

POLITECNICO DI TORINO

Corso di Laurea Magistrale in Ingegneria Biomedica



Tesi di Laurea Magistrale

Valutazione automatica delle alterazioni vocali nella malattia di Parkinson: sviluppo di un protocollo per la raccolta dati ed analisi acustica

Relatore

Prof.ssa Gabriella OLMO

Correlatore

Dott.ssa Federica AMATO

Candidata

Sabrina Scimeca

Marzo 2023

Sommario

La disfunzione vocale è uno dei primi sintomi a manifestarsi durante l'evoluzione della malattia di Parkinson (MP). Diversi studi hanno dimostrato l'efficacia della parola come un potente biomarcatore delle anomalie del linguaggio associate alla neurodegenerazione, e l'analisi è ancora oggi oggetto di interesse della comunità scientifica.

Le caratteristiche della voce, quindi, possono essere impiegate come strumento di diagnosi e di monitoraggio dei fenomeni legati alla patologia. Poiché, però, vari fattori influenzano la produzione vocale, quali la lingua, il genere, l'età, il metodo di registrazione, è necessario definire un protocollo procedurale per fornire un metodo generale, standard, robusto e facilmente riproducibile di acquisizione, classificazione ed interpretazione dei dati.

La presente tesi mira a fornire una metodologia per la valutazione delle registrazioni vocali dei pazienti con MP e ad investigare l'influenza dei vari fattori esterni, quali le caratteristiche del parlante e delle condizioni di registrazione, sulle feature vocali. Infatti, se opportunamente validati, i risultati qui proposti consentirebbero di identificare un set di compiti vocali e feature acustiche capace di catturare le alterazioni della voce, minimamente dipendente dal genere, dalla lingua e dalle tecniche di registrazione. Questo consentirebbe la validazione cross-corpora degli algoritmi di analisi vocale e lo sviluppo di strumenti utili ai medici durante la diagnosi e il monitoraggio della malattia di Parkinson.

Il database utilizzato in questa ricerca è composto da: (a) 4 dataset in lingua italiana, con registrazioni relative a 104 soggetti MP e 106 soggetti controlli sani (CS) (2 dataset registrati utilizzando un microfono professionale e 2 dataset registrati tramite smartphone); (b) 1 dataset in lingua spagnola con registrazioni relative a 50 soggetti MP e 50 soggetti CS (registrati tramite microfono professionale); (c) 1 dataset in lingua ceca con registrazioni relative a 22 soggetti MP e 22 soggetti CS (registrati tramite microfono professionale). Sono state analizzate le voci di soggetti registrate durante le letture di frasi e le fonazioni prolungate delle vocali /a/ o /e/.

I risultati complessivi di questa tesi possono essere di supporto ai medici per identificare l'insorgenza della MP, per monitorarla attraverso test vocali semplici e non invasivi e per migliorare il processo decisionale clinico nei soggetti con MP.

*Ai miei genitori,
a mia sorella Fabiola,
per essermi sempre vicino.*

Indice

Elenco delle tabelle	IX
Elenco delle figure	XI
1 Introduzione	1
1.1 Produzione della voce	2
1.2 Disturbi del linguaggio dovuti al Parkinson	3
1.3 Scopo e struttura della tesi	4
2 Stato dell'arte	6
2.1 Limitazioni	8
3 Materiali e metodi	10
3.1 Database	10
3.2 Compiti vocali	12
3.2.1 Fonazione sostenuta	12
3.2.2 Frasi	12
3.3 Feature vocali	13
3.3.1 Caratteristiche basate su F0	13
3.3.2 Caratteristiche di frequenza formante (FORM)	14
3.3.3 Varianti del jitter	14
3.3.4 Varianti di shimmer	15
3.3.5 Coefficienti di frequenza Mel Cepstral (MFCC)	16
3.3.6 Rapporto tra armoniche e rumore (HNR)	16
3.3.7 Misure cepstrali e spettrali della qualità della voce (CPP) . .	16
3.3.8 Rapporto di eccitazione glottale-rumore (GNE)	16
3.3.9 Coefficienti predittivi lineari percettivi (PLP)	17
3.3.10 Analisi delle fluttuazioni senza tendenza (DFA)	17
3.4 Algoritmo di classificazione	17
3.4.1 Normalizzazione	18
3.4.2 Feature selection	19

3.4.3	Classificatori	20
3.4.4	K-fold cross-validation	22
3.4.5	Metriche di valutazione	22
3.5	Test statistici	23
3.5.1	Shapiro-Wilk test	24
3.5.2	U-Mann Whitney test	24
3.5.3	Kruskal Wallis test	25
4	Machine Learning per la classificazione dei segnali vocali	26
4.1	Valutazione dell'influenza della tecnica di registrazione	26
4.2	Identificazione compiti vocali da utilizzare nel protocollo	27
4.2.1	Scelta della vocale per il compito della fonazione sostenuta	28
4.2.2	Scelta delle frasi	29
4.2.3	Protocollo	29
5	Analisi Multilingua	31
5.1	Machine learning	31
5.2	Analisi statistica	32
5.2.1	Analisi sui 5 dataset uniti	33
5.2.2	Analisi su ogni dataset	33
5.3	Feature robuste e Machine learning	34
6	Risultati	36
6.1	Valutazione del metodo di registrazione	36
6.2	Scelta della vocale	36
6.3	Scelta delle frasi	37
6.4	Protocollo	41
6.4.1	Italiano 1 + Privato 1	42
6.4.2	Colombiano	42
6.4.3	Italiano 2	43
6.5	Machine learning per l'analisi multilingua	66
6.6	Analisi statistica	66
6.6.1	Analisi su ogni dataset	66
6.6.2	Analisi sui 5 dataset uniti	79
6.6.3	Set di feature	93
6.6.4	Feature robuste e Machine learning	94
7	Discussioni	96
7.1	Registrazione	96
7.2	Scelta vocale	97
7.3	Scelta frasi	97
7.4	Protocollo	98

7.5	Machine learning per l'analisi multilingua	99
7.6	Feature robuste e Machine learning	99
7.7	Limitazioni	101
8	Conclusioni	102
	Bibliografia	104

Elenco delle tabelle

3.1	Dati sui soggetti presenti in ogni dataset. <i>MR: microfono; SP: smartphone; V_a: vocalizzo /a/; V_e: vocalizzo /e/; F: frasi</i>	11
6.1	Accuratezza (%) media sul validation (k-fold al 70% del trainset) con 10 Classificatori e vocalizzo /a/, per modalità di registrazione .	40
6.2	Accuratezza (%) media sul validation (k-fold al 70% del trainset) con 10 Classificatori e frasi Italiano 1, per modalità di registrazione	40
6.3	Accuratezza (%) media sul validation (k-fold al 70% del trainset) con 10 Classificatori e frasi Italiano 1+Privato 1, per modalità di registrazione	40
6.4	Accuratezza (%) media sul validation (k-fold al 70% del trainset) con 10 Classificatori e vocalizzo /a/ e vocalizzo /e/	41
6.5	Accuratezza (%) media sul validation (k-fold al 70% del trainset), con 10 Classificatori, 10 frasi Italiano 1+Privato 1	41
6.6	Accuratezza (%) media sul validation (k-fold al 70% del trainset) su 10 Classificatori, 6 frasi Colombiano	41
6.7	Accuratezza (%) media sul validation (k-fold al 70% del trainset) su 10 Classificatori, 3 frasi Italiano 2	41
6.8	Risultati delle metriche di valutazione ottenuti su 10 Classificatori, vocalizzo /a/ su Italiano1+Privato1	54
6.9	Risultati delle metriche di valutazione ottenuti su 10 Classificatori, frase migliore su Italiano1+Privato1	55
6.10	Risultati delle metriche di valutazione ottenuti su 10 Classificatori, vocalizzo /a/ unito alla frase migliore su Italiano1+Privato1	56
6.11	Risultati delle metriche di valutazione ottenuti su 10 Classificatori, vocalizzo /a/ unito alle 2 frasi migliori su Italiano1+Privato1	57
6.12	Risultati delle metriche di valutazione ottenuti su 10 Classificatori, vocalizzo /a/ su Colombiano	58
6.13	Risultati delle metriche di valutazione ottenuti su 10 Classificatori, frase migliore su Colombiano	59

6.14	Risultati delle metriche di valutazione ottenuti su 10 Classificatori, vocalizzo /a/ unito alla frase migliore su Colombiano	60
6.15	Risultati delle metriche di valutazione ottenuti su 10 Classificatori, vocalizzo /a/ unito alle 2 frasi migliori su Colombiano	61
6.16	Risultati delle metriche di valutazione ottenuti su 10 Classificatori, vocalizzo /e/ su Italiano 2	62
6.17	Risultati delle metriche di valutazione ottenuti su 10 Classificatori, frase migliore su Italiano 2	63
6.18	Risultati delle metriche di valutazione ottenuti su 10 Classificatori, vocalizzo /e/ unito alla frase migliore su Italiano 2	64
6.19	Risultati delle metriche di valutazione ottenuti su 10 Classificatori, vocalizzo /e/ unito alle 2 frasi migliori su Italiano 2	65
6.20	feature selezionate dall'algoritmo Boruta applicato al vocalizzo /a/ dataset unito multilingua	66
6.21	Risultati delle metriche di valutazione ottenuti su 10 Classificatori, vocalizzo /a/ del dataset unito multilingua	66
6.22	feature con valore di p riferito alla classe ottenuto con il test di Mann-Whitney, per i 5 dataset	67
6.22	feature con valore di p riferito alla classe ottenuto con il test di Mann-Whitney, per i 5 dataset	79
6.23	feature statisticamente discriminanti in termini di classe	79
6.23	feature che non presentano alcuna differenza statisticamente significativa in termini di classe	82
6.24	feature che non presentano differenza statisticamente significative in termini di classe	83
6.24	feature che non presentano alcuna differenza statisticamente significativa in termini di classe	92
6.25	3 gruppi di feature	93
6.26	Set completo di feature robusto individuato dall'analisi statistica . .	94
6.27	Risultati ottenuti utilizzando le feature del Gruppo 1	94
6.28	Risultati ottenuti utilizzando le feature del Gruppo 2	95
6.29	Risultati ottenuti utilizzando tutte le feature del set completo (Tabella 6.26)	95
6.30	Risultati ottenuti utilizzando tutte le feature della Tabella 6.26, applicando Boruta e aggiungendo le feature <i>registrazione</i> , <i>lingua</i> , <i>genere</i>	95

Elenco delle figure

1.1	Modello corde vocali soggetto sano [11]. <i>Da sinistra, conformazione corde vocali in fase di mormorio, sussurro, suono sonoro, suono sordo.</i>	2
1.2	Modello sistema tratto vocale semplificato [14]	3
1.3	Modello corde vocali soggetto con MP. (a) <i>Fonazione.</i> (b) <i>Inspirazione normale.</i> (c) <i>Corde vocali immobili nella posizione paramedica.</i> (d) <i>Corda vocale sinistra appare immobile nella posizione paramedica mentre la destra si muove per l'inspirazione [16]</i>	4
3.1	Grafico delle distribuzioni dei soggetti nel databse. <i>MP: Malattia di Parkinson; CS: Controlli Sani</i>	11
3.2	Diagramma di flusso algoritmo di classificazione.	18
4.1	Diagramma di flusso valutazione tecnica di registrazione	27
4.2	Diagramma di flusso scelta della vocale	28
4.3	Diagramma scelta frasi	29
4.4	Diagramma di flusso protocollo	30
5.1	Diagramma di flusso applicazione algoritmo di classificazione considerando l'unione di 5 dataset	32
5.2	Diagramma di flusso analisi statistica sul dataset unito dataset . . .	33
5.3	Diagramma di flusso analisi statistica su ogni dataset	34
5.4	Diagramma di flusso ML con sei di feature dell'analisi statistica . .	35
6.1	Lista di feature Italiano 1 e Colombiano, vocalizzo /a/	38
6.2	Lista di feature Italiano 1 e Colombiano, vocalizzo /e/	39
6.3	Lista di feature per le 2 frasi migliori Italiano 1+Privato1	44
6.4	Lista di feature per le 2 frasi migliori Colombiano	45
6.5	Lista di feature per le 2 frasi migliori Italiano 2	46
6.6	Lista di feature Italiano1+Privato1 per il vocalizzo /a/ e la frase migliore	47
6.7	Lista di feature estratte dall'algoritmo Boruta vocalizzo /a/ unito alla frase migliore su Italiano1+Privato1	48

6.8	Lista di feature estratte dall'algoritmo boruta vocalizzo /a/ unito alle 2 frasi migliori su Italiano1+Privato1	49
6.9	Lista di feature vocalizzo /a/ su Colombiano	50
6.10	Lista di feature frase migliore su Colombiano	50
6.11	Lista di feature vocalizzo /a/ unito alla frase migliore su Colombiano	51
6.12	Lista di feature vocalizzo /a/ unito alle 2 frasi migliori su Colombiano	51
6.13	Lista di feature vocalizzo /e/ su Italiano 2	52
6.14	Lista di feature frase migliore su Italiano 2	52
6.15	Lista di feature vocalizzo /e/ unito alla frase migliore su Italiano 2 .	53
6.16	Lista di feature estratte dall'algoritmo boruta vocalizzo /e/ unito alle 2 frasi migliori su Italiano 2	53

Capitolo 1

Introduzione

Il Parkinson è una malattia neurodegenerativa causata dalla progressiva e non reversibile degenerazione nella sostanza nigra pars compacta del mesencefalo [1]. La conseguente carenza di dopamina all'interno dei gangli della base porta ad un disturbo caratterizzato dai classici sintomi parkinsoniani, principalmente relativi al controllo del movimento e dell'equilibrio [2]. Il ciclo cerebello-talamo-corticale svolge un ruolo importante nella frequenza del tremore tipico parkinsoniano ed è probabilmente coinvolto nella soppressione del tremore a riposo durante i movimenti volontari [3].

Per motivi non ancora pienamente compresi, l'incidenza e la prevalenza di questa malattia sono aumentate rapidamente negli ultimi due decenni [4]. I tassi di prevalenza nelle varie popolazioni esaminate risultano variabili dai 31 per 100000 abitanti al 328 per 100000 abitanti, con prevalenza più alta in Europa, Nord America e Sud America [5]. Anche in Italia si riscontrano oscillazioni tra 66 e 257 per 100000 abitanti. In genere la MP si manifesta tra i 57/60 anni e ne sono affetti sia uomini (in prevalenza) che donne [6]. L'età è il maggior fattore di rischio per lo sviluppo della malattia di Parkinson. La prevalenza e l'incidenza aumentano quasi esponenzialmente con l'età e raggiungono il picco dopo gli 80 anni. Anche il genere è un fattore di rischio consolidato, con un rapporto maschi-femmine di circa 3:2.

I sintomi motori classici del Parkinson, sin dalla descrizione iniziale di James Parkinson nel XIX secolo [2], comprendono tremore a riposo, rigidità muscolare, bradicinesia e instabilità posturale. Per la valutazione della disabilità motoria, si utilizzano maggiormente due scale di valutazione [7]. La scala Hoehn e Yahr è comunemente usata per confrontare gruppi di pazienti e per fornire una valutazione grossolana della progressione della malattia, va dallo stadio 0 (nessun segno di malattia) allo stadio 5 (sedia a rotelle legati o costretti a letto se non assistiti) [7]; la scala Unified Parkinson's Disease Rating (UPDRS) [8] è la scala più consolidata

per valutare disabilità e disabilit , ed   formata da quattro componenti: Parte I (Mente, Comportamento e Umore); Parte II (Attivit  della vita quotidiana); Parte III (Motoria); Parte IV (Complicazioni).

I pazienti con malattia di Parkinson sviluppano anche una variet  di sintomi non motori come disturbi del sonno, depressione, disturbi cognitivi [9], e circa il 90% dei pazienti sviluppa disturbi del linguaggio, generalmente raggruppati sotto il nome di disartria ipocinetica, i quali si manifestano nella fonazione, nella formazione di suoni chiari e distinti e nella produzione prosodica [10].

1.1 Produzione della voce

La voce umana   il prodotto dei movimenti coordinati del sistema respiratorio (l'aria necessaria per produrre il suono), della laringe (dove vi sono allocate le corde vocali) e di cavit  come faringe, naso, bocca, che fungono da risuonatori, cio  da modificatori di frequenze, e influiscono sul timbro vocale.

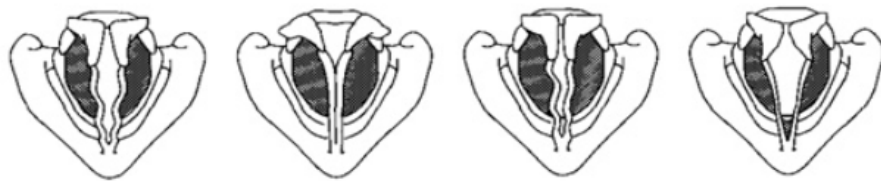


Figura 1.1: Modello corde vocali soggetto sano [11].

Da sinistra, conformazione corde vocali in fase di mormorio, sussurro, suono sonoro, suono sordo.

La laringe ospita le pieghe vocali, che vibrano al passaggio dell'aria generando la voce. Le corde vocali a riposo, cio  in fase di silenzio, presentano un'apertura triangolare, detta glottide rappresentata in Figura 1.1. Durante la fonazione le corde vocali si accostano, chiudendo pi  o meno la glottide. La produzione della voce, e quindi del linguaggio, necessita anche della partecipazione attiva di altre strutture tra le quali labbra, denti, lingua e uola, detti articolatori [12].

In un sistema di produzione vocale semplificato, come mostrato nella Figura 1.2, la produzione del suono inizia con la creazione del flusso d'aria dai polmoni. Tale flusso d'aria viene poi forzato verso l'area glottale. La fonazione   il risultato delle vibrazioni delle corde vocali e del cambiamento della pressione dell'aria. Con la modifica delle strutture nel tratto fonatorio superiore, come labbra, denti e lingua,

si determina infine la formazione di suoni chiari e distinti nel parlato [13]. La frequenza alla quale vibrano le corde vocali è chiamata frequenza fondamentale (F_0) o pitch.

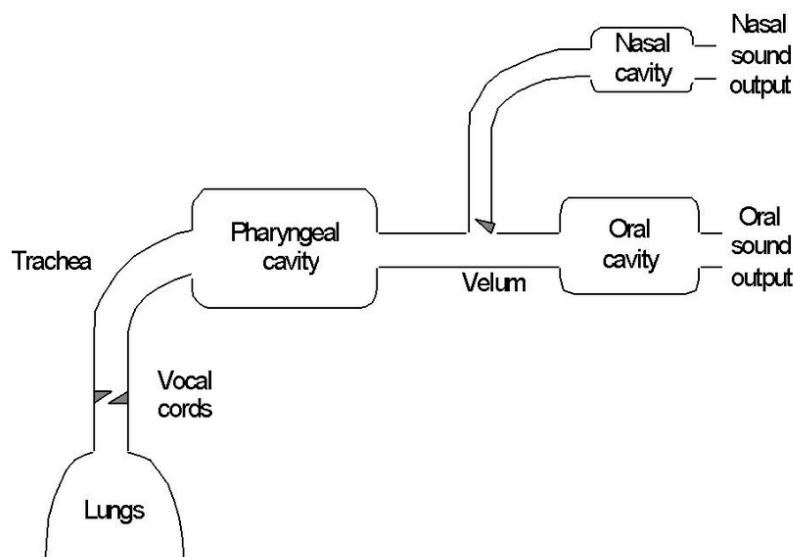


Figura 1.2: Modello sistema tratto vocale semplificato [14]

L'intensità vocale dipende da quanta energia viene ceduta dalle vibrazioni delle corde vocali all'aria. Per una voce di intensità normale, l'andamento nel tempo della velocità del volume d'aria emesso, è un segnale quasi periodico di forma approssimativamente triangolare, caratterizzato da due istanti di discontinuità, uno iniziale ed uno finale, ovvero gli istanti di apertura e chiusura glottale.

1.2 Disturbi del linguaggio dovuti al Parkinson

È ben noto come i sintomi della malattia di Parkinson siano associati alle alterazioni nei circuiti dei gangli della base, causate dalla diminuzione della dopamina nella substantia nigra [1]. Si manifestano così tremori, rigidità, lentezza nei movimenti e, non meno comuni, disturbi del linguaggio. Col termine *disartria ipocinetica* si indicano tutti i cambiamenti della voce dovuti alla malattia di Parkinson, quali riduzione del volume della voce, range tonale ridotto, ruvidità, durezza e soffio nella voce, produzioni imprecise di consonanti e vocali [15].

Numerosi studi sono stati condotti per identificare la correlazione fisiologica tra i disturbi fonatori e la patologia. Da un esame di *laringoscopia* è emerso che la glottide nei pazienti con MP non si chiude completamente a causa dell'irrigidimento delle corde vocali. L'incompleta chiusura della glottide, come mostrato in Figura 1.3, provoca perdite d'aria attraverso la glottide aperta, e durante la fonazione causa anomalie come soffio e rugosità nella voce [13].

