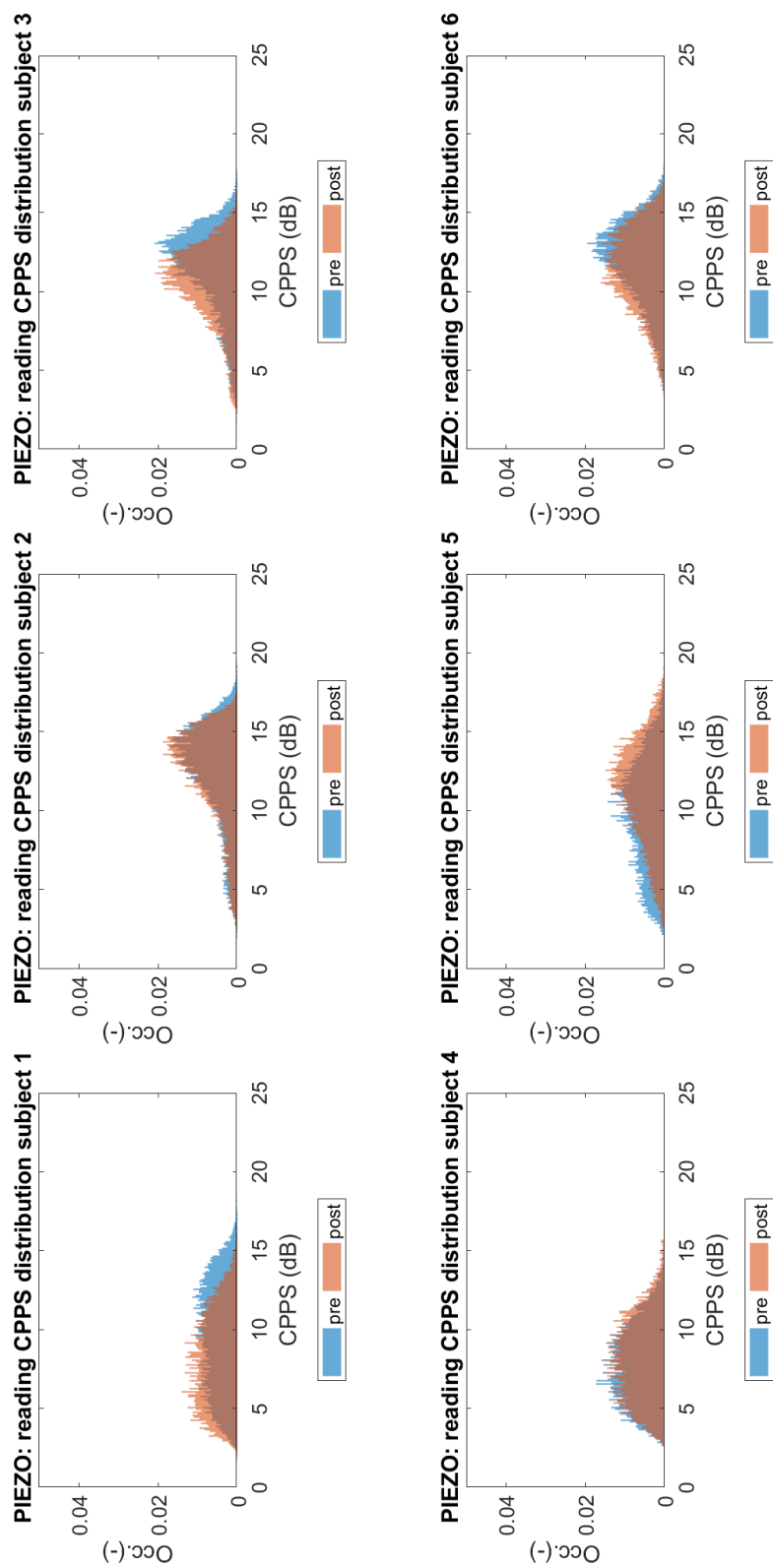


Figura 49 Distribuzioni relative al CPPS ricavate dall'analisi del segnale vocale derivante dalla lettura del brano (microfono a contatto)

5.2.2 Microfono in aria

La matrice di correlazione, relativa ai parametri ricavati dalla lettura del brano registrata con il microfono in aria, ha evidenziato le stesse correlazioni identificate per il caso precedente. Ad eccezione del CPPS_{5prc} e CPPS_{95prc} che sono entrambi correlati con la *skewness* (tabella 27).