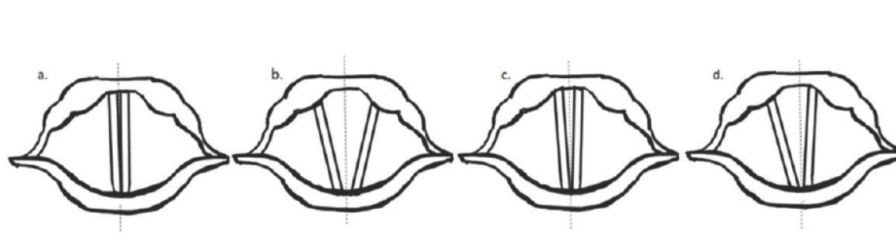


Figura 1.3: Modello corde vocali soggetto con MP.

(a) *Fonazione.* (b) *Inspirazione normale.* (c) *Corde vocali immobili nella posizione paramedica.* (d) *Corda vocale sinistra appare immobile nella posizione paramedica mentre la destra si muove per l'inspirazione [16]*

Nello studio con *elettromiografia laringea* [17], è stato evidenziato come le fibre muscolari laringee dei pazienti con MP, implicate nella fonazione, tendano a mantenere un pattern contrattile attivo, anche durante il riposo vocale, suggerendo una difficoltà di rilassamento (o rigidità muscolare) tipica della MP.

Infine, studi aerodinamici documentano riduzioni della pressione subglottale, della resistenza laringea e del flusso d'aria di picco. La combinazione di deficit glottale e ridotta pressione respiratoria potrebbe spiegare i cambiamenti nella qualità della voce [18].

1.3 Scopo e struttura della tesi

In questa tesi si indagano le tecniche di elaborazione dei segnali e di apprendimento automatico, per l'estrazione di informazioni clinicamente utili dai segnali vocali. Due sono gli obiettivi: differenziare i soggetti con MP dai controlli sani utilizzando un compito vocale o una combinazione di più compiti vocali che tenga conto dell'affaticamento dei pazienti, ed identificare, tramite analisi statistica, le feature acustiche maggiormente discriminanti e al contempo quanto più indipendenti da fattori esterni, quali lingua, genere, strumenti e modalità di registrazione.

Le fasi per raggiungere questo obiettivo includono: (a) lo sviluppo di nuovi algoritmi

di elaborazione del segnale vocale, (b) l'indagine su algoritmi robusti di selezione delle feature per identificare il sottoinsieme delle caratteristiche più utili alla discriminazione, (c) il confronto tra il sottoinsieme delle feature estratte dall'algoritmo di selezione e le feature selezionate dall'analisi statistica.

Nel capitolo 2 si fornisce una rassegna della letteratura sugli algoritmi di elaborazione dei segnali vocali più popolari utilizzati nelle applicazioni biomediche e si riporta un breve resoconto delle tecniche di apprendimento automatico in uso.

Nel capitolo 3 si descrive una metodologia di apprendimento automatico efficace, applicabile ad una vasta gamma di problemi che si occupano di dati ad alta dimensione.

Nel capitolo 4 vengono descritte le prove di Machine learning svolte sui vari compiti vocali.

Nel capitolo 5 si illustra l'analisi multilingua svolta considerando il compito della fonazione sostenuta /a/ su 3 lingue diverse. Vengono descritte e confrontate le prove di Machine learning effettuate considerando le feature selezionate dagli algoritmi di feature selection con le feature selezionate dall'analisi statistica.

Nel capitolo 6 si riportano, per confronto, i risultati ottenuti dall'apprendimento automatico, con i valori di accuratezza migliori, e i risultati dell'analisi statistica.

Nel capitolo 7 si confrontano i risultati ottenuti con i risultati presenti in letteratura e si suggeriscono aree di potenziale interesse per il lavoro futuro.

Nel capitolo 8 vengono tratte le conclusioni dello studio.

Capitolo 2

Stato dell'arte

Per ottenere un quadro generale del decorso della malattia Parkinson, i ricercatori hanno tentato di quantificare i disturbi neurologici manifestati dai pazienti monitorandoli con l'uso di vari tipi di sensori e sottoponendoli a diverse procedure di analisi dei segnali. La voce è risultata essere un promettente strumento per il monitoraggio della patologia perché la sua produzione risente molto delle lievi perturbazioni causate dai disturbi neurologici [19].

L'impatto della MP sulle capacità di parlare dei pazienti può essere evidenziato, secondo lo studio di J. R. Orozco-Arroyave et al. [20], nelle tre principali "dimensioni" del discorso: fonazione, articolazione e prosodia.

La fonazione è definita come la vibrazione delle corde vocali per produrre suono, l'articolazione comprende la modifica della posizione e della forma degli organi e dei tessuti coinvolti nella produzione del linguaggio, e la prosodia è la variazione di volume, altezza, e tempistica nella produzione del discorso naturale. Dal punto di vista clinico, i problemi di fonazione sono dovuti principalmente alla chiusura incompleta o asimmetrica delle corde vocali.

I deficit di articolazione si manifestano come ridotte ampiezza e velocità dei movimenti articolari delle labbra, della mandibola e della lingua, e i problemi della prosodia si manifestano con monotono, e cambiamenti nella velocità del linguaggio e con difficoltà nell'esprimere emozioni attraverso la parola.

Per valutare tali sintomi, i compiti vocali usati più frequentemente e riportati in letteratura secondo Quoc Cuong Ngo et al. [21] sono la fonazione sostenuta, ripetizioni rapide delle sillabe /pa/ /ta/ /ka/ e lettura di frasi o monologhi liberamente parlati. Il compito della fonazione sostenuta è il più usato perché le caratteristiche della sorgente vocale, del tratto vocale e degli articolatori sono

relativamente invarianti nel tempo, di conseguenza si opera in un ambiente più controllato in cui i parametri di perturbazione del rumore sono facilmente calcolabili [22].

La produzione di un discorso (un monologo, la lettura di frasi o di un testo), invece, contiene variazioni nel tono e nel volume che sono importati indicatori di qualità della voce patologica [23], ma presentano parametri difficili da ponderare. I segnali vocali sono poi comunemente analizzati utilizzando diverse misure vocali quali livello di pressione sonora, frequenza fondamentale, frequenze formanti, frequenza vocale, ritmo e altre [24].

Negli ultimi anni sono state introdotte molte tecniche di estrazione delle caratteristiche per modellare i disturbi del linguaggio e della voce. Misure acustiche come jitter e shimmer sono state utilizzate con l'obiettivo di modellare la variazione di frequenza e ampiezza durante la vibrazione delle corde vocali. Esistono parametri relativi alla respirazione come il quoziente glottide (GQ), il rapporto armonico/rumore (HNR), il rapporto rumore/armonico (NHR), il rapporto eccitazione glottale/rumore (GNE) e il rapporto di eccitazione delle corde vocali (VFER). Inoltre, la regolarità nelle fonazioni vocali è anche modellata da diversi parametri dinamici non lineari tra cui Detrended Fluctuation Analysis (DFA), Recurrence [25].

Ad oggi, come viene confermato da Jan Ruzs et al. [26], le prove sulle disfunzioni vocali si basano principalmente sulle valutazioni su una sola lingua. Ciò può sovrastimare la sensibilità delle misure sull'influenza della MP sulla parola sebbene la precisione dell'analisi acustica nel differenziare soggetti MP da soggetti CS sia elevata. Lo studio di Jan Ruzs et al. [26] valuta in che modo le anomalie motorie della parola potrebbero differire attraverso l'uso di 5 lingue diverse, e focalizza le analisi dello studio solo sul compito del parlato perché meglio si presta alla standardizzazione dello studio multilingua. Si dimostra che ci sono delle influenze del linguaggio nella qualità della voce, nell'articolare le parole e nella velocità del parlato.

L'analisi basata sul machine learning può essere correlata con caratteristiche percettive come la qualità della voce, il volume, l'altezza e la risonanza. Alcune delle feature di base, valutate e trovate adatte per la voce MP sono intensità vocale, jitter (variabilità di frequenza), shimmer (variabilità di ampiezza), armoniche al rapporto di rumore (HNR), F0, e profili di frequenza formante [27]. Le caratteristiche di produzione vocale estratte dalla forma d'onda glottidale rimuovono l'effetto dell'articolazione sul segnale acustico. La perdita di qualità della voce dei soggetti MP è associata principalmente alla chiusura incompleta delle corde vocali che porta alla fuga d'aria e quindi alla presenza di un rumore relativamente più alto nella

voce, all'abbassamento dell'intensità e alla predominanza della fase aperta del ciclo glottale. Le persone con MP hanno un jitter più alto e un HNR più basso, associati all'aperiodicità della vibrazione delle corde vocali.

Data la natura di tali parametri (quantificazioni del grado di periodicità di un segnale) è opportuno che le loro valutazioni avvengano su segnali quanto più possibile regolari, come quelli prodotti durante l'esecuzione di un vocalizzo. Numerosi studi, infatti, concentrano le loro analisi su dataset che si limitano alla fonazione sostenuta. Tra questi lo studio di Nemuel D. Pah et al. [27]; gli autori hanno ottenuto un risultato in termini di accuratezza del 59%, utilizzando un classificatore SVM e considerando feature come Jitt(abs), Jitt(rel), Shim(abs), Shim(rel), SD(pitch), HNR, NHR.

Lo studio di Rohit Lamba e al. [28] ha valutato le prestazioni sulle vocali ottenendo un'accuratezza di 97%, con 17 feature base utilizzando un classificatore XGBoost. In un altro studio, Adedolapo Aishat Teye e Suryaprakash Kompalli [29], hanno condotto due esperimenti: (1) sono state valutate 11 feature acustiche di base utilizzando a fonazione sostenuta della vocale /a/, e il classificatore SVM ha ottenuto le migliori prestazioni con accuratezza di 87,2%; (2) hanno combinato le feature di base con 24 coefficienti MFCC, e il classificatore SVM (accuratezza 98,3%) [29]. Studi che hanno analizzato il compito della lettura di frasi, come Qiyue Wang et al.[30], hanno dimostrato che l'unione delle frasi al vocalizzo migliora le prestazioni della classificazione.

Un'analisi multilingua condotta sia sul vocalizzo che sulla lettura di frasi è stata svolta da Daniel Kovac et al. [31]. Sulla base di tutte le caratteristiche estratte, gli autori hanno rilevato alterazioni vocali dovute a MP con una precisione che andava dal 67% all'85%. Per tutte le lingue combinate (inglese, israeliano, spagnolo e italiano) il modello ha raggiunto una precisione del 75% usando come classificatore XGBoost. Inoltre, lo studio ha determinato quali feature acustiche fossero importanti per le diverse lingue in termini di precisione di classificazione e quali sufficientemente robuste e indipendenti dalla lingua del parlante.

2.1 Limitazioni

Le limitazioni più importanti degli studi sopra citati sono: (a) l'eterogeneità dei set di dati analizzati, (b) il numero diverso di soggetti, (c) le registrazioni della voce e delle parole dei soggetti nei vari set di dati utilizzando protocolli di acquisizione diversi e (d) l'hardware (microfono) utilizzato.

Negli studi multilingua, i modi di svolgere compiti vocali sono diversi in ogni

paese, e non tutti i dataset presentano gli stessi compiti vocali. Inoltre, bisogna considerare che le classi non sempre sono bilanciate, con una conseguente influenza sui risultati della convalida linguistica dei modelli di apprendimento automatico.

Capitolo 3

Materiali e metodi

In questo capitolo viene descritto lo studio che è stato condotto sull'analisi acustica automatizzata dei segnali vocali.

Si inizia col rappresentare le caratteristiche del database utilizzato e col riportare le tipologie dei compiti vocali svolti dai soggetti MP e CS all'interno del database. Seguono due proposte di analisi. La prima proposta è incentrata sullo sviluppo di un protocollo di analisi vocale, in cui si utilizzano algoritmi di machine learning di classificazione. La seconda proposta è un'analisi cross-corpora multilingua, in cui si confrontano le feature selezionate dagli algoritmi di feature selection, con le feature individuate dall'analisi statistica finalizzata ad individuare le feature più robuste e meno influenzate da parametri esterni alla patologia.

3.1 Database

Il database esaminato (Tabella 3.1) comprende 354 registrazioni vocali (176 di soggetti con MP e 178 soggetti CS), ed è composto da:

1. Dataset Italiano 1 [32]: (26 soggetti con MP e 21 soggetti CS), contiene vocalizzo /a/, vocalizzo /e/ e 10 frasi;
2. Dataset Privato 1: (17 soggetti con MP e 13 soggetti CS), contiene vocalizzo /a/ e 10 frasi;
3. Dataset Colombiano [33]: (50 soggetti con MP e 13 soggetti CS), contiene vocalizzo /a/, vocalizzo /e/ e 6 frasi;
4. Dataset Ceco [34]: (22 soggetti con MP e 22 soggetti CS), contiene vocalizzo /a/;

5. Dataset Privato 2: (15 soggetti con MP e 15 soggetti CS), contiene vocalizzo /a/;
6. Dataset Italiano 2: (46 soggetti con MP e 57 soggetti CS), contiene vocalizzo /e/ e 3 frasi.

Parametri	Italiano 1	Privato 1	Colombiano	Ceco	Privato 2	Italiano 2
MP	26	17	50	22	15	46
CS	21	13	50	22	15	56
Maschi MP	17	12	25	10	11	33
Femmine MP	9	5	25	12	4	13
Maschi CS	9	8	25	11	11	15
Femmine CS	12	5	25	11	4	41
REC	MR	SP	MR	MR	SP	MR
Età media MP	67,38 ± 6,57	70,35 ± 7,23	61,14 ± 7,78	65,4 ± 9,1	70,38 ± 7,77	65 ± 6,81
Età media CS	67 ± 4,38	63,62 ± 5,8	60,9 ± 6,8	54,5 ± 17,7	59,93 ± 15,15	50,07 ± 9,66
Compiti vocali	$V_a + V_e + F$	$V_a + F$	$V_a + V_e + F$	V_a	V_a	$V_e + F$

Tabella 3.1: Dati sui soggetti presenti in ogni dataset.

MR: microfono; SP: smartphone; V_a : vocalizzo /a/; V_e : vocalizzo /e/; F : frasi

Tutti i dataset hanno una numerosità bilanciata di soggetti MP e CS, come mostrato in Figura 3.1. I dataset Italiano 1, Privato 2 e Italiano 2 risultano essere sbilanciati in termini di distribuzione di genere, con una prevalenza maschile per i soggetti con MP mentre il dataset Italiano 2 ha una prevalenza di soggetti femminili per i controlli sani.

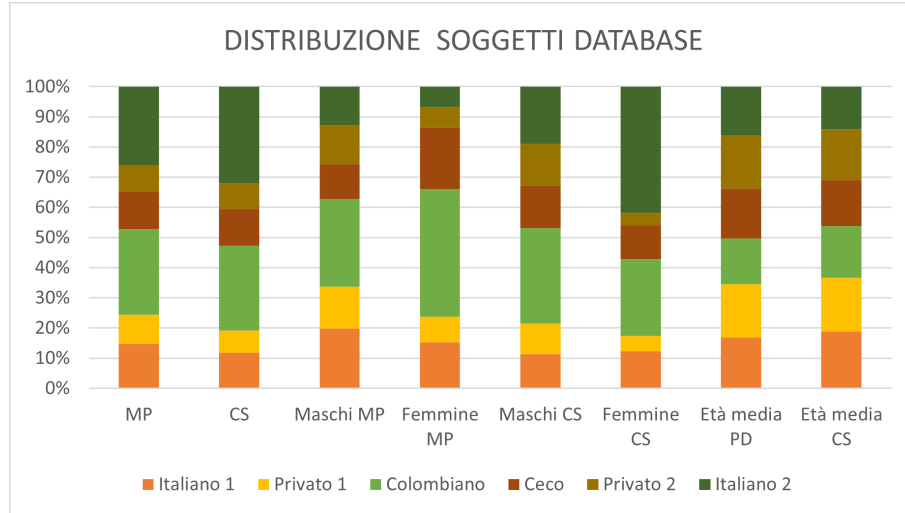


Figura 3.1: Grafico delle distribuzioni dei soggetti nel database.

MP: Malattia di Parkinson; CS: Controlli Sani

Non tutti i dataset sono stati registrati alla stessa frequenza di campionamento. Privato 1 e Ceco sono stati registrati a 48 kHz, Colombiano e Italiano 2 sono stati registrati a 44.1 kHz e Italiano 1 è stato registrato a 16 kHz. E' stato effettuato un ricampionamento alla frequenza di campionamento più bassa, 16 kHz, per garantire omogeneità nelle analisi.

3.2 Compiti vocali

Per la valutazione dei disturbi del linguaggio, sono stati utilizzati due diversi compiti vocali: la *fonazione sostenuta della vocale* (o vocalizzo) e le *frasi* [19]. I dataset utilizzati sono Italiano 1, Privato 1, Colombiano e Italiano 2, perché contenenti entrambi i compiti vocali.

3.2.1 Fonazione sostenuta

In letteratura sono trattati prevalentemente studi in cui la vocale maggiormente impiegata è la /a/, ciò perché è il suono più semplice da produrre [35]. In questo studio si è deciso di effettuare un preliminare confronto tra 2 vocali per valutare la presenza di eventuali differenze; è stato quindi impiegato anche il vocalizzo con la vocale /e/.

Il compito vocale che il paziente assolve, consiste nell'eseguire in modo costante e il più a lungo possibile, una fonazione sostenuta della vocale (/a/ oppure /e/).

Questo compito vocale si basa sull'ipotesi che, mentre un soggetto sano può suscitare una fonazione stazionaria, un soggetto con compromissione vocale non possa effettuarlo. I silenzi iniziali e finali, non necessari per l'analisi, sono stati rimossi manualmente.

3.2.2 Frasi

I dataset sono stati realizzati utilizzando frasi diverse, ad eccezione dei dataset Italiano 1 e Privato 1 che utilizzano le stesse frasi. L'organizzazione dei dataset è la seguente:

1. Italiano 1 e Privato 1: una registrazione contiene 10 frasi;
2. Colombiano: una registrazione per ogni frase (6 registrazioni per 6 frasi);
3. Italiano 2: una registrazione contiene 3 frasi;

Per ottenere una registrazione per frase, come prima operazione, sono state segmentate manualmente le registrazioni dei dataset Italiano 1, Privato 1 e Italiano 2;

successivamente, sono stati isolati i segmenti sonori tramite segmentazione automatica, utilizzando il software PRAAT [36], e uniti per formare una registrazione unica con tutti i segmenti sonori.

3.3 Feature vocali

Per questo studio, in totale, sono state considerate 498 feature vocali che possono essere classificate nelle seguenti categorie:

- Caratteristiche basate su F0
- Caratteristiche di frequenza formante (FORM)
- Varianti del jitter
- Varianti di shimmer
- Coefficienti di frequenza Mel Cepstral (MFCC)
- Rapporto tra armoniche e rumore (HNR)
- Misure temporali e spettrali della qualità della voce (CPP, SPEC)
- Rapporto di eccitazione glottale-rumore (GNE)
- Coefficienti predittivi lineari percettivi (PLP)
- Analisi delle fluttuazioni senza tendenza (DFA)

Per ogni categoria, fatta eccezione per jitter e shimmer che forniscono già un'informazione relativa all'interno del segnale, le feature sono state estratte su finestre di 40 ms (20 ms di overlap), quindi sono stati valutati gli indici statistici di media, mediana, deviazione standard, skewness e kurtosis. Per i coefficienti MFCC e PLP, sono state calcolate, per ogni indice statistico, le derivate prima e seconda nel tempo.

3.3.1 Caratteristiche basate su F0

La frequenza fondamentale (o pitch) dell'emissione vocale di un parlante dipende dalle caratteristiche fisiche delle corde vocali dell'individuo; essa varia da soggetto a soggetto e può essere modificata, dallo stesso parlante, con azioni fisiche variando il livello di tensione delle corde vocali. Le corde vocali degli uomini sono lunghe tra i 20 e 25 mm, mentre quelle delle donne e dei bambini sono lunghe tra i 15 e 20 mm; il pitch degli uomini, quindi, è generalmente più basso [37].

- Range frequenza fondamentale maschile: 50 - 200 Hz
- Range frequenza fondamentale femminile: 150 - 300 Hz
- Range frequenza fondamentale bambini: 200 - 500 Hz

Dalla letteratura si ricava che le alterazioni nella fonazione nei soggetti con MP sono dovute ai cambiamenti in F0, alla variabilità di F0 e all'intensità di F0. L'irrigidimento della muscolatura laringea, che si verifica con MP, potrebbe portare all'aumento di F0 [38].

3.3.2 Caratteristiche di frequenza formante (FORM)

Le frequenze di risonanza sono dette frequenze formanti, e sono determinate dalla forma del condotto vocale che dipende dalla posizione degli organi mobili, dall'età e dal sesso dell'individuo. Donne e bambini hanno un apparato vocale più corto degli uomini, di conseguenza i valori delle frequenze formanti saranno più elevati.

Le formanti dipendono quasi esclusivamente dalla conformazione istantanea delle cavità epilaringee e risultano indipendenti dalle armoniche, le quali dipendono dalle caratteristiche del suono prodotto dal movimento delle corde vocali [39].

Le formanti si ottengono per la risonanza del tratto vocale. Nei pazienti affetti da MP le frequenze formanti sono influenzate dal movimento limitato degli articolatori. Le frequenze formanti calcolate sono tre: F1, F2, F3.

3.3.3 Varianti del jitter

Il *Jitter* rappresenta la famiglia di caratteristiche che misura la perturbazione a breve termine di F0 [40]. Le varianti considerate sono:

Jitter absolute: è la media dei valori assoluti delle differenze del generico periodo T_0^i del segnale vocale e il successivo T_0^{i+1} , diviso per il numero totale di periodi N del segnale vocale [40].

$$Jitter_abs = \frac{1}{N-1} \sum_{i=1}^{N-1} |T_0^i - T_0^{i+1}|$$

Jitter relative: rapporto tra la differenza media assoluta tra due periodi consecutivi e il periodo medio [40].

$$Jitter_rel = \frac{\frac{1}{N-1} \sum_{i=1}^{N-1} |T_0^i - T_0^{i+1}|}{\frac{1}{N} \sum_{i=1}^N T_0^i}$$

Relative average perturbation (RAP): variabilità del pitch con un fattore di smoothing di tre periodi [40].

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

Pitch Period perturbation quotient (PPQ5): variabilità del pitch con un fattore di attenuazione di cinque periodi [40].

$$PPQ5 = \frac{\frac{1}{N-4} \sum_{i=3}^{N-2} \left| \frac{T_0^{i-2} + T_0^{i-1} + T_0^i + T_0^{i+1} + T_0^{i+2}}{5} - T_0^i \right|}{\frac{1}{N} \sum_{i=1}^N T_0^i}$$

3.3.4 Varianti di shimmer

Shimmer è l'analogo del jitter ma considera come variabile l'ampiezza massima di picco del segnale [41]. Le varianti considerate sono :

Shimmer (absolute): misura la variabilità dell'ampiezza A^i picco-picco su N impulsi considerati.

$$Shimmer_abs = \frac{1}{N-1} \sum_{i=1}^{N-1} |A^i - A^{i+1}|$$

Shimmer (rel): rapporto tra la media dei valori assoluti delle differenze tra le ampiezze A^i dei periodi consecutivi e l'ampiezza media su N impulsi estratti [40].

$$Shimmer_rel = \frac{\frac{1}{N-1} \sum_{i=1}^{N-1} |A^i - A^{i+1}|}{\frac{1}{N} \sum_{i=1}^N A^i}$$

Shimmer (abs, dB): Shimmer (abs) in decibel [40].

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

Shimmer (APQ3): differenza media assoluta tra l'ampiezza di un periodo e la media delle ampiezze del periodo e dei suoi 3 periodi vicini, diviso per l'ampiezza media [40]

$$APQ3 = \frac{\frac{1}{N-2} \sum_{i=2}^{N-1} \left| \frac{A^{i-1} + A^i + A^{i+1}}{3} - A^i \right|}{\frac{1}{N} \sum_{i=1}^N A^i}$$

Shimmer (APQ5): differenza media assoluta tra l'ampiezza di un periodo e la media delle ampiezze del periodo e dei suoi 5 periodi vicini, diviso per l'ampiezza media [40]

$$APQ5 = \frac{\frac{1}{N-4} \sum_{i=3}^{N-2} \left| \frac{A^{i-2} + A^{i-1} + A^i + A^{i+1} + A^{i+2}}{5} - A^i \right|}{\frac{1}{N} \sum_{i=1}^N A^i}$$

3.3.5 Coefficienti di frequenza Mel Cepstral (MFCC)

Studi psicoacustici indicano una relazione non lineare tra la frequenza reale di un suono e la frequenza effettiva percepita; sulla base di tale relazione si determinano i coefficienti di frequenza MFCC. Il calcolo degli MFCC si basa sulla scala Mel che stima la frequenza percepita dell'orecchio umano [42].

Nelle applicazioni con voci patologiche, i coefficienti MFCC hanno dimostrato la capacità di simulare il movimento irregolare o la mancanza di chiusura delle corde vocali. Infatti, le alterazioni legate alla vibrazione delle corde vocali (rumore di modulazione) si riflettono nelle bande basse dell'MFCC, mentre le componenti di turbolenza (rumore additivo) si riflettono sulle le bande più alte dell'MFCC.

3.3.6 Rapporto tra armoniche e rumore (HNR)

Le misure relative al rumore sono calcolate per stimare il grado di rumore presente nel segnale vocale. Il parametro prescelto è l'HNR. La quantificazione dell'HNR si basa sul calcolo separato delle componenti armoniche e del rumore del segnale. Per ogni frame, si determina il logaritmo del rapporto tra l'energia armonica e il rumore. La media di tutti i valori fornisce l'HNR finale [40].

3.3.7 Misure cepstrali e spettrali della qualità della voce (CPP)

Le misure spettrali sono altamente correlate con la valutazione della fonazione patologica, sia nei compiti di vocale sostenuta che nel compito delle frasi [43]. Esempi di parametri spettrali sono il centro di gravità, la deviazione standard, la skewness e la kurtosis, descrittori dello spettro del segnale vocale. Parametro significativo per valutare il grado della patologia è il *Cepstral peak prominence* (CPP) che risulta in forte correlazione con la mancanza del respiro sia nelle vocali sostenute che nelle frasi. Il CPP è una misura normalizzata dell'ampiezza di picco cepstrale, che confronta il livello di organizzazione armonica del discorso con il rumore di fondo cepstrale risultante dall'aspirazione.

3.3.8 Rapporto di eccitazione glottale-rumore (GNE)

Glottal-to-noise excitation ratio (GNE) è un'affidabile misura del livello di rumore, anche in presenza di forti variazioni di ampiezza e periodicità. Il GNE si basa sulla correlazione fra gli involuppi di Hilbert su diversi canali di frequenza. Ad ogni chiusura della glottide, tutti i canali di frequenza vengono eccitati. Ricavando gli involuppi di Hilbert da tutti i canali, gli involuppi, condividendo la stessa forma,

possiedono un'alta correlazione tra loro. La forma di ogni impulso di eccitazione è praticamente indipendente dagli impulsi precedenti o successivi. In caso di segnali turbolenti (rumore, sussurro) un rumore a banda stretta viene eccitato in ogni canale di frequenza. Questi rumori a banda stretta non sono correlati. In questo modo si può ottenere che le deviazioni dalla periodicità non influenzano il grado di turbolenza misurato da GNE [44].

3.3.9 Coefficienti predittivi lineari percettivi (PLP)

Perceptual linear predictive coefficients: i coefficienti PLP possono essere considerati una codifica percettiva del segnale vocale; essi mirano a replicare i meccanismi percettivi dell'udito [40].

3.3.10 Analisi delle fluttuazioni senza tendenza (DFA)

Il *Detrended fluctuations analysis (DFA)* è uno stimatore di autosimilarità che serve a calcolare le componenti stocastiche nel parlato dovute al rumore di aspirazione. Se $DFA < 0.5$ il segnale è non correlato (caratteristica del rumore bianco); se $DFA = 0.5$, il segnale è anti-correlato (tendenze negative); se $DFA > 0.5$, ci sono correlazioni positive (tendenze positive) [40].

3.4 Algoritmo di classificazione

Dopo aver estratto le 498 feature acustiche, con l'utilizzo della libreria Parselmouth di Python [45], è stato implementato un algoritmo di classificazione, (Figura 3.2) con i seguenti step:

1. Divisione 70% trainset e 30% testset;
2. Normalizzazione (z-score, massimo-minimo);
3. Feature selection (due algoritmi: Boruta e Cross-correlazione);
4. Classificatori considerati:
 - a K Neighbors Classifier (KNN)
 - b Linear SVM (LSVM)
 - c Gaussian Process Classifier (GPr)
 - d Decision Tree Classifier (DT)
 - e Random Forest Classifier (RF)
 - f AdaBoost Classifier (AdaB)

- g Gaussian Naive Bayes (NB)
 - h Linear Discriminant Analysis (LDA)
 - i XGB Classifier (XGB)
 - l Artificial Neural Net (NN)
5. K-fold cross-validation applicato al training set ($k=5$);
 6. Calcolo delle prestazioni

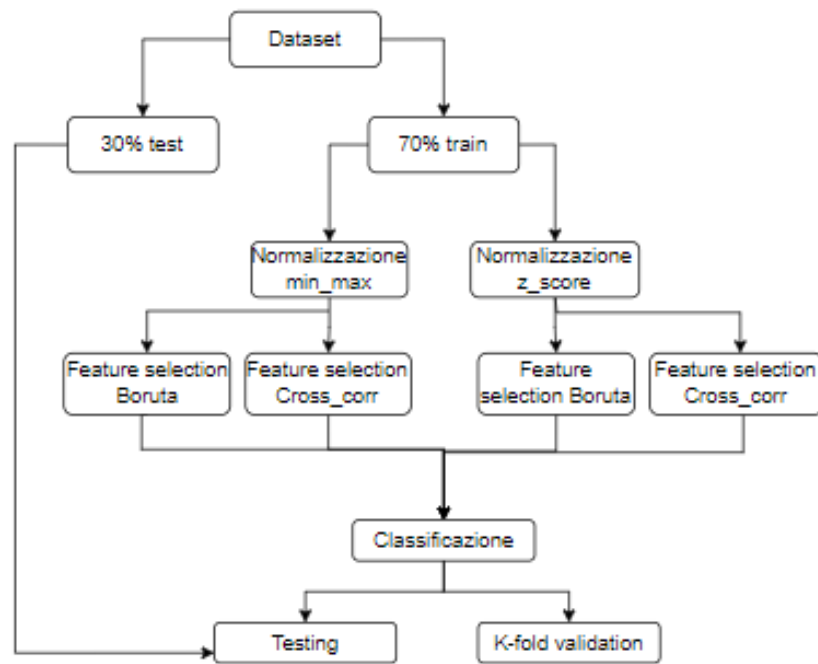


Figura 3.2: Diagramma di flusso algoritmo di classificazione.

3.4.1 Normalizzazione

Lo scopo della normalizzazione consiste nel modificare i valori delle colonne numeriche (nel nostro caso le feature acustiche) nel set di dati, utilizzando una scala comune senza distorsioni negli intervalli di valori o perdita di informazioni. Con la normalizzazione si conservano la distribuzione e le proporzioni generali dei dati di origine, mantenendo i valori all'interno di una scala applicata in tutte le colonne usate nel modello [46].

E' una delle operazioni di pre-processing più utili perché migliora la qualità dei risultati finali. Inoltre, la normalizzazione riduce il tempo in cui l'algoritmo di apprendimento converge al risultato finale e migliora l'efficacia del modello. In questo studio si sono utilizzate due normalizzazioni: Z_score e Minimo_massimo. L'operazione di normalizzazione è stata applicata dopo la divisione in 70% trainset e 30% testset.

Z_score

La normalizzazione Z_score si riferisce al processo di normalizzazione di ogni colonna di un set di dati in modo tale che la media di tutti i valori sia 0 e la deviazione standard sia 1. Basandosi sul calcolo di media e deviazione standard, consente di minimizzare l'influenza degli outliers eventualmente presenti.

$$Z = \frac{x - \text{mean}(x)}{\text{dev_stand}(x)}$$

Minimo_massimo

La normalizzazione min-max riscalda in modo lineare ogni colonna di un set di dati nell'intervallo $[0,1]$.

$$Z = \frac{x - \min(x)}{\max(x) - \min(x)}$$

3.4.2 Feature selection

Feature selection è il processo di riduzione del numero di feature di input durante lo sviluppo di un modello predittivo. È importante ridurre il numero di variabili di input sia per ridurre il costo computazionale della modellazione sia, in alcuni casi, per migliorare le prestazioni del modello [47].

Cross-correlazione

L'algoritmo di feature selection basato sulla cross-correlazione individua il set di feature maggiormente correlate con la classe da discriminare. Vengono settate due soglie:

- th1: valore di soglia della cross-correlazione con la classe, settata a 0.03;
- th2: valore di soglia del p_value, settata a 0.05

Il primo step dell'algoritmo è il calcolo della cross-correlazione delle feature del trainset; di queste feature correlate, si escludono quelle che, in valore assoluto, presentano una correlazione con la classe $<$ di th1 e un p_value associato minore

di 0.05. Infine si verifica l'esistenza di feature ridondanti: si valuta la correlazione tra coppie di feature e, in presenza di valori di correlazione intra-feature superiori rispetto ai valori feature-classe, si seleziona solo la feature maggiormente correlata con la classe.

Boruta

In questo studio è stata usata la versione di Boruta per Python, chiamata BorutaPy, funzionalità di `SelectFromModel` della libreria di `scikit-learn` [48]. L'algoritmo si basa sull'addestramento del classificatore RF su un set di dati ricavato dall'unione del trainset e di un dataframe creato mescolando casualmente tutte le feature del trainset, chiamate feature d'ombra. Dopo l'addestramento, si calcola Mean Decrease Accuracy per determinare l'importanza di ciascuna feature.

Ad ogni iterazione, l'algoritmo Boruta controlla se una feature reale supera la soglia di importanza, definita come l'importanza più alta tra le feature d'ombra. Per cui, il set di feature che individua Boruta sarà formato dalle feature reali che hanno un'importanza maggiore della soglia. Il numero di ripetizioni dell'algoritmo è settato a 100.

3.4.3 Classificatori

La scelta di un classificatore appropriato può significativamente migliorare il risultato della classificazione [21]. In questo studio sono stati utilizzati i classificatori più usati in letteratura per discriminare soggetti con MP e soggetti CS [49].

- **K Neighbors:** è un metodo di apprendimento basato su istanze o apprendimento non generalizzante. Il classificatore non tenta di costruire un modello interno generale, ma memorizza semplicemente istanze dei dati di addestramento. La classificazione è calcolata da un "voto di maggioranza" dei *nearest neighbors* di ogni "punto di interrogazione", ovvero viene assegnato il target che si presenta più frequentemente attorno a un dato punto di interrogazione. L'insieme dei punti che formano il punto di interrogazione è settato dall'utente, (k=5 di default) [50].
- **SVM:** è un metodo di apprendimento supervisionato utilizzato per la classificazione. L'obiettivo del SVM con kernel lineare è quello di trovare un iperpiano che separi, al miglior grado possibile, i punti di dati di una classe da quelli di un'altra classe. Per "miglior grado possibile" si intende l'iperpiano che ha il margine maggiore tra le due classi. Il margine maggiore è la larghezza massima della linea parallela all'iperpiano che non presenta punti di dati interni [51].

- **Gaussian Process:** è un di apprendimento supervisionato non parametrico utilizzato per la classificazione. La classificazione del processo gaussiano (GPC) si basata sull'approssimazione di Laplace, la quale approssima il posteriore non gaussiano con una gaussiana. Come impostazione predefinita è stato usato l kernel "1.0 * RBF(1.0)" [52].
- **Decision Tree:** è un metodo di apprendimento supervisionato non parametrico utilizzato per la classificazione. L'obiettivo è creare un modello che preveda il valore di una variabile target apprendendo semplici regole decisionali dedotte dalle caratteristiche dei dati [53]. Il parametro `max_depth` permette di controllare le dimensioni dell'albero per evitare l'overfitting, ed è stato settato a 5.
- **Random Forest:** modello di apprendimento supervisionato ensemble, basato sull'utilizzo di più modelli addestrati sugli stessi dati, calcolando la media dei risultati di ciascun modello alla fine trovando un risultato di previsione/classificazione più potente. Nel caso del RF i modelli utilizzati sono i DT. I principali parametri da regolare quando si utilizzano questi metodi sono `n_estimators` e `max_feature`. Il primo è il numero di alberi nella foresta, settato a 10. Il secondo è la dimensione dei sottoinsiemi casuali di feature da considerare quando si divide un nodo. Il valore di default settato è `max_feature=1` [54].
- **Ada Boost :** è un metodo di ensemble iterativo che combina più classificatori per aumentare la precisione dei classificatori utilizzati. Il classificatore AdaBoost combina più classificatori con prestazioni non ottimali in modo da ottenere un classificatore robusto ad alta precisione. Il concetto di base di Adaboost è impostare i pesi dei classificatori e addestrare il campione di dati ad ogni iterazione in modo tale da garantire previsioni accurate [55].
- **Gaussian Naive Bayes:** è un metodo probabilistico di apprendimento automatico utilizzato per la classificazione e si basa sul teorema di Bayes con l'assunzione "naive" di indipendenza condizionale tra ogni coppia di caratteristiche dato il valore della variabile di classe [56].
- **Linear Discriminant Analysis:** è un classificatore di apprendimento supervisionato classico che utilizza una superficie di decisione lineare, proiettando i dati di input su un sottospazio lineare costituito dalle direzioni che massimizzano la separazione tra le classi [57].
- **XGB:** è un algoritmo di potenziamento basato sull'algoritmo dei DT potenziati dal gradiente. Consente l'ottimizzazione di funzioni di perdita differenziabili arbitrarie. In ogni fase il numero di alberi di regressione sono adattati al gradiente negativo della funzione di perdita. La funzione di perdita settata di default è la funzione logaritmica [58].

- **Artificial Neural Network:** sono algoritmi supervisionati. Il modello di rete utilizzato è il perceptrone multistrato (MLP) feedforward che mappa insiemi di dati di input su un insieme di output appropriati. Un MLP è costituito da più livelli e ogni livello è completamente connesso a quello successivo. I nodi degli strati sono neuroni che utilizzano funzioni di attivazione non lineari, ad eccezione dei nodi dello strato di input. Possono esserci uno o più livelli nascosti non lineari tra il livello di input e quello di output [59]. Il numero di iterazioni settato è 1000.

3.4.4 K-fold cross-validation

Nell'apprendimento supervisionato, per evitare la situazione di overfitting, si utilizza un "set di validazione" che viene utilizzato dopo l'addestramento per effettuare una valutazione dei risultati.

Nell'approccio di base, chiamato k-fold Cross-validation, il training set è suddiviso in k set più piccoli e per ogni k, il modello viene addestrato su k-1 set di dati per l'allenamento e si testa sul k restante. La misura delle prestazioni della validazione consiste nella media dei valori calcolati nel ciclo per tutti i k. Per la procedura è stata utilizzata la funzione `cross_val_score` di `scikit-learn`. [60]

3.4.5 Metriche di valutazione

Le metriche di valutazione utilizzate per misurare il successo dei classificatori sono accuratezza, sensibilità, specificità e MCC [61].

Accuratezza

Per un classificatore binario, l'accuratezza è il rapporto tra le istanze corrette e quelle intere totali:

$$\text{Accuratezza} = \frac{TP+TN}{TP+TN+FP+FN}$$

True Positive, o TP, sono casi con target positivo che sono stati correttamente classificati come positivi.

True Negative, o TN, sono casi con target negativo che sono stati correttamente classificati come negativi.

False Positive, o FP, sono casi con target negativo che sono stati erroneamente classificati come positivi.

False Negative, o FN, sono casi con target positivo che sono stati erroneamente classificati come negativi.

Sensibilità

Sensibilità è una misura statistica di istanze positive correttamente classificate:

$$\text{Sensibilità} = \frac{TP}{TP+FN}$$

Specificità

Specificità è una misura statistica di istanze negative correttamente classificate:

$$\text{Specificità} = \frac{TN}{TN+FP}$$

MCC

MCC è una misura che mostra la qualità della classificazione, facendo la correlazione tra le classificazioni binarie previste e osservate, ottenendo un valore tra -1 e +1.

$$\text{MCC} = \frac{TP*TN-FP*FN}{\sqrt{(TP+FP)(TP+FN)(TN+FP)(TN+FN)}}$$

Area under the ROC curve (AUC)

La curva ROC è un grafico che mostra le prestazioni di un modello di classificazione per tutte le soglie di classificazione. L'AUC misura l'area bidimensionale sotto l'intera curva ROC [62] e fornisce una misura del rendimento di un classificatore al variare delle possibili soglie di classificazione. L'AUC ha valori compresi tra 0 e 1.

F1 score

Il punteggio F1 può essere interpretato come una media armonica della precisione e del recall, dove un punteggio F1 raggiunge il suo valore migliore a 1 e il punteggio peggiore a 0 [63].

$$\text{F1_score} = \frac{2 * \text{precision} * \text{recall}}{\text{precision} + \text{recall}}$$

3.5 Test statistici

In questo paragrafo si descrivono i test statistici che sono stati utilizzati per l'analisi statistica delle feature.

Prima di procedere con i test di significatività, è stata verificata la normalità delle feature all'interno dei vari dataset con il test di **Shapiro-Wilk** [64].