				CPPS								
	F0	HNR	V/Uv	mean	median	mode	std	range	5prc	95prc	skewness	kurtosis
F0	1											
HNR		1	0,90									
V/Uv			1									
mean				1	0,99	0,90			0,92	0,88	-0,95	0,57
median					1	0,91		0,47	0,88	0,90	-0,97	0,54
mode						1			0,78	0,89	-0,93	
std							1	0,57				-0,57
range								1		0,52		
5prc									1	0,72	-0,83	0,53
95prc										1	-0,85	
skewness											1	-0,51
kurtosis												1

Tabella 27 Matrice di correlazione relativa ai parametri calcolati lettura del brano acquisita con il microfono in aria

I parametri ricavati dall'acquisizione del segnale vocale relativo alla lettura del brano acquisito con il microfono in aria hanno evidenziato in diversi soggetti lo stesso andamento riscontrato con il microfono a contatto. Infatti, i soggetti numero uno e due, sebbene con valori differenti, hanno mostrato la stessa variazione evidenziata nell'analisi precedente. Il soggetto numero tre, anche in questo caso, ha riportato un aumento del rapporto *voiced/unvoiced*, mentre gli altri parametri hanno mantenuto pressoché il loro valore. Il soggetto numero quattro, così come è stato riscontrato con l'analisi effettuata tramite microfono a contatto, ha mostrato i valori più bassi per tutti i parametri *cepstrali*, che sono anche in questa analisi aumentati post trattamento. Anche per il soggetto numero cinque, in analogia con quanto riscontrato in precedenza, si è evinto un aumento del rapporto *voiced/unvoiced* post trattamento; aumento che ha coinvolto in questo caso solo la media, a differenza degli altri parametri che sono diminuiti. L'ultimo soggetto, così come visto per il

microfono a contatto, ha riportato un andamento circa costante di tutti i parametri. Per rendere maggiormente visibili le differenze di tali parametri pre e post trattamento sono state riportate in figura 50 le distribuzioni relative al segnale vocale di ogni soggetto in esame.

			Subject					
			1	2	3	4	5	6
CPPS	V/Uv (%)	PRE	91,3	93,9	77,1	97,4	65,8	98,5
		POST	89,0	92,8	98,1	99,0	74,5	90,5
	Mean (dB)	PRE	9,1	11,6	11,9	7,3	8,7	11,7
		POST	7,6	12,7	11,1	8,7	10,9	10,8
	Std (dB)	PRE	3,7	3,5	2,6	2,5	3,3	2,9
		POST	2,8	3,3	2,7	3,0	2,9	2,7
	Range (dB)	PRE	17,6	17,2	16,1	12,9	17,1	15,7
		POST	14,3	17,6	17,6	14,9	17,0	15,8
	Kurtosis	PRE	-1,0	0,0	0,2	-0,8	-0,8	-0,2
		POST	-0,5	0,1	0,0	-0,8	-0,4	-0,4

Tabella 28 Parametri pre e post trattamento logopedico ricavati dalla lettura del brano registrata con il microfono in aria

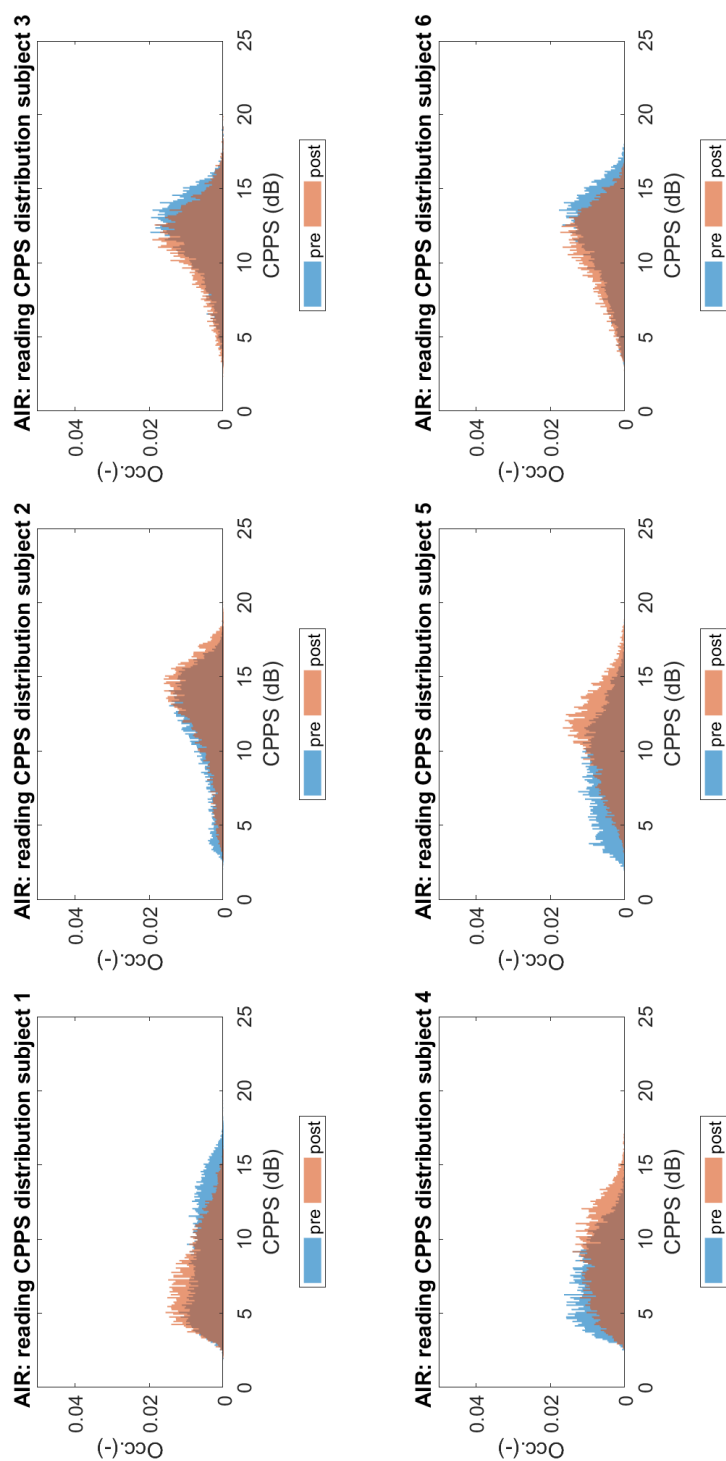


Figura 50 Distribuzioni relative al CPPS ricavate dall'analisi del segnale vocale derivante dalla lettura del brano (microfono in aria)

CONCLUSIONI

La prima parte di questo lavoro di tesi è consistito nel creare dei modelli predittivi in grado di diagnosticare un soggetto come sano o come Parkinsoniano. Durante tale fase dell'attività è stato acquisito il segnale vocale di 22 soggetti Parkinsoniani presso l'ambulatorio Parkinson dell'ospedale Molinette di Torino. Tali soggetti hanno seguito un protocollo della durata di circa cinque minuti che ha previsto la ripetizione di tre vocali sostenute /a/, la ripetizione dei fonemi /pa/ta/ka, la lettura di un brano foneticamente bilanciato ed un eloquio spontaneo. Durante la fase di acquisizione, ad ogni soggetto è stato fatto indossare un microfono a contatto a giro collo ed un microfono in aria a guancia. Solo i segnali relativi alle vocali /a/ acquisite con il microfono a contatto ed alla lettura del brano acquisiti mediante i due diversi microfoni sono stati elaborati. Lo script Matlab appositamente implementato, ha permesso di estrarre parametri nel dominio del tempo, quali *jitter* e *shimmer*, nel dominio della frequenza (*Harmonic-to-Noise-Ratio*: HNR) e nel dominio della *quefreny* (*Cepstral-Peak-Prominence-Smoothed*: CPPS). Ottenuti tutti i parametri, si sono selezionati solo quelli maggiormente significativi nell'identificazione di classificatori di voci sane e Parkinsoniane. Utilizzando la matrice di correlazione sono stati inizialmente esclusi i parametri ridondanti individuandone così un primo sottoinsieme, dal quale sono stati estratti i parametri rilevanti attraverso l'applicativo Matlab "*classification learner*". Per ciascun tipo di materiale vocale analizzato sono stati costruiti due modelli di classificatore basati sugli algoritmi *coarse tree* e regressione logistica. Per la vocale sostenuta, acquisita tramite microfono a contatto, utilizzando l'algoritmo *coarse tree* si è ottenuta una *accuracy* del 90% una *sensitivity* del 92% ed una *specificity* dell'87% con quattro parametri rilevanti: PPQ, vF_0 , *Shim* e F_0 . L'algoritmo di regressione logistica, sempre per la stesa fase del protocollo analizzata, ha mostrato invece una *accuracy* dell'83% una *sensitivity* dell'84%, una *specificity* dell'83% ed una AUC pari a 0.87. I parametri rilevanti, identificati per quest'ultimo modello, sono risultati essere il PPQ ed il vAm . Per la lettura del brano attraverso il microfono a contatto, l'algoritmo *coarse tree* ha permesso di avere una *accuracy* del 95% una *sensitivity* dell'89% ed una *specificity* del 100% con $CPPS_{range}$, $CPPS_{95prc}$ e $CPPS_{kurtosis}$ come parametri rilevanti. La regressione logistica, sempre per la stessa analisi, ha