Dopo aver valutato la normalità, sono stati eseguiti i test di significatività non parametrici di **U-Mann Whitney** [65] e di **Kruskal Wallis** [66]. Il primo è stato utilizzato per valutare la significatività statistica delle feature tra i due gruppi MP

e CS. Il secondo test è stato utilizzato per valutare la significatività statistica delle feature con i parametri esterni, quali il genere, la tecnica di registrazione e la lingua. I test sopra citati sono stati svolti utilizzando il software di analisi statistica Jamovi [67] ed Excel.

3.5.1 Shapiro-Wilk test

Si tratta di un test per valutare la normalità dei dati all'interno di un dataset.

L'**ipotesi nulla (H0)** del test di Shapiro-Wilk afferma che tutti i valori all'interno del dataset abbiano una distribuzione normale.

Il test respinge l'ipotesi nulla quando il p_value è inferiore o uguale a 0.05, e permette di affermare che i dati non si adattano alla distribuzione normale per cui è stato trovato un allontanamento significativo dalla normalità.

Quando il p_value è maggiore di 0.05, l'ipotesi nulla non è rifiutata e il test consente di dichiarare che non è stato trovato un allontanamento significativo dalla normalità dei dati.

3.5.2 U-Mann Whitney test

Il test di Mann-Whitney è stato usato per confrontare se esiste una differenza nelle variabili dipendenti, le feature, per i due gruppi indipendenti, soggetti con MP e CS [68]. I presupposti per effettuare il test sono che i due gruppi devono essere indipendenti e che le variabili dipendenti siano numeriche.

L'**ipotesi nulla (H0)** del test di Mann-Whitney afferma che i campioni di una popolazione abbiano distribuzione di probabilità eguali.

L'**ipotesi alternativa (H1)** afferma che i campioni di una stessa popolazione abbiano una distribuzione di probabilità statisticamente diversa.

Se il p_value è inferiore al valore 0.05, il test rifiuta l'ipotesi nulla e si può presumere che esiste almeno una differenza statisticamente significativa tra i campioni della stessa popolazione.

Se il p_value è maggiore di 0.05, l'ipotesi nulla non è rifiutata e si può dichiarare che non è stata trovata differenza statisticamente significativa tra i campioni della stessa popolazione.

3.5.3 Kruskal Wallis test

Il test di Kruskal-Wallis è un'estensione del test di Mann-Whitney, nel caso in cui si considerino più di due gruppi di variabili indipendenti, ed è la versione non parametrica del test ANOVA [69].

L'**ipotesi nulla (H0)** del test Kruskal-Wallis afferma che i campioni provengano da popolazioni statisticamente identiche.

L'**ipotesi alternativa (H1)** afferma che almeno un campione proviene da una popolazione con una distribuzione statisticamente diversa.

Se il p-value calcolato è minore del livello di significatività di 0.05, l'ipotesi nulla viene rifiutata.

Se il p-value è maggiore, l'ipotesi nulla viene mantenuta per cui si può affermare che i campioni provengono da popolazioni statisticamente identiche.

Capitolo 4

Machine Learning per la classificazione dei segnali vocali

In questa sezione sono descritte le prove eseguite per determinare quali compiti vocali siano utili per la discriminazione tra soggetti con MP e soggetti CS.

Lo scopo è quello di identificare un compito vocale o una combinazione di compiti vocali in grado di analizzare le diverse caratteristiche della voce, minimizzando la possibilità di affaticamento dei pazienti.

4.1 Valutazione dell'influenza della tecnica di registrazione

Data la presenza di 2 dataset differenziati solo dalla tecnica di registrazione, è stata condotta un'analisi preliminare volta a valutare l'influenza delle condizioni di acquisizione sulle prestazioni del modello. Successivamente, al fine di limitare la complessità delle analisi, i due dataset verranno uniti e trattati come unico per gli esperimenti successivi.

Obiettivo: valutare come l'utilizzo di strumentazione a basso costo influisca sulle prestazioni del modello;

Set di dati utilizzato: per la prova sono stati utilizzati i dataset Privato 1 e Italiano 1, caratterizzati dallo stesso protocollo di registrazione, salvo per la strumentazione utilizzata;

Analisi effettuata: la prova è stata condotta in due fasi. La prima prevede l'applicazione dell'algoritmo di classificazione al dataset Italiano 1. La seconda, viene applicata al dataset somma di Italiano 1 e Privato 1. Una volta aggiunto il dataset Privato 1, procede come nella fase prima applicando l'algoritmo del punto 3.4. Le prestazioni ottenute dalle fasi sono state poi confrontate per valutare l'influenza della tecnica di registrazione, come mostrato in Figura 4.1.

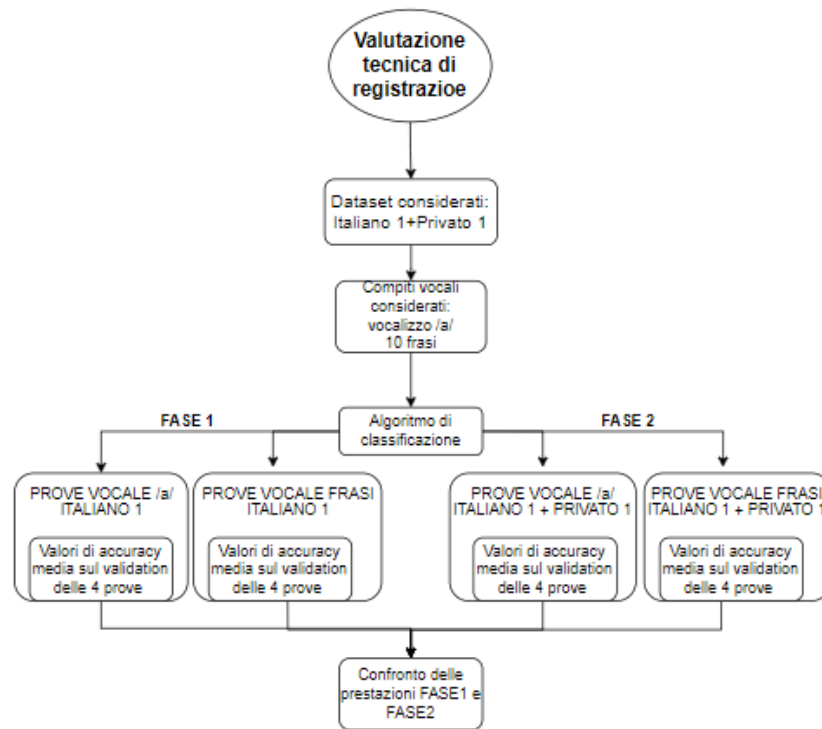


Figura 4.1: Diagramma di flusso valutazione tecnica di registrazione

4.2 Identificazione compiti vocali da utilizzare nel protocollo

Questo secondo esperimento è stato condotto al fine di valutare l'esistenza di una combinazione di esercizi vocali capaci di catturare le alterazioni dei pazienti MP riducendo lo sforzo fisico richiesto.

4.2.1 Scelta della vocale per il compito della fonazione sostenuta

Obiettivo: valutare quale vocale, tra /a/ ed /e/, sia la più indicata nel compito del vocalizzo per discriminare al meglio pazienti MP dai soggetti CS;

Set di dati utilizzato: per la prova sono stati utilizzati i dataset Italiano 1 e Colombiano, perché contengono le registrazioni dei due vocalizzi /a/ ed /e/;

Analisi effettuata: applicando l'algoritmo di classificazione descritto al punto 3.4 ai dataset Italiano 1 e Colombiano, per i vocalizzi /a/ ed /e/, in totale, sono state effettuate 16 prove (Figura 4.2), 4 per ognuno dei due dataset al variare di algoritmi di classificazione e feature selection. Ottenuti i valori dell'accuratezza, sono stati messi a confronto per valutare quale vocale avesse accuratezza maggiore sul valid.

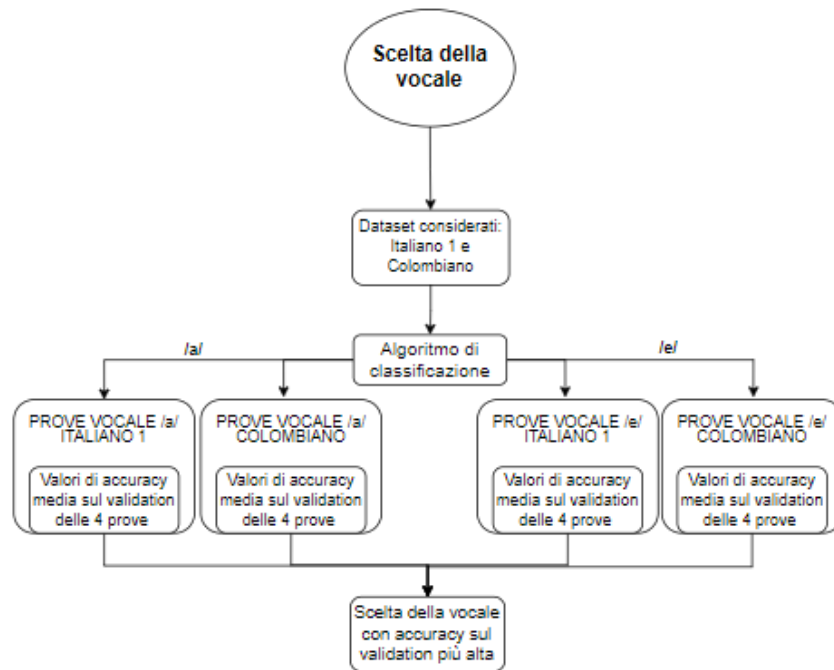


Figura 4.2: Diagramma di flusso scelta della vocale

4.2.2 Scelta delle frasi

Obiettivo: trovare le frasi che consentano di discriminare al meglio i pazienti MP dai soggetti CS e valutare se l'utilizzo di una singola frase sia sufficiente.

Set di dati utilizzato: i dataset utilizzati per la prova sono stati Italiano 1, Privato 1, Colombiano e Italiano 2. Italiano 1 e Privato 1 sono stati considerati come un unico dataset così da formare un dataset con una numerosità più consistente. La divisione in train e test è stata fatta bilanciando equamente le percentuali dei due dataset.

Analisi effettuata: ad ogni frase dei dataset è stato applicato l'algoritmo di classificazione, i risultati sono stati messi a confronto e sono state individuate le due frasi con valore di accuratezza media maggiore Figura 4.3.

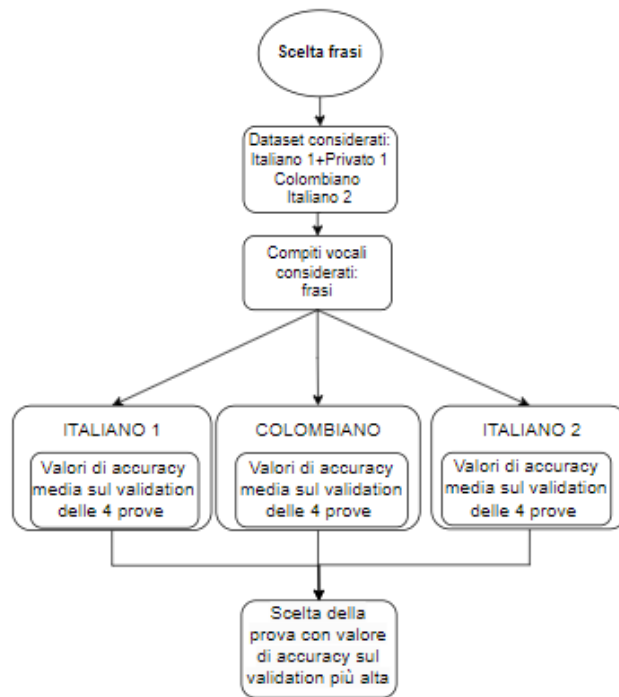


Figura 4.3: Diagramma scelta frasi

4.2.3 Protocollo

Obiettivo: individuare il compito vocale o la combinazione vocalizzo e frasi, che fornisca la migliore discriminazione tra pazienti MP e soggetti CS.

Set di dati utilizzato: sono stati considerati i dataset contenenti sia le frasi che la fonazione sostenuta della vocale che dal setup precedente ha riportato le prestazioni migliori (Italiano e Privato 1, Colombiano, Italiano 2).

Analisi effettuata: sono state confrontate le prestazioni dei compiti vocali del vocalizzo e della frase migliori, della combinazione fonazione sostenuta unita ad una frase e della combinazione fonazione sostenuta unita a due frasi, col fine di individuare il compito vocale più discriminante e che tenga conto del tempo di esecuzione. Nel caso della combinazione dei 2 compiti vocali, la fusione è stata effettuata tramite tecnica di *early fusion* (i set di feature vengono uniti prima della classificazione). E' stato, infatti, dimostrato in uno studio precedente che tale tecnica consente di ottenere risultati migliori rispetto alla *late fusion* [70]. In Figura 4.4 si illustra il diagramma di flusso dell'analisi.

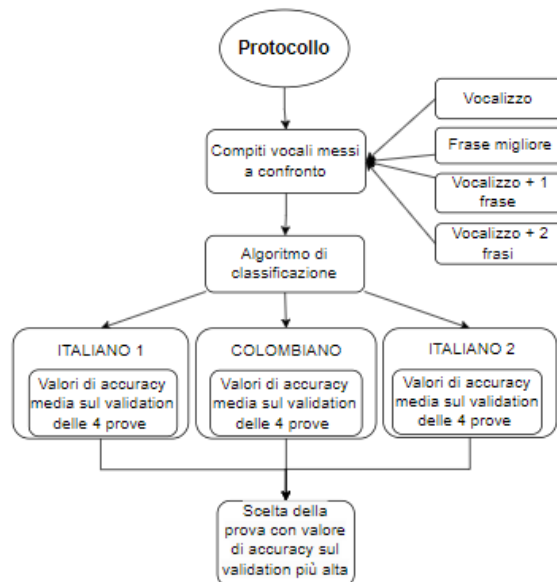


Figura 4.4: Diagramma di flusso protocollo

Capitolo 5

Analisi Multilingua

Questa sezione riguarda l'analisi cross-corpora multilingua, che esamina il compito vocale della fonazione sostenuta della vocale /a/. Essendo gli obiettivi di questa analisi l'esplorazione e l'individuazione dei parametri che influiscono maggiormente nella classificazione tra soggetti con MP e soggetti di CS e siano al contempo indipendenti dal genere, dalla lingua e dalla tecnica di registrazione, ci si è soffermati solo sul vocalizzo così da aumentare la numerosità dei dati analizzati (il vocalizzo /a/, infatti, è riportato in ampia misura in vari dataset disponibili online). I dataset presi in esame sono Italiano 1, Privato 1, Colombiano, Ceco, Privato 2; essi sono considerati come un unico dataset trilingue: Italiano, Spagnolo e Ceco.

L'analisi Multilingua si sviluppa in 3 step:

1. **analisi di machine learning**, in cui si applica l'algoritmo di classificazione e si ricavano le prestazioni del modello;
2. **analisi statistica**, effettuata su tutte le feature acustiche, in cui si punta a valutare il subset più robusto considerando la lingua, il genere e l'appartenenza al dataset come parametri esterni;
3. **feature robuste e machine learning**, in cui si applica l'algoritmo di classificazione al set di feature individuato dall'analisi statistica e si ricavano le prestazioni del modello.

5.1 Machine learning

Sono state apportate delle modifiche all'algoritmo di classificazione a causa della numerosità elevata del dataset: (1) l'algoritmo di feature selection considerato è Boruta e (2) le ripetizioni per il K-fold cross-validation sono 20. Sono state calcolate quindi le metriche di valutazione del modello, come mostrato in Figura 5.1.

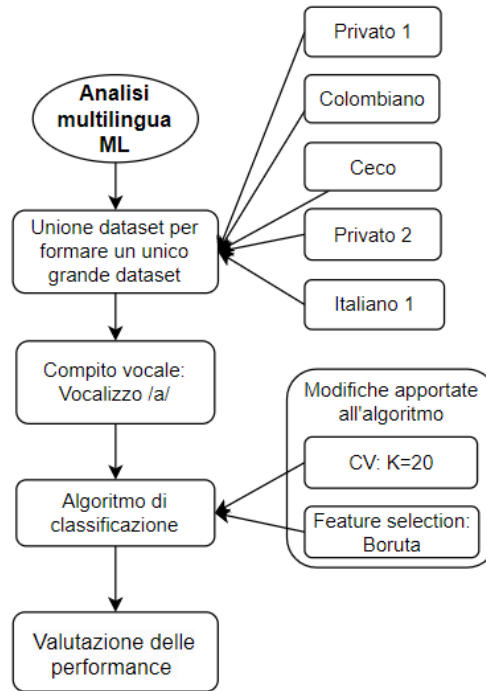


Figura 5.1: Diagramma di flusso applicazione algoritmo di classificazione considerando l'unione di 5 dataset

5.2 Analisi statistica

L'analisi statistica è finalizzata a trovare un set feature acustiche robuste per valutare le alterazioni della voce causate dalla MP. Sono stati condotti due esperimenti paralleli:

1. Analisi sui 5 dataset uniti;
2. Analisi su ogni dataset;

L'esperimento 1 (Analisi sui 5 dataset uniti), individua il gruppo di feature significative in almeno 1 dataset e confronta poi i risultati ottenuti sui vari insiemi di dati; mentre l'esperimento 2 (Analisi su ogni dataset), individua il gruppo di feature che, oltre ad essere significativamente discriminante (CS e MP) non sia statisticamente dipendente dalla lingua, dal genere e dal dataset di provenienza.

Individuati i 2 set di feature maggiormente robusti per ogni esperimento, si è applicato l'algoritmo di classificazione utilizzando i set di feature trovati per verificare se le prestazioni dei classificatori migliorassero rispetto alla semplice applicazione della feature selection.

5.2.1 Analisi sui 5 dataset uniti

Considerando i 5 dataset uniti, è stato utilizzato il test di Kruskal-Wallis per confrontare le distribuzioni delle feature nei soggetti CS e con MP in termini di lingua e dipendenza dal dataset, mentre è stato usato Mann Whitney per valutare la dipendenza dal genere.

Ottenute le feature che discriminano correttamente la classe, è stata testata la robustezza di queste feature considerando il genere, la dipendenza dal dataset e la lingua, utilizzando, anche in questo caso, il test di Kruskal-Wallis. Per queste prove, il dataset unito è stato diviso in soggetti CS e soggetti con MP per evitare che la dipendenza dalla classe influenzasse i risultati.

Dalla prova si individuano 2 set di feature:

1. Gruppo 1: feature con valore di $p < 0,05$ per la classe e valore di $p > 0,05$ per lingua, genere e dataset;
2. Gruppo 2: feature con valore di $p < 0,05$ per la classe e per il genere (sia soggetti MP che CS) e valore di $p > 0,05$ per lingua e dataset.

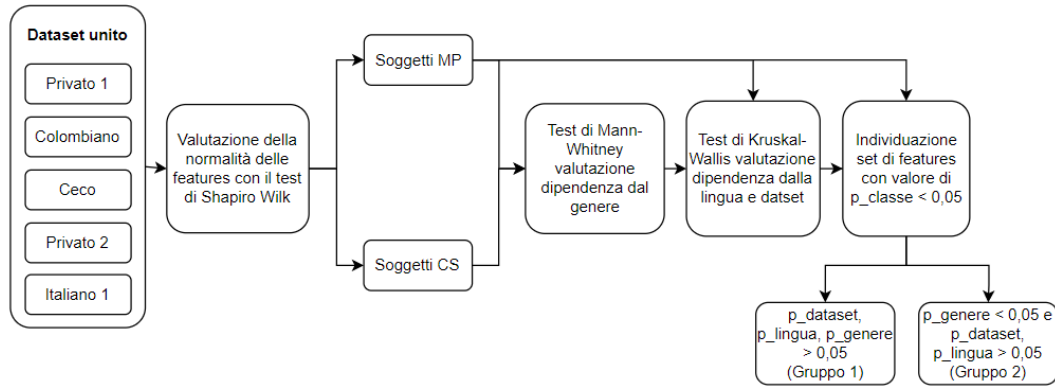


Figura 5.2: Diagramma di flusso analisi statistica sul dataset unito dataset

5.2.2 Analisi su ogni dataset

Il test di Shapiro-Wilk è stato dapprima applicato per valutare la normalità delle feature. Successivamente sono stati applicati i test di U-Mann Whitney e di Kruskal Wallis per valutare la rilevanza dei parametri. E' stato utilizzato il test U di Mann-Whitney per confrontare le distribuzioni delle feature nei soggetti CS e con MP, Figura 5.3.

Dalla prova si individua il set di feature, chiamato Gruppo 3, formato dalle feature che hanno valore di $p_classe < 0,05$ in almeno 3 dataset.