prodotto una *accuracy* del 76% una *sensitivity* del 79%, una *specificity* del 74% ed un AUC pari a 0,79, identificando come parametri rilevanti il rapporto *voiced/unvoiced* ed il $CPPS_{range}$. Per la lettura del brano, registrato attraverso il microfono in aria, l'algoritmo *coarse tree* ha fornito una *accuracy*, una *sensitivity* ed una *specificity* rispettivamente pari all'89 95 ed 82 %. I parametri rilevanti che sono stati utilizzati sono tre: HNR, $CPPS_{std}$ e $CPPS_{5pre}$. La regressione logistica in quest'ultimo caso, ha mostrato una *accuracy* del 73% una *sensitivity* del 68% una *specificity* del 77% ed una AUC pari a 0,75. Come ci si aspettava, avendo acquisito i segnali in un ambiente affetto da rumore, il microfono a contatto ha mostrato prestazioni superiori in termini di accuratezza rispetto al microfono in aria poiché poco influenzato dal rumore di fondo. Nella seconda parte di questo lavoro di tesi si è valutata l'efficacia di un percorso logopedico acquisendo, con lo stesso protocollo e con la stessa catena di acquisizione utilizzata nella prima parte, un totale di sei soggetti presso il centro "amici Parkinsoniani Piemonte onlus" all'inizio e dopo due mesi di trattamento. Estratti tutti i parametri, ci si è serviti anche in questo caso della matrice di correlazione con lo scopo di evidenziare ed escludere i parametri ridondanti. Dall'analisi del segnale relativo alle vocali sostenute si sono ricavati dei *bar diagrams* per i parametri di perturbazione quali *Jitt*, PPQ, *vF0*, *Shim* e *vAm* e per gli indici statistici relativi al dominio cepstrale quali $CPPS_{mean}$, $CPPS_{std}$, $CPPS_{skewness}$ e $CPPS_{kurtosis}$. Per i parametri di perturbazione, nella maggior parte dei casi, si è avuto un miglioramento pre e post "trattamento", specialmente per il soggetto numero uno che ha mostrato una diminuzione significativa in tutti i parametri evidenziati. Mentre, per i parametri cepstrali, si è notata una diminuzione in termini assoluti della $CPPS_{kurtosis}$ e della $CPPS_{std}$. Per i risultati relativi all'elaborazione del segnale vocale derivante dalla lettura del brano, sia per il microfono in aria che il microfono a contatto, la matrice di correlazione ha permesso di identificare gli stessi parametri: rapporto *voiced/unvoiced*, $CPPS_{mean}$, $CPPS_{std}$, $CPPS_{range}$ e $CPPS_{kurtosis}$. I parametri appena esposti, oltre che ad essere gli stessi per entrambi i microfoni, hanno presentato un andamento simile pre e post "trattamento" sia con l'utilizzo del microfono a contatto che con l'utilizzo del microfono in aria. I risultati relativi a quest'ultima parte del lavoro di tesi, oltre che riguardare un numero limitato di soggetti, sono evidentemente soggettivi, in quanto, il fine ultimo del corso logopedico è stato quello di insegnare ai soggetti interessati degli esercizi da svolgere autonomamente

e con costanza in casa. Perciò, la valutazione delle variazioni relative ai parametri estratti pre e post “trattamento” hanno evidenziato dei valori differenti intersoggetto e dovranno essere ripetute successivamente essendo un percorso con effetti a medio-lungo termine, considerando magari un numero maggiore di soggetti e tenendo anche traccia della frequenza giornaliera di allenamento degli stessi. In conclusione, questo lavoro di tesi ha permesso di identificare su quali parametri concentrarsi maggiormente al fine di discriminare un soggetto come sano o come Parkinsoniano in base al segnale vocale, ed ha indirizzato l'utilizzo del microfono a contatto come sistema hardware con il quale effettuare le registrazioni, sia al fine di creare un modello predittivo che al fine di valutare l'efficacia di percorsi logopedici. Tuttavia, lo stesso lavoro, si pone ancora alla base di un progetto futuro, che per consolidare i risultati ottenuti richiederà l'estensione del *data set*.

APPENDICE

Brano foneticamente bilanciato

Avevo un bulldog che si chiamava Bulka. Era tutto nero salvo una macchia bianca all'estremità delle zampe anteriori. Nei cani di questa razza, la mandibola è sempre prominente, così i denti superiori vengono a collocarsi dietro a quelli inferiori. Ma quella di Bulka era tanto grossa che tra gli uni e gli altri denti rimaneva molto spazio. Aveva il muso largo, grandi occhi neri e brillanti e i canini sempre scoperti, perfettamente bianchi. Somigliava a un grugno. Bulka era assai forte. E se afferrava qualcosa tra i denti non c'era verso che mollasse la sua preda. Stretti i canini nella carne dell'avversario, serrava la mascella e rimaneva sospeso come un cencio ad un chiodo: attaccato come una sanguisuga. Un giorno che era stato lanciato contro un orso, gli afferrò tra i denti un orecchio. L'orso cercava di colpirlo con una zampa, scuoteva la testa, ma non se ne poteva sbarazzare: finì per rovesciare il testone in terra per schiacciarvi il cane. Su quest'ultimo, però, perché lasciasse la presa, dovemmo gettare una secchia di acqua gelata. Lo avevo avuto da ragazzo e gli davo da mangiare io stesso. Quando dovetti partire a prestar servizio nel Caucaso, decisi di non prenderlo con me e cercai di andarmene senza che lo sapesse. Ordinai che lo tenessero rinchiuso. Ero giunto alla prima tappa, stavo per ripartire con i cavalli freschi, quando ad un tratto notai una palla nera e brillante che avanzava velocissima sulla strada. Era Bulka col suo collare di rame al collo. Correva a perdifiato; si gettò su di me, mi lecco la mano e poi, con la lingua a ciondolini, si stese all'ombra sotto la vettura. Seppi più tardi che aveva rotto un vetro per seguirmi; era saltato dalla finestra: aveva percorso venti chilometri d'estate, sotto un sole bruciante.