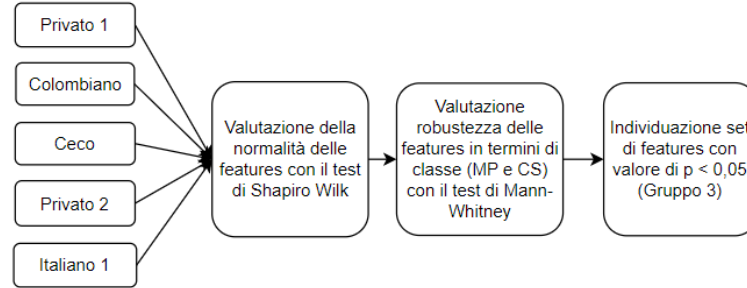


Figura 5.3: Diagramma di flusso analisi statistica su ogni dataset

5.3 Feature robuste e Machine learning

Gli ultimi esperimenti condotti, considerano i 3 gruppi di feature robuste individuate dall'analisi statistica e utilizzano l'algoritmo di classificazione per la discriminazione pazienti MP e CS.

In totale, come mostrato in Figura 5.4, gli esperimenti effettuati sono 4:

1. algoritmo di classificazione considerando il set di feature del Gruppo 1;
2. algoritmo di classificazione considerando il set di feature del Gruppo 1 + Gruppo 2;
3. algoritmo di classificazione considerando il set di feature del Gruppo 1 + Gruppo 2 + Gruppo 3;
4. algoritmo di classificazione in cui è stato utilizzato Boruta come algoritmo di feature selection, applicato al set di feature del Gruppo 1 + Gruppo 2 + Gruppo 3;

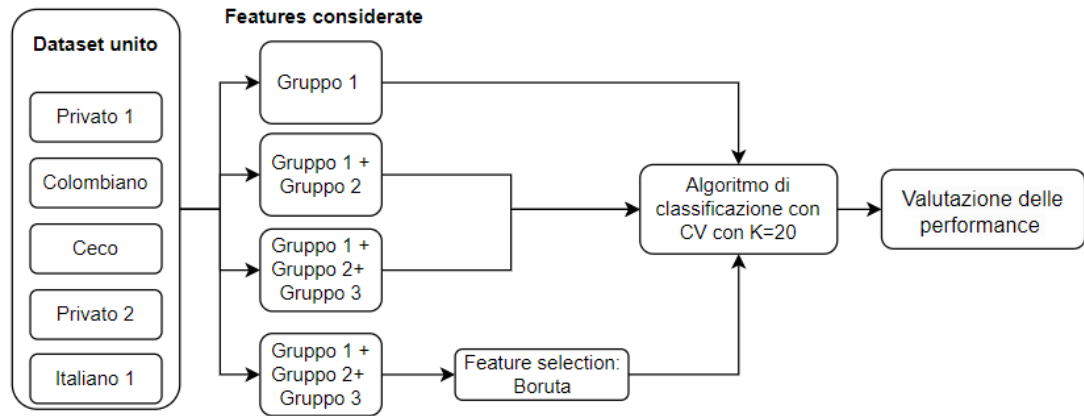


Figura 5.4: Diagramma di flusso ML con sei di feature dell'analisi statistica

Capitolo 6

Risultati

Nel presente capitolo sono fornite le valutazioni dei metodi descritti in termini di accuratezza per ciascuno dei 10 classificatori utilizzati. Le scelte delle prove migliori ricadono sulla prova che ha media dell'accuratezza sul validation più alta e deviazione standard minore.

L'analisi di accuratezza è stata eseguita per i due compiti vocali, fonazione sostenuta della vocale e due frasi, fornendo informazioni sul compito vocale migliore da utilizzare per le applicazioni cliniche, tenendo in conto anche, come parametro di scelta, il tempo di esecuzione.

6.1 Valutazione del metodo di registrazione

In questa sezione si valuta quanto il metodo di registrazione influisca sulle prestazioni del validation. Si considerano separatamente i compiti del vocalizzo e delle frasi.

In Tabella 6.1 si riportano i risultati del compito del vocalizzo sul dataset Italiano 1 e sull'unione di Italiano 1 e Privato 1. In Tabella 6.2 si riportano le accuratezze medie del compito delle frasi del dataset Italiano 1. Infine, in Tabella 6.3 si riportano le accuratezze medie del compito delle frasi del dataset Italiano 1+ Privato 1.

6.2 Scelta della vocale

In Tabella 6.4 si riportano i risultati, per i dataset Italiano 1 e Colombiano, delle 16 prove effettuate per la scelta della vocale (/a/ oppure /e/). In Figura 6.1 si

riporta la lista di feature selezionate per la vocale /a/ e in Figura 6.2 la lista di feature selezionate per la vocale /e/.

6.3 Scelta delle frasi

Si riportano di seguito i risultati ottenuti con l'applicazione dell'algoritmo di classificazione al compito delle frasi.

Italiano 1 + Privato 1

In Tabella 6.5 sono riportati i valori di accuratezza su tutte le frasi. Le feature selezionate sono rappresentate in Figura 6.3

Colombiano

In Tabella 6.6 sono riportati i valori di accuratezza su tutte le frasi. Le feature selezionate sono rappresentate in Figura 6.4.

Italiano 2

In Tabella 6.7 sono riportati i valori di accuratezza su tutte le frasi. Le feature selezionate sono rappresentate in Figura 6.5.

CROSS CORR_ZSCORE	['CPP_mean_a','CPP_median_a','FORM3_skewness_a','FORM_BAND1_std_a','FORM_BAND2_std_a','GNE_median_a','HNR_kurtosis_a','HNR_mean_a',' 'HNR_skewness_a','HNR_std_a','INT_AVG_a','MFCC11_mean_a','MFCC11_median_a','MFCC11_skewness_a','MFCC11_Δ_mean_a','MFCC11_mean_a',' 'MFCC11_median_a','MFCC11_skewness_a','MFCC11_ΔΔ_skewness_a','MFCC2_mean_a','MFCC2_median_a','MFCC3_mean_a','MFCC3_median_a','MFCC3_std_a',' 'MFCC3_Δ_std_a','MFCC4_kurtosis_a','MFCC4_skewness_a','MFCC4_Δ_mean_a','MFCC4_ΔΔ_skewness_a','MFCC4_Δ_std_a','MFCC4_ΔΔ_std_a','MFCC6_skewness_a',' 'MFCC7_mean_a','MFCC7_median_a','MFCC7_Δ_std_a','MFCC7_ΔΔ_mean_a','MFCC7_ΔΔ_std_a','MFCC8_mean_a','MFCC8_median_a','MFCC8_Δ_mean_a',' 'MFCC8_Δ_std_a','MFCC9_Δ_mean_a','PLP0_mean_a','PLP0_skewness_a','PLP0_Δ_mean_a','PLP0_ΔΔ_skewness_a','PLP0_ΔΔ_mean_a','PLP10_mean_a','PLP10_std_a',' 'PLP10_Δ_std_a','PLP10_ΔΔ_std_a','PLP11_std_a','PLP11_Δ_std_a','PLP11_ΔΔ_std_a','PLP12_std_a','PLP12_Δ_std_a','PLP12_ΔΔ_std_a',' 'PLP1_kurtosis_a','PLP1_mean_a','PLP1_Δ_std_a','PLP1_ΔΔ_std_a','PLP2_mean_a','PLP2_skewness_a','PLP2_std_a','PLP2_Δ_std_a','PLP2_ΔΔ_mean_a',' 'PLP2_ΔΔ_skewness_a','PLP2_ΔΔ_std_a','PLP3_skewness_a','PLP3_std_a','PLP3_Δ_std_a','PLP3_ΔΔ_skewness_a','PLP3_ΔΔ_std_a','PLP4_std_a','PLP4_Δ_std_a',' 'PLP4_ΔΔ_mean_a','PLP4_ΔΔ_std_a','PLP5_std_a','PLP5_Δ_std_a','PLP5_ΔΔ_std_a','PLP6_std_a','PLP6_Δ_std_a','PLP6_ΔΔ_std_a','PLP7_mean_a','PLP7_std_a',' 'PLP7_Δ_std_a','PLP7_ΔΔ_std_a','PLP8_std_a','PLP8_Δ_mean_a','PLP8_ΔΔ_std_a','PLP8_ΔΔ_mean_a','PLP8_ΔΔ_std_a','PLP9_std_a','PLP9_Δ_std_a','PLP9_ΔΔ_std_a',' 'SHIM_APQ11_a','SHIM_APO5_a','SHIM_DB_a','SHIM_a','SPEC_CENT_GRAVITY_mean_a','SPEC_CENT_GRAVITY_median_a','SPEC_CENT_GRAVITY_std_a',' 'SPEC_FLUX_mean_a','SPEC_FLUX_median_a','SPEC_FLUX_std_a','SPEC_KURTOSIS_mean_a','SPEC_KURTOSIS_median_a','SPEC_KURTOSIS_std_a',' 'SPEC_SKEWNESS_mean_a','SPEC_SKEWNESS_median_a','SPEC_STD_kurtosis_a','SPEC_STD_skewness_a','STE_mean_a','STE_std_a']
CROSS CORR_MINMAX	['CPP_mean_a','CPP_median_a','FORM_BAND1_std_a','FORM_BAND2_std_a','GNE_median_a','HNR_kurtosis_a','HNR_skewness_a','INT_AVG_a',' 'MFCC10_skewness_a','MFCC11_Δ_mean_a','MFCC11_ΔΔ_skewness_a','MFCC1_mean_a','MFCC11_median_a','MFCC1_skewness_a','MFCC11_Δ_mean_a',' 'MFCC1_ΔΔ_mean_a','MFCC1_ΔΔ_skewness_a','MFCC4_Δ_mean_a','MFCC4_ΔΔ_mean_a','MFCC6_skewness_a','MFCC7_mean_a','MFCC7_median_a',' 'MFCC8_Δ_mean_a','MFCC9_median_a','PLP0_mean_a','PLP0_skewness_a','PLP0_Δ_mean_a','PLP0_ΔΔ_mean_a','PLP0_ΔΔ_mean_a','PLP10_mean_a',' 'PLP10_std_a','PLP10_Δ_std_a','PLP10_ΔΔ_std_a','PLP12_Δ_mean_a','PLP1_kurtosis_a','PLP1_mean_a','PLP1_Δ_kurtosis_a','PLP1_Δ_std_a','PLP2_skewness_a',' 'PLP3_skewness_a','PLP6_std_a','PLP6_Δ_std_a','PLP6_ΔΔ_skewness_a','PLP6_ΔΔ_std_a','PLP7_mean_a','SHIM_APQ11_a','SHIM_APO3_a','SHIM_APO5_a',' 'SHIM_DB_a','SHIM_a','SPEC_CENT_GRAVITY_mean_a','SPEC_CENT_GRAVITY_std_a','SPEC_FLUX_mean_a','SPEC_FLUX_median_a','SPEC_FLUX_std_a',' 'SPEC_KURTOSIS_mean_a','SPEC_KURTOSIS_median_a','SPEC_KURTOSIS_std_a','SPEC_SKEWNESS_mean_a','SPEC_SKEWNESS_median_a','SPEC_SKEWNESS_std_a']
BORUTA_ZSCORE	['SHIM_DB_a','SHIM_APO5_a','SHIM_APQ11_a','INT_AVG_a','STE_mean_a','SPEC_CENT_GRAVITY_mean_a','PLP0_mean_a','STE_median_a',' 'SPEC_CENT_GRAVITY_median_a','SPEC_FLUX_median_a','PLP0_median_a','PLP0_ΔΔ_mean_a','STE_std_a','SPEC_SLOPE_std_a','PLP2_std_a','PLP4_std_a',' 'PLP6_std_a','PLP2_Δ_std_a','PLP4_Δ_std_a','PLP7_Δ_std_a','PLP2_ΔΔ_std_a','PLP4_ΔΔ_std_a','PLP7_ΔΔ_std_a','PLP0_skewness_a','PLP0_ΔΔ_skewness_a',' 'PLP0_ΔΔ_skewness_a']
BORUTA_MINMAX	['SHIM_DB_a','SHIM_APO3_a','SHIM_APQ11_a','SPEC_FLUX_mean_a','PLP0_mean_a','STE_median_a','GNE_median_a','SPEC_FLUX_median_a','PLP0_median_a',' 'PLP2_ΔΔ_mean_a','STE_std_a','PLP2_std_a','PLP3_std_a','PLP10_std_a','PLP1_Δ_std_a','PLP2_Δ_std_a','PLP3_Δ_std_a','PLP1_ΔΔ_std_a','PLP2_ΔΔ_std_a',' 'PLP6_ΔΔ_std_a','PLP0_ΔΔ_std_a','PLP0_ΔΔ_skewness_a']

CROSSCORR_ZSCORE	['CPP_skewness_a','F0_std_a','GNE_mean_a','GNE_median_a','GNE_std_a','MFCC10_std_a','MFCC10_Δ_std_a','MFCC10_ΔΔ_std_a','MFCC11_Δ_std_a','MFCC11_ΔΔ_std_a','MFCC12_Δ_std_a','MFCC12_ΔΔ_std_a','MFCC13_Δ_std_a','MFCC13_ΔΔ_std_a','MFCC14_Δ_std_a','MFCC14_ΔΔ_std_a','MFCC2_Δ_std_a','MFCC2_ΔΔ_std_a','MFCC3_Δ_std_a','MFCC3_ΔΔ_std_a','MFCC4_Δ_std_a','MFCC4_ΔΔ_std_a','MFCC5_Δ_std_a','MFCC5_ΔΔ_std_a','MFCC6_Δ_std_a','MFCC6_ΔΔ_std_a','MFCC7_mean_a','MFCC7_Δ_std_a','MFCC7_ΔΔ_std_a','MFCC8_Δ_mean_a','MFCC8_Δ_std_a','MFCC8_ΔΔ_std_a','MFCC9_Δ_std_a','MFCC9_ΔΔ_std_a','PLP10_mean_a','PLP10_median_a','PLP11_Δ_std_a','PLP11_ΔΔ_std_a','PLP12_Δ_std_a','PLP12_ΔΔ_std_a','PLP1_std_a','PLP1_Δ_std_a','PLP1_ΔΔ_std_a','PLP2_ΔΔ_std_a','PLP4_std_a','PLP4_Δ_std_a','PLP4_ΔΔ_std_a','PLP5_ΔΔ_std_a','PLP6_ΔΔ_std_a','PLP7_ΔΔ_std_a','PLP8_std_a','PLP8_Δ_std_a','PLP8_ΔΔ_std_a','PLP9_Δ_std_a','PLP9_ΔΔ_std_a','SHIM_APQ11_a','SPEC_STD_std_a']
CROSSCORR_MINMAX	['CPP_mean_a','CPP_median_a','F0_std_a','FORM3_mean_a','FORM_BAND3_skewness_a','GNE_mean_a','GNE_median_a','GNE_std_a','HNR_mean_a','HNR_std_a','JIT_DDP_a','JIT_L_a','JIT_PPQ5_a','JIT_RAP_a','MFCC10_std_a','MFCC10_Δ_std_a','MFCC10_ΔΔ_std_a','MFCC11_Δ_std_a','MFCC11_ΔΔ_std_a','MFCC12_Δ_std_a','MFCC12_ΔΔ_std_a','MFCC13_Δ_std_a','MFCC13_ΔΔ_std_a','MFCC1_Δ_std_a','MFCC1_ΔΔ_std_a','MFCC2_Δ_std_a','MFCC2_ΔΔ_std_a','MFCC3_Δ_std_a','MFCC3_ΔΔ_std_a','MFCC4_Δ_std_a','MFCC4_ΔΔ_std_a','MFCC5_Δ_std_a','MFCC5_ΔΔ_std_a','MFCC6_Δ_std_a','MFCC6_ΔΔ_std_a','MFCC7_Δ_mean_a','MFCC8_Δ_mean_a','PLP0_Δ_std_a','PLP0_ΔΔ_std_a','PLP10_mean_a','PLP10_median_a','PLP11_Δ_std_a','PLP11_ΔΔ_std_a','PLP12_Δ_std_a','PLP1_std_a','PLP1_Δ_std_a','PLP1_ΔΔ_std_a','PLP2_Δ_std_a','PLP2_ΔΔ_mean_a','PLP2_ΔΔ_skewness_a','PLP2_ΔΔ_std_a','PLP4_std_a','PLP4_Δ_std_a','PLP4_ΔΔ_std_a','PLP5_Δ_std_a','PLP5_ΔΔ_std_a','PLP6_Δ_std_a','PLP6_ΔΔ_std_a','PLP8_std_a','PLP8_Δ_mean_a','PLP8_Δ_std_a','PLP8_ΔΔ_std_a','PLP9_mean_a','PLP9_skewness_a','PLP9_Δ_std_a','PLP9_ΔΔ_std_a','SHIM_APQ11_a','SHIM_a','SPEC_FLUX_skewness_a','SPEC_KURTOSIS_skewness_a','SPEC_ROLLOFF_skewness_a','SPEC_SLOPE_kurtosis_a','SPEC_SLOPE_std_a']
BORUTA_ZSCORE	['FORM1_std_a','PLP2_ΔΔ_std_a','PLP9_ΔΔ_std_a','PLP5_kurtosis_a','HNR_skewness_a']
BORUTA_MINMAX	['JIT_PPQ5_a','SHIM_APQ11_a','PLP9_mean_a','PLP10_mean_a','PLP2_ΔΔ_mean_a','FORM1_std_a','MFCC10_ΔΔ_std_a','MFCC11_ΔΔ_std_a','PLP9_ΔΔ_std_a','PLP2_Δ_kurtosis_a','HNR_skewness_a']

CROSSCORR_ZSCORE	['CPP_mean_e','F0_std_e','FORM1_kurtosis_e','FORM2_skewness_e','GNE_mean_e','HNR_kurtosis_e','HNR_mean_e','HNR_skewness_e','HNR_std_e','INT_AVG_e','MFCC10_kurtosis_e','MFCC10_skewness_e','MFCC10_Δ_std_e','MFCC10_ΔΔ_std_e','MFCC12_kurtosis_e','MFCC13_kurtosis_e','MFCC1_mediana_e','MFCC1_Δ_std_e','MFCC1_ΔΔ_kurtosis_e','MFCC1_ΔΔ_std_e','MFCC3_ΔΔ_std_e','MFCC5_mean_e','MFCC5_mediana_e','MFCC5_Δ_mean_e','MFCC5_ΔΔ_skewness_e','MFCC5_Δ_std_e','MFCC5_ΔΔ_std_e','MFCC6_Δ_std_e','MFCC6_ΔΔ_std_e','PLP0_mean_e','PLP0_skewness_e','PLP0_Δ_mean_e','PLP0_ΔΔ_skewness_e','PLP0_ΔΔ_mean_e','PLP0_ΔΔ_mediana_e','PLP0_ΔΔ_skewness_e','PLP10_kurtosis_e','PLP10_Δ_kurtosis_e','PLP10_Δ_mean_e','PLP12_std_e','PLP12_Δ_std_e','PLP12_ΔΔ_std_e','PLP1_kurtosis_e','PLP1_mean_e','PLP1_skewness_e','PLP1_Δ_skewness_e','PLP2_kurtosis_e','PLP2_skewness_e','PLP4_mean_e','PLP4_skewness_e','PLP5_mean_e','PLP5_skewness_e','PLP5_std_e','PLP5_Δ_skewness_e','PLP5_ΔΔ_std_e','PLP5_ΔΔ_std_e','PLP6_std_e','PLP7_Δ_mediana_e','PLP8_kurtosis_e','PLP8_Δ_mediana_e','PLP9_kurtosis_e','PLP9_skewness_e','SHIM_APQ11_e','SHIM_APQ5_e','SHIM_e','SPEC_CENT_GRAVITY_mediana_e','SPEC_FLUX_mean_e','SPEC_FLUX_mediana_e','SPEC_FLUX_std_e','SPEC_ROLLOFF_mean_e','SPEC_SLOPE_mediana_e','SPEC_STD_kurtosis_e','SPEC_STD_mean_e','SPEC_STD_skewness_e','SPEC_STD_std_e','STE_mediana_e','STE_std_e']
CROSSCORR_MINMAX	['CPP_mean_e','F0_std_e','FORM1_kurtosis_e','FORM2_skewness_e','GNE_mean_e','HNR_mean_e','HNR_skewness_e','HNR_std_e','INT_AVG_e','MFCC13_kurtosis_e','MFCC1_mediana_e','MFCC1_Δ_mean_e','MFCC1_ΔΔ_skewness_e','MFCC1_ΔΔ_std_e','MFCC2_kurtosis_e','MFCC3_Δ_std_e','MFCC3_ΔΔ_std_e','MFCC4_mean_e','MFCC4_mediana_e','MFCC5_mean_e','MFCC5_mediana_e','MFCC5_Δ_mean_e','MFCC5_ΔΔ_skewness_e','MFCC6_Δ_mean_e','MFCC6_ΔΔ_std_e','PLP0_mean_e','PLP0_skewness_e','PLP0_Δ_mediana_e','PLP0_ΔΔ_skewness_e','PLP0_ΔΔ_mediana_e','PLP10_mean_e','PLP10_std_e','PLP11_std_e','PLP11_Δ_std_e','PLP11_ΔΔ_std_e','PLP1_mean_e','PLP1_skewness_e','PLP1_Δ_skewness_e','PLP2_skewness_e','PLP4_mean_e','PLP5_mean_e','PLP5_skewness_e','PLP5_std_e','PLP5_Δ_skewness_e','PLP5_ΔΔ_std_e','PLP7_mean_e','PLP7_std_e','PLP9_skewness_e','SHIM_APQ11_e','SHIM_APQ3_e','SHIM_APQ5_e','SHIM_DB_e','SHIM_e','SPEC_CENT_GRAVITY_mediana_e','SPEC_FLUX_mean_e','SPEC_FLUX_mediana_e','SPEC_SKEWNESS_kurtosis_e','SPEC_SLOPE_mediana_e','SPEC_STD_mean_e','SPEC_STD_mediana_e','SPEC_STD_std_e','STE_mediana_e','STE_std_e']
BORUTA_ZSCORE	['INT_AVG_e','STE_mean_e','MFCC2_Δ_mean_e','MFCC5_Δ_mean_e','PLP0_mean_e','PLP5_mean_e','STE_mediana_e','SPEC_FLUX_mediana_e','STE_std_e','SPEC_STD_std_e','PLP6_Δ_std_e','PLP5_ΔΔ_std_e','SPEC_SKEWNESS_kurtosis_e','MFCC13_kurtosis_e','PLP0_skewness_e','PLP1_skewness_e','PLP0_ΔΔ_skewness_e']
BORUTA_MINMAX	['MFCC5_mean_e','PLP7_mean_e','STE_mediana_e','SPEC_FLUX_mediana_e','FORM_BAND3_mediana_e','MFCC5_mediana_e','MFCC13_kurtosis_e','PLP0_ΔΔ_skewness_e']