BIBLIOGRAFIA

Testo

- [1] W. De Colle. VoceComputer, Analisi acustica digitale del segnale verbale (Il sistema CSL-MDVP). Omega Edizioni, 2001.
- [2] Fahn S, Elton RL, UPDRS Program Members. Unified Parkinson's disease rating scale. In: Fahn S, Marsden CD, Goldstein M, Calne DB, editors. Recent developments in Parkinson's disease, Vol. 2. Florham Park, NJ: Macmillan Healthcare Information; 1987. p 153–163, 293–304
- [3] Ramaker C, Marinus J, Stiggelbout AM, van Hilten BJ. Systematic evaluation of rating scales for impairment and disability in Parkinson's disease. *Mov Disord* 2002;17:867–876).
- [4] <https://www.my-personaltrainer.it/salute/diagnosi-morbo-di-parkinson.html>
- [5] Goetz CG, Fahn S, Martinez-Martin P, et al. Movement disorder society-sponsored revision of the Unified Parkinson's disease rating scale (MDS-UPDRS): process, format, and clinimetric testing plan. *Mov Disord* 2007;22:41–47
- [6] G. Deuschl, C. Schade-Brittinger, P. Krack, A Randomized Trial of Deep-Brain Stimulation for Parkinson's Disease, *N. Engl J Med.*, 21 September 2006
- [7] Ho AK, Iannsek R, Marigliani C, Bradshaw JL, Gates S (1999) Speech impairment in a large sample of patients with Parkinson's disease. *J Behav Neurol* 11:131–137 [CrossRefGoogle Scholar](#)
- [8] Darley FL, Aronson AE, Brown JR (1969a) Differential diagnostic patterns of dysarthria. *J Speech Hear Res* 12:246–269 [PubMedCrossRefGoogle Scholar](#)
- [9] Kegl J, Cohen H, Poizner H (1999) Articulatory consequences of Parkinson's Disease: perspectives from two modalities. *Brain Cogn* 40:355–386 [PubMedCrossRefGoogle Scholar](#).
- [10] Hornykiewicz O (1998) Biochemical aspects of Parkinson's disease. *Neurology* 51:S2–S9 [PubMedCrossRefGoogle Scholar](#)
- [11] Baker KK, Ramig LO, Luschei ES, Smith ME (1998) Thyroarytenoid muscle activity associated with hypophonia in Parkinson's disease and aging. *Neurology* 51:1592–1598 [PubMedCrossRefGoogle Scholar](#)

- [12] (Spencer KA, Rogers MA (2005) Speech motor programming in hypokinetic and ataxic dysarthria. *Brain Lang* 94:347–366PubMedCrossRefGoogle Scholar)
- [13] Roy N, Nissen SL, Dromey C, Sapir S (2009) Articulatory changes in muscle tension dysphonia: evidence of vowel space expansion following manual circumlaryngeal therapy. *J Commun Disord* 42:124–135PubMedCrossRefGoogle Scholar)
- [14] (Moretti R, Torre P, Antonello RM, Capus L, Gioulis M, Zambito Marsala S, Cazzato G, Bava A (2003) Speech initiation hesitation following subthalamic nucleus stimulation in a patient with parkinson's disease. *Eur Neurol* 49:251–253PubMedCrossRefGoogle Scholar
- [15] Gentil M, Pollak P, Perret J (1995) Parkinsonian dysarthria. *Rev Neurol* 151:105–112PubMedGoogle Scholar
- [16] Berg D, Postuma RB, Adler CH et al (2015) MDS research criteria for prodromal Parkinson's disease. *Mov Disord* 30:1600–1611. doi: 10.1002/mds.26431PubMedCrossRefGoogle Scholar
- [17] G.J. CanterSpeech characteristics of patients with Parkinson's disease: II. Physiological support for speech *J Speech Hear Disord*, 30 (1965), pp. 44-49
- [18] F.J. Jiménez-Jiménez, J. Gamboa, A. Nieto, *et al.*Acoustic voice analysis in untreated patients with Parkinson's disease *Parkinsonism Relat Disord*, 3 (1997), pp. 111-116
- [19] Youri Maryn, Paul Corthals, Paul Van Cauwenberge, Nelson Roy, and Marc De Bodt. Toward improved ecological validity in the acoustic measurement of overall voice quality: combining continuous speech and sustained vowels.*Journal of voice*, 24(5):540-555, 2010.
- [20] H. Oguz, M. A. Kilic, M. Safak. Comparison of results in two acoustic analysis programs: Praat and MDVP. *Turk J Med Sci*, 2010.
- [21] Meike Brockmann, Michael J Drinnan, Claudio Storck, and Paul N Card-ing. Reliable jitter and shimmer measurements in voice clinics: the relevance of vowel, gender, vocal intensity, and fundamental frequency effects in a typical clinical task.*Journal of voice* , 25(1):4453, 2011.

- [22] Joao Paulo Teixeira, Carla Oliveira, and Carla Lopes. Vocal acoustic analysis jitter, shimmer and hnr parameters. *Procedia Technology*, 9:11121122, 2013.
- [23] E. Yumoto, W. J. Gould e T. Baer, «Harmonics-to-noise ratio as an index of the degree of hoarseness,» *The journal of the Acoustical Society of America*, vol. 6, n. 71, pp. 1544-1550, 1982.
- [24] Y. Qui e R. E. Hillman, «Temporal and spectral estimations od harmonics-to noise ratio in human voice signals,» *The Journal of the Acoustical Society of America*, vol. 1, n. 102, pp. 537-543, 1997
- [25] Y. D. Heman-Ackah, D. D. Michael e G. S. Goding, «The relationship between cepstral peak prominence and selected parameters of dysphonia,» *Journal of voice*, vol.1, n. 16, pp. 20-27, 2002
- [26] Antonella Castellana, Alessio Carullo, Simone Corbellini, and Arianna Astol. Discriminating Pathological Voice From Healthy Voice Using Cep-stral Peak Prominence Smoothed Distribution in Sustained Vowel. *IEEE Transaction on Instrumentation and Masurement*. IEEE, 2018.
- [27] A. Ferreira, V. Fernandes, Consistency of the F0, Jitter, Shimmer and HNR voice parameters in GSM and VOIP communication, 2017 22nd International Conference on Digital Signal Processing, Proceedings, August 2017
- [28] F. E. Ferrero, R. Lanni e W. De Colle, Primi risultati di uno studio per la validazione del sistema MDVP come strumento per una caratterizzazione multiparametrica della voce, *Acta Phon. Lat.*, 1995
- [29] J. W. Tukey, B. P. Borget e M. Healy, «The frequency analysis of time series for echoes: Cepstrum, pseudo-autocovariance, cross-cepstrum and saphe cracking, time series analysis,» 1963.
- [30] Lior Rokach, Oded Maimon, *Data mining with decision trees theory and application*, World Scientific, 2nd Edition
- [31] Breiman L., Friedman J., Olshen R., and Stone C. *Classification and Regression Trees*. Wadsworth Int. Group, 1984
- [32] G. D'Arrigo, F. Provenzano, C. Torino, C. Zoccali e G. Tripepi, «I test diagnostici e l'analisi della curva ROC,» *G Ital Nefrol*, vol. 28, n. 6, pp. 642-647, 20.

Figure

[1] [http://www2.pd.istc.cnr.it/piero/Biblos/images/FoneticaAcustica \(Fig102\).htm](http://www2.pd.istc.cnr.it/piero/Biblos/images/FoneticaAcustica (Fig102).htm)

[2] <http://sextuple.ga/paveh/dopamina-e-parkinson-poki.php>

[3] <https://www.lastampa.it/2017/03/20/scienza/parkinson-oltre-al-tremore-disturbi-cognitivi-e-organici-che-cambiano-la-vita-gorZJrH3pv6cYRs8rMGVtO/pagina.html>

[4] <https://www.cienciasinseso.com/en/tag/roc-curve/>

RINGRAZIAMENTI

Un immenso ringraziamento va al Professor Alessio Carullo per avermi supportato nel lavoro di tesi con professionalità e per avermi trasmesso serenità ed umiltà, due rare doti che, oltre alla gentilezza, lo contraddistinguono. Un sentito ringraziamento va al Professor Alberto Vallan ed al Dott. Alessio Atzori per avermi aiutato a portare a termine questo lavoro. Grazie a Francesco, Viviana, Andon e Vittorio per avermi sopportato durante il percorso di tesi e per aver alleggerito le giornate in laboratorio. Grazie al Dott. Gaetano Marino per avermi dato l'opportunità di vivere in prima persona la medicina chirurgica urologica nella quale è un luminare. Un grandissimo grazie è rivolto a tutte le persone con le quali ho avuto a che fare in questi due anni; in particolare Domenico, con cui ho avuto la fortuna ed il piacere di sostenere gli esami di gruppo. Grazie a tutti i miei coinquilini che, durante le nostre brevi convivenze, mi ha fatto sentire a casa, ognuno a suo modo. Un pensiero immenso va a tutti i miei amici "di giù" per avermi sempre accolto calorosamente ogni qualvolta sono ritornato a casa. Grazie a tutti i miei zii, cugini e splendidi nonni. Non basterebbero infiniti ringraziamenti a mio Padre, mia Madre e mio Fratello. È per questo che mi limito a dirvi che ogni mio gesto d'amore, ogni mio sorriso ed ogni mia lacrima di gioia è solo merito vostro. Siete e sarete per sempre la mia forza ed il mio orgoglio oltre che la mia patria del cuore.