(a) Italiano 1

CROSSCORR_ZSCORE	['CPP_mean_e','CPP_std_e','GNE_mean_e','GNE_median_e','GNE_std_e','MFCC10_Δ_std_e','MFCC10_ΔΔ_std_e','MFCC11_Δ_std_e','MFCC11_ΔΔ_std_e','MFCC12_Δ_std_e','MFCC12_ΔΔ_std_e','MFCC13_std_e','MFCC13_Δ_std_e','MFCC13_ΔΔ_std_e','MFCC1_Δ_std_e','MFCC1_ΔΔ_std_e','MFCC2_mean_e','MFCC2_median_e','MFCC2_ΔΔ_std_e','MFCC3_Δ_std_e','MFCC3_ΔΔ_std_e','MFCC4_ΔΔ_std_e','MFCC5_std_e','MFCC5_Δ_std_e','MFCC5_ΔΔ_std_e','MFCC6_Δ_std_e','MFCC6_ΔΔ_std_e','MFCC7_Δ_kurtosis_e','MFCC7_Δ_std_e','MFCC7_ΔΔ_std_e','MFCC8_Δ_kurtosis_e','MFCC8_Δ_std_e','MFCC8_ΔΔ_std_e','MFCC9_std_e','MFCC9_Δ_std_e','MFCC9_ΔΔ_std_e','PLP2_median_e','SHIM_APQ11_e','SHIM_APQ5_e','SHIM_DB_e','SHIM_e','SPEC_FLUX_median_e','SPEC_KURTOSIS_std_e']
CROSSCORR_MINMAX	['GNE_mean_e','GNE_median_e','GNE_std_e','MFCC10_std_e','MFCC10_Δ_std_e','MFCC10_ΔΔ_std_e','MFCC11_Δ_std_e','MFCC11_ΔΔ_std_e','MFCC12_Δ_std_e','MFCC12_ΔΔ_std_e','MFCC13_Δ_std_e','MFCC13_ΔΔ_std_e','MFCC2_ΔΔ_std_e','MFCC3_Δ_std_e','MFCC3_ΔΔ_std_e','MFCC4_ΔΔ_std_e','MFCC5_std_e','MFCC5_Δ_std_e','MFCC5_ΔΔ_std_e','MFCC6_ΔΔ_std_e','MFCC7_std_e','MFCC7_Δ_std_e','MFCC7_ΔΔ_std_e','MFCC9_std_e','MFCC9_Δ_std_e','MFCC9_ΔΔ_std_e','PLP0_ΔΔ_skewness_e','PLP10_ΔΔ_std_e','PLP11_Δ_std_e','PLP11_ΔΔ_std_e','PLP1_Δ_std_e','PLP1_ΔΔ_std_e','PLP2_kurtosis_e','PLP2_median_e','PLP2_ΔΔ_skewness_e','PLP2_ΔΔ_std_e','PLP4_ΔΔ_skewness_e','PLP5_ΔΔ_std_e','SHIM_APQ11_e','SHIM_DB_e','SHIM_e']
BORUTA_ZSCORE	['MFCC4_ΔΔ_mean_e','SPEC_FLUX_median_e','SPEC_SLOPE_kurtosis_e']
BORUTA_MINMAX	['PLP8_ΔΔ_mean_e','MFCC13_Δ_std_e','PLP2_ΔΔ_std_e','CPP_skewness_e']

(b) Colombian

Figura 6.2: Lista di feature Italiano 1 e Colombiano, vocalizzo /e/

Prove	Italiano 1	Italiano 1 + Privato 1
Cross-corr + z-score	95 ±2.4	71 ± 4.7
Cross-corr + max-min	90 ± 5.4	72 ± 5.1
Boruta + z-score	92 ± 4.6	79 ±5.3
Boruta + max-min	93 ± 5.0	78 ± 5.2

Tabella 6.1: Accuratezza (%) media sul validation (k-fold al 70% del trainset) con 10 Classificatori e vocalizzo /a/, per modalità di registrazione

Italiano 1	Cross-corr + z-score	Cross-corr + max-min	Boruta + z-score	Boruta + max-min
Frase 0	94 ± 2.1	94 ± 5.5	89 ± 11.1	85 ± 14.7
Frase 1	93 ± 4.4	90 ± 5.0	96 ± 3.5	90 ± 7.8
Frase 2	92 ± 3.3	93 ± 2.1	97 ± 2.5	96 ± 3.9
Frase 3	94 ± 2.4	95 ± 3.8	96 ± 2.9	93 ± 5.8
Frase 4	92 ± 4.1	90 ± 6.6	92 ± 8.5	92 ± 7.3
Frase 5	90 ± 5.7	89 ± 9.6	92 ± 2.2	87 ± 10.1
Frase 6	93 ± 7.7	88 ± 7.2	93 ± 5.9	88 ± 9.8
Frase 7	91 ± 5.4	92 ± 7.0	93 ± 3.9	92 ± 8.0
Frase 8	90 ± 8.6	94 ± 4.7	93 ± 4.0	92 ± 4.8
Frase 9	92 ± 6.3	89 ± 5.8	98 ±2.0	93 ± 6.0

Tabella 6.2: Accuratezza (%) media sul validation (k-fold al 70% del trainset) con 10 Classificatori e frasi Italiano 1, per modalità di registrazione

Italiano 1+Privato 1	Cross-corr+z-score	Cross-corr+max-min	Boruta+z-score	Boruta+max-min
Frase 0	78 ± 4.7	74 ± 7.5	81 ± 3.0	78 ± 7.2
Frase 1	66 ± 5.3	67 ± 5.3	78 ± 4.1	81 ± 6.3
Frase 2	76 ± 4.8	76 ± 6.3	83 ± 3.0	81 ± 8.3
Frase 3	80 ± 5.7	73 ± 7.4	86 ±2.4	84 ± 7.3
Frase 4	83 ± 5.1	80 ± 5.7	81 ± 3.9	77 ± 5.3
Frase 5	69 ± 4.0	68 ± 6.5	76 ± 3.8	82 ± 6.2
Frase 6	79 ± 5.8	70 ± 7.5	80 ± 4.1	75 ± 3.4
Frase 7	72 ± 4.4	70 ± 4.8	81 ± 4.6	77 ± 6.1
Frase 8	78 ± 4.6	68 ± 6.6	84 ± 2.6	72 ± 6.3
Frase 9	78 ± 4.1	74 ± 8.5	86 ± 3.7	75 ± 6.3

Tabella 6.3: Accuratezza (%) media sul validation (k-fold al 70% del trainset) con 10 Classificatori e frasi Italiano 1+Privato 1, per modalità di registrazione

Prove	Italiano 1 /a/	Colombiano /a/	Italiano 1 /e/	Colombiano /e/
Cross-corr + z-score	95 ±2.4	65 ± 5.2	94 ± 3.1	66 ± 5.8
Cross-corr + max-min	91 ± 5.4	71 ± 5.0	87 ± 6.8	63 ± 4.6
Boruta + z-score	92 ± 4.6	72 ± 2.4	89 ± 9.3	67 ± 3.9
Boruta + max-min	93 ± 5.0	79 ±3.2	90 ± 6.2	70 ± 4.3

Tabella 6.4: Accuratezza (%) media sul validation (k-fold al 70% del trainset) con 10 Classificatori e vocalizzo /a/ e vocalizzo /e/

Italiano 1+Privato 1	Cross-corr+z-score	Cross-corr+max-min	Boruta+z-score	Boruta+max-min
Frase 0	78 ± 4.7	74 ± 7.5	81 ± 3.0	78 ± 7.2
Frase 1	66 ± 5.3	67 ± 5.3	78 ± 4.1	81 ± 6.3
Frase 2	76 ± 4.8	76 ± 6.3	83 ± 3.0	81 ± 8.3
Frase 3	80 ±5.7	73 ±7.4	86 ±2.4	84 ±7.3
Frase 4	83 ±5.1	80 ±5.7	81 ±3.9	77 ±5.3
Frase 5	69 ± 4.0	68 ± 6.5	76 ± 3.8	82 ± 6.2
Frase 6	79 ± 5.8	70 ± 7.5	80 ± 4.1	75 ± 3.4
Frase 7	72 ± 4.4	70 ± 4.8	81 ± 4.6	77 ± 6.1
Frase 8	78 ± 4.6	68 ± 6.6	84 ± 2.6	72 ± 6.3
Frase 9	78 ± 4.1	74 ± 8.5	86 ± 3.7	75 ± 6.3

Tabella 6.5: Accuratezza (%) media sul validation (k-fold al 70% del trainset), con 10 Classificatori, 10 frasi Italiano 1+Privato 1

Colombiano	Cross-corr+z-score	Cross-corr+max-min	Boruta+z-score	Boruta+max-min
Frase 0	72 ± 3.0	70±2.6	79 ± 4.2	77 ± 2.4
Frase 1	81 ±5.1	70 ±3.4	78 ±2.3	80 ±7.0
Frase 2	73 ± 6.1	66 ± 2.9	80 ± 3.6	74 ± 7.3
Frase 3	75 ± 4.4	65 ± 6.5	74 ± 3.8	65 ± 4.6
Frase 4	63 ± 2.2	69 ± 2.7	74 ± 2.9	68 ± 7.4
Frase 5	71 ±3.3	74 ±5.8	82 ±2.3	75 ±5.0

Tabella 6.6: Accuratezza (%) media sul validation (k-fold al 70% del trainset) su 10 Classificatori, 6 frasi Colombiano

Italiano 2	Cross-corr+z-score	Cross-corr+max-min	Boruta+z-score	Boruta+max-min
Frase 0	72 ± 5.0	80 ± 4.9	85 ± 3.2	82 ± 5.8
Frase 1	76 ±4.4	83 ±4.9	87 ±1.8	78 ±8.2
Frase 2	78 ±4.9	81 ±6.6	89 ±4.3	85 ±7.6

Tabella 6.7: Accuratezza (%) media sul validation (k-fold al 70% del trainset) su 10 Classificatori, 3 frasi Italiano 2

6.4 Protocollo

In questa sezione si riportano i valori di accuratezza sul validation e, solo sul test, i valori di specificità, sensibilità, AUC ed F1. Ad ogni compito vocale sono associate

una tabella (in cui sono riportate le prestazioni di ogni classificatore) e la rispettiva lista di feature selezionate.

6.4.1 Italiano 1 + Privato 1

Vocalizzo

In Tabella 6.8, si riportano i risultati del vocalizzo /a/ per i 10 classificatori. In Figura 6.6 (a) si riporta la lista di feature selezionate.

Frase migliore

In Tabella 6.9, si riportano i risultati della frase migliore per i 10 classificatori. In Figura 6.6 (b) si riporta la lista di feature selezionate.

Vocalizzo unito alla frase migliore

In Tabella 6.10, si riportano i risultati dell'unione del vocalizzo /a/ e delle frase migliore per i 10 classificatori. In Figura 6.7 si riporta la lista di feature selezionate.

Vocalizzo unito alle due frasi migliori

In Tabella 6.11, si riportano i risultati dell'unione del vocalizzo /a/ e delle due frasi migliori per i 10 classificatori. In Figura 6.8 si riporta la lista di feature selezionate.

6.4.2 Colombiano

Vocalizzo

In Tabella 6.12, si riportano i risultati del vocalizzo /a/ per i 10 classificatori. In Figura 6.9 si riporta la lista di feature selezionate.

Frase migliore

In Tabella 6.13, si riportano i risultati della frase migliore per i 10 classificatori. In Figura 6.10 si riporta la lista di feature selezionate.

Vocalizzo unito alla frase migliore

In Tabella 6.14, si riportano i risultati dell'unione del vocalizzo /a/ e delle frase migliore per i 10 classificatori. In Figura 6.11 si riporta la lista di feature selezionate.

Vocalizzo unito alle due frasi migliori

In Tabella 6.15, si riportano i risultati dell'unione del vocalizzo /a/ e delle due frasi migliori per i 10 classificatori. In Figura 6.12 si riporta la lista di feature selezionate.

6.4.3 Italiano 2

Vocalizzo

In Tabella 6.16, si riportano i risultati del vocalizzo /e/ per i 10 classificatori. In Figura 6.13 si riporta la lista di feature selezionate.

Frase migliore

In Tabella 6.17, si riportano i risultati della frase migliore per i 10 classificatori. In Figura 6.14 si riporta la lista di feature selezionate.

Vocalizzo unito alla frase migliore

In Tabella 6.18, si riportano i risultati dell'unione del vocalizzo /e/ e delle frase migliore per i 10 classificatori. In Figura 6.15 si riporta la lista di feature selezionate.

Vocalizzo unito alle due frasi migliori

In Tabella 6.19, si riportano i risultati dell'unione del vocalizzo /e/ e delle due frasi migliori per i 10 classificatori. In Figura 6.16 si riporta la lista di feature selezionate.

CROSSCORR_ZSCORE	['CPP_mean_frame' 'CPP_std_frame' 'DFA_COEFF_frame' 'FORM1_mean_frame' 'FORM1_median_frame' 'FORM1_skewness_frame' 'HNR_mean_frame' 'MFCC13_skewness_frame' 'MFCC1_mean_frame' 'MFCC1_median_frame' 'MFCC1_skewness_frame' 'MFCC1_Δ_mean_frame' 'MFCC1_Δ_median_frame' 'MFCC2_mean_frame' 'MFCC2_median_frame' 'MFCC2_Δ_mean_frame' 'MFCC2_Δ_median_frame' 'MFCC3_mean_frame' 'MFCC3_median_frame' 'MFCC3_Δ_kurtosis_frame' 'MFCC4_mean_frame' 'MFCC4_median_frame' 'MFCC5_mean_frame' 'MFCC6_mean_frame' 'MFCC6_median_frame' 'MFCC7_mean_frame' 'MFCC7_median_frame' 'MFCC8_ΔΔ_mean_frame' 'PLP0_kurtosis_frame' 'PLP0_Δ_kurtosis_frame' 'PLP0_Δ_std_frame' 'PLP0_ΔΔ_kurtosis_frame' 'PLP0_ΔΔ_std_frame' 'PLP1_skewness_frame' 'PLP1_Δ_mean_frame' 'PLP1_Δ_skewness_frame' 'PLP1_ΔΔ_mean_frame' 'PLP2_std_frame' 'PLP2_Δ_mean_frame' 'PLP2_Δ_std_frame' 'PLP2_ΔΔ_std_frame' 'PLP3_Δ_std_frame' 'PLP3_ΔΔ_std_frame' 'PLP5_std_frame' 'PLP5_Δ_std_frame' 'PLP5_ΔΔ_std_frame' 'PLP6_std_frame' 'PLP6_Δ_std_frame' 'PLP6_ΔΔ_std_frame' 'PLP7_Δ_std_frame' 'PLP7_ΔΔ_std_frame' 'PLP9_std_frame' 'PLP9_Δ_std_frame' 'PLP9_ΔΔ_std_frame' 'SHIM_APO3_frame' 'SHIM_APO5_frame' 'SHIM_DB_frame' 'SPEC_CENT_GRAVITY_mean_frame' 'SPEC_CENT_GRAVITY_median_frame' 'SPEC_CENT_GRAVITY_std_frame' 'SPEC_KURTOSIS_mean_frame' 'SPEC_KURTOSIS_median_frame' 'SPEC_KURTOSIS_std_frame' 'SPEC_ROLLOFF_mean_frame' 'SPEC_ROLLOFF_median_frame' 'SPEC_SKEWNESS_mean_frame' 'SPEC_SKEWNESS_median_frame' 'SPEC_SKEWNESS_std_frame']
CROSSCORR_MINMAX	['DFA_COEFF_frame' 'FORM1_kurtosis_frame' 'FORM1_skewness_frame' 'HNR_mean_frame' 'HNR_median_frame' 'HNR_std_frame' 'MFCC12_std_frame' 'MFCC1_median_frame' 'MFCC1_Δ_mean_frame' 'MFCC2_mean_frame' 'MFCC2_median_frame' 'MFCC2_std_frame' 'MFCC2_Δ_std_frame' 'MFCC3_Δ_std_frame' 'MFCC3_ΔΔ_std_frame' 'MFCC4_mean_frame' 'MFCC4_median_frame' 'MFCC6_mean_frame' 'MFCC7_mean_frame' 'MFCC7_median_frame' 'PLP0_Δ_kurtosis_frame' 'PLP0_Δ_std_frame' 'PLP0_ΔΔ_kurtosis_frame' 'PLP0_ΔΔ_std_frame' 'PLP1_Δ_mean_frame' 'PLP1_Δ_skewness_frame' 'PLP2_std_frame' 'PLP2_Δ_mean_frame' 'PLP2_Δ_std_frame' 'PLP2_ΔΔ_std_frame' 'PLP3_mean_frame' 'PLP3_Δ_std_frame' 'PLP3_ΔΔ_std_frame' 'PLP4_Δ_std_frame' 'PLP4_ΔΔ_std_frame' 'PLP6_std_frame' 'PLP6_Δ_std_frame' 'SHIM_APO3_frame' 'SHIM_APO5_frame' 'SHIM_DB_frame' 'SPEC_CENT_GRAVITY_mean_frame' 'SPEC_CENT_GRAVITY_median_frame' 'SPEC_CENT_GRAVITY_std_frame' 'SPEC_KURTOSIS_mean_frame' 'SPEC_KURTOSIS_median_frame' 'SPEC_ROLLOFF_mean_frame' 'SPEC_SKEWNESS_mean_frame' 'SPEC_SKEWNESS_median_frame' 'SPEC_SKEWNESS_std_frame']
BORUTA_ZSCORE	['SHIM_APO5_frame', 'DFA_COEFF_frame', 'HNR_mean_frame', 'SPEC_CENT_GRAVITY_mean_frame', 'MFCC2_mean_frame', 'PLP0_mean_frame', 'PLP4_Δ_mean_frame', 'SPEC_CENT_GRAVITY_median_frame', 'MFCC2_median_frame', 'PLP1_Δ_mean_frame', 'f0_std_frame']
BORUTA_MINMAX	['SHIM_APO5_frame', 'DFA_COEFF_frame', 'HNR_mean_frame', 'MFCC2_mean_frame', 'MFCC4_mean_frame', 'PLP0_mean_frame', 'MFCC2_median_frame', 'PLP12_Δ_mean_frame', 'MFCC3_Δ_std_frame', 'MFCC3_ΔΔ_std_frame', 'PLP2_std_frame', 'PLP4_ΔΔ_std_frame', 'PLP2_Δ_skewness_frame']

(a) Frase 3

CROSSCORR_ZSCORE	['CPP_mean_frame' 'DFA_COEFF_frame' 'FORM1_mean_frame' 'FORM1_median_frame' 'FORM2_std_frame' 'FORM3_mean_frame' 'HNR_mean_frame' 'HNR_median_frame' 'INT_SD_frame' 'MFCC11_ΔΔ_mean_frame' 'MFCC12_kurtosis_frame' 'MFCC1_mean_frame' 'MFCC1_median_frame' 'MFCC1_ΔΔ_skewness_frame' 'MFCC2_mean_frame' 'MFCC2_median_frame' 'MFCC2_Δ_kurtosis_frame' 'MFCC2_ΔΔ_kurtosis_frame' 'MFCC3_mean_frame' 'MFCC3_median_frame' 'MFCC4_mean_frame' 'MFCC4_median_frame' 'MFCC5_mean_frame' 'MFCC5_median_frame' 'MFCC7_mean_frame' 'MFCC7_Δ_mean_frame' 'PLP0_Δ_kurtosis_frame' 'PLP0_Δ_std_frame' 'PLP0_ΔΔ_kurtosis_frame' 'PLP0_ΔΔ_std_frame' 'PLP10_Δ_std_frame' 'PLP10_ΔΔ_std_frame' 'PLP2_Δ_kurtosis_frame' 'PLP2_Δ_std_frame' 'PLP2_ΔΔ_kurtosis_frame' 'PLP3_std_frame' 'PLP3_Δ_std_frame' 'PLP3_ΔΔ_std_frame' 'PLP5_std_frame' 'PLP5_Δ_mean_frame' 'PLP5_Δ_std_frame' 'PLP5_ΔΔ_mean_frame' 'PLP6_Δ_mean_frame' 'PLP6_ΔΔ_mean_frame' 'PLP8_Δ_std_frame' 'PLP9_Δ_std_frame' 'PLP9_ΔΔ_std_frame' 'SHIM_APO3_frame' 'SHIM_APO5_frame' 'SHIM_DB_frame' 'SHIM_frame' 'SPEC_CENT_GRAVITY_mean_frame' 'SPEC_CENT_GRAVITY_median_frame' 'SPEC_CENT_GRAVITY_std_frame' 'SPEC_KURTOSIS_mean_frame' 'SPEC_KURTOSIS_median_frame' 'SPEC_KURTOSIS_std_frame' 'SPEC_ROLLOFF_mean_frame' 'SPEC_ROLLOFF_std_frame' 'SPEC_SKEWNESS_mean_frame' 'SPEC_SKEWNESS_median_frame' 'SPEC_SKEWNESS_std_frame']
CROSSCORR_MINMAX	['CPP_mean_frame' 'CPP_median_frame' 'DFA_COEFF_frame' 'FORM1_kurtosis_frame' 'FORM1_mean_frame' 'FORM1_median_frame' 'FORM_BAND3_skewness_frame' 'HNR_mean_frame' 'MFCC10_ΔΔ_kurtosis_frame' 'MFCC1_kurtosis_frame' 'MFCC1_mean_frame' 'MFCC1_median_frame' 'MFCC1_skewness_frame' 'MFCC1_Δ_kurtosis_frame' 'MFCC1_Δ_mean_frame' 'MFCC1_ΔΔ_kurtosis_frame' 'MFCC1_ΔΔ_skewness_frame' 'MFCC2_mean_frame' 'MFCC2_median_frame' 'MFCC2_Δ_kurtosis_frame' 'MFCC2_ΔΔ_kurtosis_frame' 'MFCC3_mean_frame' 'MFCC3_median_frame' 'MFCC3_skewness_frame' 'MFCC3_std_frame' 'MFCC4_mean_frame' 'MFCC4_median_frame' 'MFCC5_mean_frame' 'MFCC5_median_frame' 'MFCC8_kurtosis_frame' 'PLP0_ΔΔ_std_frame' 'PLP10_std_frame' 'PLP10_Δ_kurtosis_frame' 'PLP10_Δ_std_frame' 'PLP10_ΔΔ_std_frame' 'PLP12_Δ_std_frame' 'PLP1_kurtosis_frame' 'PLP1_Δ_kurtosis_frame' 'PLP1_Δ_skewness_frame' 'PLP1_ΔΔ_kurtosis_frame' 'PLP2_kurtosis_frame' 'PLP2_Δ_kurtosis_frame' 'PLP2_ΔΔ_kurtosis_frame' 'PLP3_std_frame' 'PLP3_Δ_skewness_frame' 'PLP3_Δ_std_frame' 'PLP3_ΔΔ_std_frame' 'PLP5_std_frame' 'PLP5_Δ_mean_frame' 'PLP5_Δ_std_frame' 'PLP5_ΔΔ_std_frame' 'PLP6_Δ_std_frame' 'PLP7_Δ_kurtosis_frame' 'SHIM_APO11_frame' 'SHIM_APO3_frame' 'SHIM_APO5_frame' 'SHIM_DB_frame' 'SHIM_frame' 'SPEC_CENT_GRAVITY_mean_frame' 'SPEC_CENT_GRAVITY_median_frame' 'SPEC_KURTOSIS_mean_frame' 'SPEC_KURTOSIS_median_frame' 'SPEC_KURTOSIS_std_frame' 'SPEC_ROLLOFF_mean_frame' 'SPEC_ROLLOFF_std_frame' 'SPEC_SKEWNESS_mean_frame' 'SPEC_SKEWNESS_median_frame']
BORUTA_ZSCORE	['DFA_COEFF_frame', 'SPEC_CENT_GRAVITY_mean_frame', 'MFCC2_mean_frame', 'PLP3_mean_frame', 'PLP1_Δ_mean_frame', 'HNR_median_frame', 'MFCC4_median_frame', 'MFCC1_ΔΔ_kurtosis_frame', 'PLP10_Δ_kurtosis_frame', 'MFCC3_skewness_frame']
BORUTA_MINMAX	['DFA_COEFF_frame', 'CPP_mean_frame', 'SPEC_CENT_GRAVITY_mean_frame', 'MFCC2_mean_frame', 'MFCC4_mean_frame', 'MFCC5_mean_frame', 'PLP1_Δ_mean_frame', 'SPEC_CENT_GRAVITY_median_frame', 'MFCC2_median_frame', 'MFCC4_median_frame', 'MFCC9_std_frame', 'PLP3_Δ_std_frame', 'PLP0_Δ_kurtosis_frame', 'FORM2_skewness_frame']

(b) Frase 4

Figura 6.3: Lista di feature per le 2 frasi migliori Italiano 1+Privato1

CROSSCORR_ZSCORE	['GNE_mean_frased' 'GNE_median_frased' 'GNE_std_frased' 'MFCC10_Δ_std_frased' 'MFCC10_ΔΔ_std_frased' 'MFCC1_Δ_std_frased' 'MFCC1_ΔΔ_kurtosis_frased' 'MFCC1_Δ_std_frased' 'MFCC2_mean_frased' 'MFCC2_Δ_std_frased' 'MFCC2_ΔΔ_std_frased' 'MFCC2_Δ_std_frased' 'MFCC3_std_frased' 'MFCC3_Δ_std_frased' 'MFCC3_ΔΔ_std_frased' 'MFCC5_Δ_median_frased' 'MFCC5_Δ_std_frased' 'MFCC6_std_frased' 'MFCC6_Δ_std_frased' 'MFCC7_Δ_std_frased' 'MFCC7_ΔΔ_std_frased' 'PLP0_Δ_std_frased' 'PLP0_ΔΔ_std_frased' 'PLP1_Δ_mean_frased' 'PLP2_Δ_std_frased' 'PLP2_ΔΔ_std_frased' 'PLP3_std_frased' 'PLP3_Δ_std_frased' 'PLP3_ΔΔ_std_frased' 'PLP6_median_frased' 'PLP6_Δ_std_frased' 'PLP6_ΔΔ_std_frased' 'PLP7_ΔΔ_std_frased' 'SPEC_CENT_GRAVITY_mean_frased' 'SPEC_CENT_GRAVITY_median_frased' 'SPEC_ROLLOFF_mean_frased' 'SPEC_SKEWNESS_median_frased' 'SPEC_SLOPE_median_frased']
CROSSCORR_MINMAX	['GNE_mean_frased' 'GNE_median_frased' 'GNE_std_frased' 'MFCC1_skewness_frased' 'MFCC2_mean_frased' 'MFCC2_Δ_std_frased' 'MFCC3_std_frased' 'MFCC3_Δ_std_frased' 'MFCC4_Δ_std_frased' 'MFCC4_ΔΔ_std_frased' 'MFCC6_std_frased' 'MFCC6_Δ_std_frased' 'MFCC7_Δ_std_frased' 'PLP0_std_frased' 'PLP0_Δ_std_frased' 'PLP0_ΔΔ_std_frased' 'PLP11_Δ_std_frased' 'PLP11_ΔΔ_std_frased' 'PLP1_kurtosis_frased' 'PLP1_Δ_mean_frased' 'PLP2_Δ_median_frased' 'PLP2_ΔΔ_std_frased' 'PLP3_std_frased' 'PLP3_Δ_std_frased' 'PLP3_ΔΔ_std_frased' 'PLP5_Δ_std_frased' 'PLP5_ΔΔ_std_frased' 'PLP7_Δ_std_frased' 'PLP7_ΔΔ_std_frased' 'SPEC_CENT_GRAVITY_mean_frased' 'SPEC_CENT_GRAVITY_median_frased' 'SPEC_KURTOSIS_median_frased' 'SPEC_ROLLOFF_mean_frased' 'SPEC_ROLLOFF_median_frased' 'SPEC_ROLLOFF_std_frased' 'SPEC_SKEWNESS_mean_frased' 'SPEC_SKEWNESS_median_frased' 'SPEC_SLOPE_median_frased' 'SPEC_STD_mean_frased']
BORUTA_ZSCORE	['GNE_mean_frased', 'PLP11_ΔΔ_mean_frased', 'MFCC5_Δ_median_frased', 'MFCC3_std_frased', 'MFCC6_std_frased', 'MFCC6_Δ_std_frased', 'PLP3_Δ_std_frased', 'PLP0_ΔΔ_std_frased', 'MFCC10_Δ_kurtosis_frased']
BORUTA_MINMAX	['GNE_mean_frased', 'PLP11_ΔΔ_mean_frased', 'MFCC5_Δ_median_frased', 'MFCC3_std_frased', 'MFCC4_Δ_std_frased', 'MFCC6_Δ_std_frased', 'PLP3_Δ_std_frased', 'PLP6_ΔΔ_std_frased']

(a) Frase 1

CROSSCORR_ZSCORE	['FORM2_std_frased' 'GNE_mean_frased' 'GNE_median_frased' 'GNE_std_frased' 'MFCC1_kurtosis_frased' 'MFCC1_std_frased' 'MFCC1_Δ_std_frased' 'MFCC1_ΔΔ_median_frased' 'MFCC2_std_frased' 'MFCC2_Δ_std_frased' 'MFCC2_ΔΔ_skewness_frased' 'MFCC3_std_frased' 'PLP0_std_frased' 'PLP0_Δ_std_frased' 'PLP0_ΔΔ_std_frased' 'PLP11_kurtosis_frased' 'PLP1_mean_frased' 'PLP1_std_frased' 'PLP1_Δ_std_frased' 'PLP2_std_frased' 'PLP2_Δ_std_frased' 'PLP2_ΔΔ_std_frased' 'PLP3_std_frased' 'PLP4_mean_frased' 'PLP5_std_frased' 'PLP5_Δ_std_frased' 'PLP5_ΔΔ_std_frased' 'SPEC_CENT_GRAVITY_mean_frased' 'SPEC_CENT_GRAVITY_median_frased' 'SPEC_KURTOSIS_median_frased' 'SPEC_SKEWNESS_mean_frased' 'SPEC_SKEWNESS_median_frased']
CROSSCORR_MINMAX	['GNE_kurtosis_frased' 'GNE_mean_frased' 'GNE_median_frased' 'GNE_skewness_frased' 'GNE_std_frased' 'HNR_std_frased' 'MFCC1_kurtosis_frased' 'MFCC1_std_frased' 'MFCC1_Δ_std_frased' 'MFCC1_ΔΔ_std_frased' 'MFCC2_std_frased' 'MFCC2_Δ_std_frased' 'MFCC2_ΔΔ_std_frased' 'MFCC3_std_frased' 'MFCC3_Δ_std_frased' 'MFCC3_ΔΔ_std_frased' 'MFCC8_median_frased' 'PLP0_std_frased' 'PLP0_Δ_std_frased' 'PLP0_ΔΔ_std_frased' 'PLP1_kurtosis_frased' 'PLP1_std_frased' 'PLP1_ΔΔ_kurtosis_frased' 'PLP2_std_frased' 'PLP2_Δ_std_frased' 'PLP2_ΔΔ_std_frased' 'PLP3_std_frased' 'PLP3_Δ_std_frased' 'PLP3_ΔΔ_std_frased' 'PLP4_std_frased' 'SPEC_CENT_GRAVITY_mean_frased' 'SPEC_SKEWNESS_median_frased']
BORUTA_ZSCORE	['GNE_mean_frased', 'MFCC1_ΔΔ_median_frased', 'MFCC1_std_frased', 'MFCC3_std_frased', 'PLP0_std_frased', 'PLP0_Δ_std_frased', 'MFCC1_kurtosis_frased', 'MFCC3_kurtosis_frased', 'PLP10_Δ_kurtosis_frased', 'MFCC2_Δ_skewness_frased']
BORUTA_MINMAX	['GNE_mean_frased', 'PLP8_Δ_mean_frased', 'PLP2_ΔΔ_mean_frased', 'GNE_median_frased', 'MFCC1_ΔΔ_median_frased', 'PLP8_Δ_median_frased', 'MFCC1_std_frased', 'MFCC3_std_frased', 'MFCC8_std_frased', 'PLP0_Δ_std_frased', 'MFCC3_kurtosis_frased', 'MFCC9_kurtosis_frased', 'PLP8_Δ_skewness_frased']

(b) Frase 5

Figura 6.4: Lista di feature per le 2 frasi migliori Colombiano

Risultati

CROSSCORR_ZSCORE	['CPP_mean_frase' 'F0_std_frase' 'FORM1_median_frase' 'MFCC11_median_frase' 'MFCC12_Δ_std_frase' 'MFCC1_mean_frase' 'MFCC1_median_frase' 'MFCC1_Δ_median_frase' 'MFCC2_std_frase' 'MFCC2_Δ_std_frase' 'MFCC3_mean_frase' 'MFCC3_median_frase' 'MFCC4_skewness_frase' 'MFCC6_std_frase' 'MFCC6_Δ_std_frase' 'MFCC7_Δ_std_frase' 'MFCC8_median_frase' 'MFCC9_Δ_std_frase' 'PLP0_std_frase' 'PLP0_Δ_kurtosis_frase' 'PLP0_Δ_skewness_frase' 'PLP1_std_frase' 'PLP1_Δ_std_frase' 'PLP1_Δ_std_frase' 'PLP2_std_frase' 'PLP2_Δ_std_frase' 'PLP2_Δ_std_frase' 'PLP3_Δ_kurtosis_frase' 'SPEC_CENT_GRAVITY_mean_frase']
CROSSCORR_MINMAX	['CPP_mean_frase' 'CPP_std_frase' 'F0_mean_frase' 'F0_median_frase' 'F0_std_frase' 'FORM1_mean_frase' 'FORM1_median_frase' 'FORM2_kurtosis_frase' 'FORM2_std_frase' 'GNE_kurtosis_frase' 'GNE_skewness_frase' 'INT_AVG_frase' 'MFCC11_mean_frase' 'MFCC11_median_frase' 'MFCC1_mean_frase' 'MFCC1_median_frase' 'MFCC2_skewness_frase' 'MFCC2_std_frase' 'MFCC2_Δ_median_frase' 'MFCC2_Δ_std_frase' 'MFCC2_Δ_std_frase' 'MFCC3_mean_frase' 'MFCC3_Δ_std_frase' 'MFCC3_Δ_mean_frase' 'MFCC3_Δ_std_frase' 'MFCC4_mean_frase' 'MFCC4_median_frase' 'MFCC4_Δ_kurtosis_frase' 'MFCC5_mean_frase' 'MFCC5_median_frase' 'MFCC6_std_frase' 'MFCC6_Δ_std_frase' 'MFCC8_mean_frase' 'MFCC8_median_frase' 'MFCC9_Δ_std_frase' 'MFCC9_Δ_std_frase' 'PLP0_skewness_frase' 'PLP0_std_frase' 'PLP0_Δ_kurtosis_frase' 'PLP0_Δ_mean_frase' 'PLP0_Δ_skewness_frase' 'PLP0_Δ_std_frase' 'PLP0_Δ_kurtosis_frase' 'PLP0_Δ_skewness_frase' 'PLP10_mean_frase' 'PLP1_Δ_median_frase' 'PLP2_std_frase' 'PLP2_Δ_std_frase' 'PLP2_Δ_std_frase' 'PLP3_Δ_kurtosis_frase' 'PLP3_Δ_skewness_frase' 'PLP3_Δ_kurtosis_frase' 'PLP4_skewness_frase' 'PLP4_std_frase' 'PLP4_Δ_std_frase' 'PLP7_mean_frase' 'SHIM_APQ11_frase' 'SPEC_CENT_GRAVITY_mean_frase' 'SPEC_CENT_GRAVITY_median_frase' 'SPEC_CENT_GRAVITY_std_frase' 'SPEC_ROLLOFF_mean_frase' 'SPEC_ROLLOFF_std_frase' 'STE_mean_frase' 'STE_median_frase' 'STE_std_frase']
BORUTA_ZSCORE	['SPEC_CENT_GRAVITY_mean_frase', 'MFCC1_mean_frase', 'MFCC3_mean_frase', 'MFCC7_Δ_mean_frase', 'PLP0_Δ_mean_frase', 'PLP7_Δ_mean_frase', 'PLP8_Δ_mean_frase', 'SPEC_CENT_GRAVITY_median_frase', 'MFCC11_median_frase', 'MFCC3_median_frase', 'PLP1_median_frase', 'MFCC1_Δ_std_frase', 'MFCC1_Δ_std_frase', 'PLP2_std_frase', 'PLP2_Δ_std_frase', 'PLP5_Δ_std_frase', 'F0_std_frase', 'SPEC_KURTOSIS_std_frase', 'PLP0_Δ_std_frase', 'MFCC9_Δ_kurtosis_frase', 'PLP7_kurtosis_frase', 'PLP0_Δ_kurtosis_frase', 'PLP3_Δ_kurtosis_frase', 'PLP0_Δ_skewness_frase']
BORUTA_MINMAX	['MFCC1_mean_frase', 'MFCC3_mean_frase', 'PLP0_Δ_mean_frase', 'PLP7_Δ_mean_frase', 'PLP8_Δ_mean_frase', 'SPEC_CENT_GRAVITY_median_frase', 'FORM1_median_frase', 'PLP1_Δ_median_frase', 'MFCC1_Δ_std_frase', 'MFCC1_Δ_std_frase', 'PLP2_std_frase', 'PLP2_Δ_std_frase', 'PLP5_Δ_std_frase', 'MFCC4_Δ_kurtosis_frase', 'PLP0_Δ_kurtosis_frase', 'PLP0_Δ_kurtosis_frase', 'FORM2_skewness_frase', 'PLP0_Δ_skewness_frase', 'PLP4_Δ_skewness_frase']

(a) Frase 2

CROSSCORR_ZSCORE	['CPP_std_frase' 'F0_std_frase' 'FORM1_mean_frase' 'FORM1_median_frase' 'FORM2_mean_frase' 'HNR_mean_frase' 'INT_AVG_frase' 'MFCC11_mean_frase' 'MFCC11_Δ_median_frase' 'MFCC13_Δ_std_frase' 'MFCC1_mean_frase' 'MFCC1_std_frase' 'MFCC1_Δ_std_frase' 'MFCC1_Δ_std_frase' 'MFCC2_std_frase' 'MFCC2_Δ_std_frase' 'MFCC3_mean_frase' 'MFCC3_median_frase' 'MFCC5_Δ_median_frase' 'PLP0_kurtosis_frase' 'PLP0_kurtosis_frase' 'PLP0_skewness_frase' 'PLP0_std_frase' 'PLP0_Δ_skewness_frase' 'PLP0_Δ_std_frase' 'PLP0_Δ_std_frase' 'PLP0_Δ_std_frase' 'PLP10_Δ_kurtosis_frase' 'PLP1_mean_frase' 'PLP1_std_frase' 'PLP1_Δ_std_frase' 'PLP1_Δ_std_frase' 'PLP2_median_frase' 'PLP2_std_frase' 'PLP2_Δ_std_frase' 'PLP2_Δ_std_frase' 'PLP4_mean_frase' 'PLP4_median_frase' 'PLP7_mean_frase' 'PLP7_median_frase' 'SPEC_CENT_GRAVITY_mean_frase' 'SPEC_CENT_GRAVITY_std_frase' 'SPEC_FLUX_mean_frase' 'SPEC_KURTOSIS_mean_frase' 'SPEC_ROLLOFF_mean_frase' 'SPEC_ROLLOFF_std_frase' 'SPEC_STD_mean_frase' 'SPEC_STD_mean_frase' 'STE_mean_frase' 'STE_mean_frase' 'STE_std_frase']
CROSSCORR_MINMAX	['F0_mean_frase' 'F0_median_frase' 'F0_std_frase' 'FORM1_kurtosis_frase' 'FORM1_mean_frase' 'FORM1_median_frase' 'FORM1_skewness_frase' 'FORM2_mean_frase' 'FORM_BAND2_std_frase' 'FORM_BAND3_std_frase' 'HNR_mean_frase' 'INT_AVG_frase' 'MFCC11_mean_frase' 'MFCC11_median_frase' 'MFCC11_Δ_mean_frase' 'MFCC13_Δ_std_frase' 'MFCC1_mean_frase' 'MFCC1_median_frase' 'MFCC1_Δ_std_frase' 'MFCC1_Δ_std_frase' 'MFCC2_std_frase' 'MFCC2_Δ_std_frase' 'MFCC2_Δ_std_frase' 'MFCC3_mean_frase' 'MFCC3_median_frase' 'MFCC6_Δ_mean_frase' 'PLP0_median_frase' 'PLP0_skewness_frase' 'PLP0_std_frase' 'PLP0_Δ_mean_frase' 'PLP0_Δ_skewness_frase' 'PLP0_Δ_std_frase' 'PLP10_Δ_std_frase' 'PLP11_Δ_mean_frase' 'PLP1_Δ_std_frase' 'PLP1_Δ_std_frase' 'PLP2_median_frase' 'PLP2_std_frase' 'PLP2_Δ_std_frase' 'PLP2_Δ_std_frase' 'PLP3_Δ_kurtosis_frase' 'PLP3_Δ_mean_frase' 'PLP4_mean_frase' 'PLP4_median_frase' 'PLP6_std_frase' 'PLP7_median_frase' 'PLP8_Δ_mean_frase' 'PLP9_Δ_skewness_frase' 'SPEC_CENT_GRAVITY_mean_frase' 'SPEC_CENT_GRAVITY_std_frase' 'SPEC_CENT_GRAVITY_std_frase' 'SPEC_FLUX_mean_frase' 'SPEC_ROLLOFF_mean_frase' 'SPEC_ROLLOFF_std_frase' 'SPEC_STD_mean_frase' 'STE_mean_frase']
BORUTA_ZSCORE	['MFCC1_mean_frase', 'MFCC3_mean_frase', 'PLP8_Δ_mean_frase', 'MFCC11_median_frase', 'MFCC3_median_frase', 'PLP2_std_frase', 'PLP2_Δ_std_frase', 'F0_skewness_frase', 'MFCC4_skewness_frase', 'PLP2_skewness_frase', 'PLP9_Δ_skewness_frase']
BORUTA_MINMAX	['HNR_mean_frase', 'MFCC1_mean_frase', 'MFCC3_mean_frase', 'STE_mean_frase', 'MFCC3_median_frase', 'F0_std_frase', 'MFCC2_std_frase', 'MFCC1_Δ_std_frase', 'PLP2_std_frase', 'PLP1_Δ_std_frase', 'PLP2_Δ_std_frase', 'PLP0_Δ_std_frase', 'PLP1_Δ_std_frase', 'MFCC4_skewness_frase']

(b) Frase 5

Figura 6.5: Lista di feature per le 2 frasi migliori Italiano 2

CROSSCORR_ZSCORE	['INT_AVG_a' 'MFCC1_mean_a' 'MFCC1_median_a' 'MFCC1_Δ_mean_a' 'MFCC2_mean_a' 'MFCC2_median_a' 'MFCC2_Δ_skewness_a' 'MFCC4_mean_a' 'MFCC5_skewness_a' 'PLP0_mean_a' 'PLP0_median_a' 'PLP0_skewness_a' 'PLP0_Δ_skewness_a' 'PLP1_mean_a' 'PLP1_skewness_a' 'PLP1_Δ_skewness_a' 'PLP2_mean_a' 'PLP2_skewness_a' 'PLP2_ΔΔ_skewness_a' 'PLP6_skewness_a' 'SPEC_CENT_GRAVITY_mean_a' 'SPEC_FLUX_mean_a' 'SPEC_FLUX_median_a' 'SPEC_SKEWNESS_mean_a' 'STE_mean_a' 'STE_median_a' 'STE_std_a']
CROSSCORR_MINIMA	['INT_AVG_a' 'MFCC1_mean_a' 'MFCC1_median_a' 'MFCC1_Δ_mean_a' 'MFCC2_mean_a' 'MFCC2_median_a' 'MFCC2_Δ_skewness_a' 'MFCC4_mean_a' 'MFCC5_skewness_a' 'PLP0_mean_a' 'PLP0_median_a' 'PLP0_skewness_a' 'PLP0_Δ_skewness_a' 'PLP1_mean_a' 'PLP1_skewness_a' 'PLP1_Δ_skewness_a' 'PLP2_mean_a' 'PLP2_skewness_a' 'PLP2_ΔΔ_skewness_a' 'PLP6_skewness_a' 'SPEC_CENT_GRAVITY_mean_a' 'SPEC_FLUX_mean_a' 'SPEC_FLUX_median_a' 'SPEC_SKEWNESS_mean_a' 'STE_mean_a' 'STE_median_a' 'STE_std_a']
BORUTA_ZSCORE	['INT_AVG_a' 'MFCC1_mean_a' 'MFCC1_median_a' 'MFCC1_Δ_mean_a' 'MFCC2_mean_a' 'MFCC2_median_a' 'MFCC2_Δ_skewness_a' 'MFCC4_mean_a' 'MFCC5_skewness_a' 'PLP0_mean_a' 'PLP0_median_a' 'PLP0_skewness_a' 'PLP0_Δ_skewness_a' 'PLP1_mean_a' 'PLP1_skewness_a' 'PLP1_Δ_skewness_a' 'PLP2_mean_a' 'PLP2_skewness_a' 'PLP2_ΔΔ_skewness_a' 'PLP6_skewness_a' 'SPEC_CENT_GRAVITY_mean_a' 'SPEC_FLUX_mean_a' 'SPEC_FLUX_median_a' 'SPEC_SKEWNESS_mean_a' 'STE_mean_a' 'STE_median_a' 'STE_std_a']
BORUTA_MINIMA	['PLP3_Δ_mean_a', 'STE_median_a', 'SPEC_FLUX_median_a', 'MFCC6_ΔΔ_mean_a', 'MFCC7_ΔΔ_mean_a', 'SPEC_SLOPE_std_a', 'PLP0_Δ_skewness_a', 'PLP0_ΔΔ_skewness_a']

(a) Vocalizzo /a/

CROSSCORR_ZSCORE	['INT_AVG_a' 'MFCC1_mean_a' 'MFCC1_median_a' 'MFCC1_Δ_mean_a' 'MFCC2_mean_a' 'MFCC2_median_a' 'MFCC2_Δ_skewness_a' 'MFCC4_mean_a' 'MFCC5_skewness_a' 'PLP0_mean_a' 'PLP0_median_a' 'PLP0_skewness_a' 'PLP0_Δ_skewness_a' 'PLP1_mean_a' 'PLP1_skewness_a' 'PLP1_Δ_skewness_a' 'PLP2_mean_a' 'PLP2_skewness_a' 'PLP2_ΔΔ_skewness_a' 'PLP6_skewness_a' 'SPEC_CENT_GRAVITY_mean_a' 'SPEC_FLUX_mean_a' 'SPEC_FLUX_median_a' 'SPEC_SKEWNESS_mean_a' 'STE_mean_a' 'STE_median_a' 'STE_std_a']
CROSSCORR_MINIMA	['INT_AVG_a' 'MFCC1_mean_a' 'MFCC1_median_a' 'MFCC1_Δ_mean_a' 'MFCC2_mean_a' 'MFCC2_median_a' 'MFCC2_Δ_skewness_a' 'MFCC4_mean_a' 'MFCC5_skewness_a' 'PLP0_mean_a' 'PLP0_median_a' 'PLP0_skewness_a' 'PLP0_Δ_skewness_a' 'PLP1_mean_a' 'PLP1_skewness_a' 'PLP1_Δ_skewness_a' 'PLP2_mean_a' 'PLP2_skewness_a' 'PLP2_ΔΔ_skewness_a' 'PLP6_skewness_a' 'SPEC_CENT_GRAVITY_mean_a' 'SPEC_FLUX_mean_a' 'SPEC_FLUX_median_a' 'SPEC_SKEWNESS_mean_a' 'STE_mean_a' 'STE_median_a' 'STE_std_a']
BORUTA_ZSCORE	['INT_AVG_a' 'MFCC1_mean_a' 'MFCC1_median_a' 'MFCC1_Δ_mean_a' 'MFCC2_mean_a' 'MFCC2_median_a' 'MFCC2_Δ_skewness_a' 'MFCC4_mean_a' 'MFCC5_skewness_a' 'PLP0_mean_a' 'PLP0_median_a' 'PLP0_skewness_a' 'PLP0_Δ_skewness_a' 'PLP1_mean_a' 'PLP1_skewness_a' 'PLP1_Δ_skewness_a' 'PLP2_mean_a' 'PLP2_skewness_a' 'PLP2_ΔΔ_skewness_a' 'PLP6_skewness_a' 'SPEC_CENT_GRAVITY_mean_a' 'SPEC_FLUX_mean_a' 'SPEC_FLUX_median_a' 'SPEC_SKEWNESS_mean_a' 'STE_mean_a' 'STE_median_a' 'STE_std_a']
BORUTA_MINIMA	['INT_AVG_a' 'MFCC1_mean_a' 'MFCC1_median_a' 'MFCC1_Δ_mean_a' 'MFCC2_mean_a' 'MFCC2_median_a' 'MFCC2_Δ_skewness_a' 'MFCC4_mean_a' 'MFCC5_skewness_a' 'PLP0_mean_a' 'PLP0_median_a' 'PLP0_skewness_a' 'PLP0_Δ_skewness_a' 'PLP1_mean_a' 'PLP1_skewness_a' 'PLP1_Δ_skewness_a' 'PLP2_mean_a' 'PLP2_skewness_a' 'PLP2_ΔΔ_skewness_a' 'PLP6_skewness_a' 'SPEC_CENT_GRAVITY_mean_a' 'SPEC_FLUX_mean_a' 'SPEC_FLUX_median_a' 'SPEC_SKEWNESS_mean_a' 'STE_mean_a' 'STE_median_a' 'STE_std_a']

(b) Frase migliore

Figura 6.6: Lista di feature Italiano1+Privato1 per il vocalizzo /a/ e la frase migliore

CROSSCORR_ZSCORE	['DFA_COEFF_frased', 'FORM1_mean_frased', 'FORM1_median_frased', 'FORM1_skewness_frased', 'FORM1_std_a', 'FORM1_std_frased', 'FORM3_kurtosis_frased', 'HNR_mean_frased', 'HNR_median_frased', 'INT_AVG_a', 'MFCC10_Δ_mean_a', 'MFCC11_Δ_skewness_a', 'MFCC12_Δ_mean_frased', 'MFCC13_skewness_frased', 'MFCC1_mean_a', 'MFCC1_mean_frased', 'MFCC1_median_a', 'MFCC1_median_frased', 'MFCC1_skewness_frased', 'MFCC1_Δ_mean_frased', 'MFCC1_ΔΔ_kurtosis_a', 'MFCC2_mean_a', 'MFCC2_mean_frased', 'MFCC2_median_a', 'MFCC2_median_frased', 'MFCC2_std_frased', 'MFCC2_ΔΔ_mean_a', 'MFCC3_mean_frased', 'MFCC3_mean_frased', 'MFCC3_Δ_kurtosis_a', 'MFCC3_Δ_skewness_a', 'MFCC4_mean_a', 'MFCC4_mean_frased', 'MFCC4_median_a', 'MFCC4_median_frased', 'MFCC5_mean_frased', 'MFCC5_mean_frased', 'MFCC5_median_frased', 'MFCC5_skewness_a', 'MFCC5_Δ_mean_frased', 'MFCC6_mean_frased', 'MFCC6_median_frased', 'MFCC7_mean_a', 'MFCC7_mean_frased', 'MFCC7_median_a', 'MFCC7_median_frased', 'MFCC7_ΔΔ_kurtosis_frased', 'MFCC9_mean_a', 'MFCC9_mean_frased', 'PLP0_mean_a', 'PLP0_mean_frased', 'PLP0_median_a', 'PLP0_median_frased', 'PLP0_skewness_a', 'PLP0_Δ_kurtosis_frased', 'PLP0_Δ_skewness_a', 'PLP0_ΔΔ_kurtosis_frased', 'PLP0_ΔΔ_skewness_a', 'PLP0_ΔΔ_std_frased', 'PLP10_skewness_a', 'PLP11_mean_a', 'PLP11_mean_frased', 'PLP11_Δ_std_frased', 'PLP11_ΔΔ_mean_frased', 'PLP11_ΔΔ_std_frased', 'PLP1_skewness_frased', 'PLP1_Δ_mean_frased', 'PLP1_Δ_skewness_frased', 'PLP1_ΔΔ_mean_frased', 'PLP1_ΔΔ_skewness_frased', 'PLP2_std_frased', 'PLP2_Δ_std_frased', 'PLP2_ΔΔ_std_frased', 'PLP3_skewness_a', 'PLP3_ΔΔ_skewness_a', 'PLP4_std_frased', 'PLP4_Δ_mean_frased', 'PLP4_Δ_std_frased', 'PLP4_ΔΔ_skewness_frased', 'PLP4_ΔΔ_std_frased', 'PLP5_std_frased', 'PLP5_Δ_mean_a', 'PLP5_Δ_std_frased', 'PLP5_ΔΔ_skewness_frased', 'PLP5_ΔΔ_std_frased', 'PLP6_std_frased', 'PLP6_Δ_std_frased', 'PLP6_ΔΔ_std_frased', 'PLP9_std_frased', 'PLP9_Δ_std_frased', 'PLP9_ΔΔ_std_frased', 'SHIM_APO3_frased', 'SHIM_APO5_frased', 'SHIM_DB_frased', 'SHIM_frased', 'SPEC_CENT_GRAVITY_mean_a', 'SPEC_CENT_GRAVITY_mean_frased', 'SPEC_CENT_GRAVITY_median_a', 'SPEC_CENT_GRAVITY_median_frased', 'SPEC_CENT_GRAVITY_std_frased', 'SPEC_DECREASE_median_frased', 'SPEC_FLUX_mean_a', 'SPEC_FLUX_mean_frased', 'SPEC_KURTOSIS_mean_frased', 'SPEC_KURTOSIS_mean_frased', 'SPEC_KURTOSIS_mean_frased', 'SPEC_KURTOSIS_std_frased', 'SPEC_ROLLOFF_mean_frased', 'SPEC_ROLLOFF_mean_frased', 'SPEC_SKEWNESS_mean_a', 'SPEC_SKEWNESS_mean_frased', 'SPEC_SKEWNESS_mean_a', 'SPEC_SKEWNESS_mean_frased', 'SPEC_SKEWNESS_std_frased', 'STE_mean_a', 'STE_mean_frased', 'STE_std_a', 'STE_std_frased']
CROSSCORR_MINMAX	['DFA_COEFF_frased', 'FORM1_mean_frased', 'FORM1_median_frased', 'FORM1_skewness_frased', 'FORM1_std_a', 'FORM1_std_frased', 'FORM_BAND1_mean_frased', 'FORM_BAND1_median_frased', 'GNE_skewness_a', 'GNE_std_a', 'HNR_kurtosis_a', 'HNR_mean_frased', 'HNR_skewness_a', 'HNR_skewness_frased', 'HNR_std_frased', 'INT_AVG_a', 'MFCC10_ΔΔ_mean_frased', 'MFCC13_skewness_frased', 'MFCC1_mean_a', 'MFCC1_mean_frased', 'MFCC1_median_a', 'MFCC1_median_frased', 'MFCC1_skewness_frased', 'MFCC1_Δ_mean_frased', 'MFCC1_Δ_mean_frased', 'MFCC2_mean_a', 'MFCC2_mean_frased', 'MFCC2_median_a', 'MFCC2_median_frased', 'MFCC2_std_frased', 'MFCC2_std_frased', 'MFCC3_mean_frased', 'MFCC3_mean_frased', 'MFCC3_std_frased', 'MFCC3_ΔΔ_kurtosis_frased', 'MFCC4_mean_a', 'MFCC4_mean_frased', 'MFCC4_median_a', 'MFCC4_median_frased', 'MFCC5_mean_frased', 'MFCC5_mean_frased', 'MFCC5_Δ_mean_frased', 'MFCC5_Δ_mean_frased', 'MFCC5_ΔΔ_skewness_frased', 'MFCC6_mean_frased', 'MFCC6_mean_frased', 'MFCC6_median_frased', 'MFCC6_skewness_a', 'MFCC7_mean_a', 'MFCC7_mean_frased', 'MFCC7_median_a', 'MFCC7_median_frased', 'MFCC7_ΔΔ_mean_a', 'PLP0_mean_a', 'PLP0_mean_frased', 'PLP0_median_a', 'PLP0_median_frased', 'PLP0_skewness_a', 'PLP0_Δ_kurtosis_frased', 'PLP0_Δ_mean_a', 'PLP0_Δ_mean_frased', 'PLP0_Δ_skewness_a', 'PLP0_Δ_std_frased', 'PLP0_ΔΔ_kurtosis_frased', 'PLP0_ΔΔ_skewness_a', 'PLP0_ΔΔ_std_frased', 'PLP10_skewness_a', 'PLP11_std_a', 'PLP11_std_frased', 'PLP1_kurtosis_a', 'PLP1_Δ_mean_frased', 'PLP1_Δ_skewness_frased', 'PLP2_mean_frased', 'PLP2_std_frased', 'PLP2_Δ_mean_frased', 'PLP2_Δ_std_frased', 'PLP2_ΔΔ_std_frased', 'PLP3_std_frased', 'PLP3_Δ_std_frased', 'PLP3_ΔΔ_std_frased', 'PLP4_std_frased', 'PLP4_Δ_std_frased', 'PLP4_ΔΔ_std_frased', 'PLP5_mean_frased', 'PLP5_mean_frased', 'PLP5_std_frased', 'PLP5_Δ_mean_frased', 'PLP5_Δ_std_frased', 'PLP5_ΔΔ_skewness_frased', 'PLP5_ΔΔ_std_frased', 'PLP6_std_frased', 'PLP6_Δ_std_frased', 'PLP6_ΔΔ_std_frased', 'PLP7_std_frased', 'PLP9_std_frased', 'PLP9_Δ_std_frased', 'PLP9_ΔΔ_std_frased', 'SHIM_APO11_frased', 'SHIM_APO3_frased', 'SHIM_APO5_frased', 'SHIM_DB_frased', 'SHIM_frased', 'SPEC_CENT_GRAVITY_mean_a', 'SPEC_CENT_GRAVITY_mean_frased', 'SPEC_CENT_GRAVITY_median_a', 'SPEC_CENT_GRAVITY_median_frased', 'SPEC_CENT_GRAVITY_std_frased', 'SPEC_DECREASE_median_frased', 'SPEC_FLUX_mean_a', 'SPEC_KURTOSIS_mean_a', 'SPEC_KURTOSIS_mean_frased', 'SPEC_KURTOSIS_mean_frased', 'SPEC_KURTOSIS_mean_a', 'SPEC_KURTOSIS_std_frased', 'SPEC_KURTOSIS_std_frased', 'SPEC_ROLLOFF_mean_frased', 'SPEC_ROLLOFF_mean_frased', 'SPEC_ROLLOFF_std_frased', 'SPEC_SKEWNESS_mean_a', 'SPEC_SKEWNESS_mean_frased', 'SPEC_SKEWNESS_mean_a', 'SPEC_SKEWNESS_mean_frased', 'SPEC_SKEWNESS_std_frased', 'SPEC_SLOPE_mean_frased', 'SPEC_STD_kurtosis_a', 'SPEC_STD_kurtosis_frased', 'STE_mean_a', 'STE_mean_frased', 'STE_std_a', 'STE_std_frased']

(a) Crosscorrelazione

BORUTA_ZSCORE	['DFA_COEFF_frased', 'HNR_mean_frased', 'MFCC2_mean_frased', 'MFCC4_mean_frased', 'PLP4_Δ_mean_frased', 'PLP1_ΔΔ_mean_frased', 'MFCC2_median_frased', 'MFCC6_kurtosis_frased', 'PLP0_Δ_kurtosis_frased', 'PLP5_Δ_mean_a', 'MFCC6_ΔΔ_mean_a', 'MFCC13_ΔΔ_std_a', 'MFCC13_Δ_kurtosis_a', 'PLP0_Δ_skewness_a', 'PLP0_ΔΔ_skewness_a']
BORUTA_MINMAX	['DFA_COEFF_frased', 'HNR_mean_frased', 'SPEC_CENT_GRAVITY_mean_frased', 'MFCC2_mean_frased', 'MFCC4_mean_frased', 'MFCC5_Δ_mean_frased', 'PLP0_mean_frased', 'PLP4_Δ_mean_frased', 'PLP5_Δ_mean_frased', 'HNR_mean_frased', 'MFCC2_median_frased', 'PLP1_Δ_mean_frased', 'FORM1_kurtosis_frased', 'MFCC13_skewness_frased', 'PLP0_mean_a', 'PLP0_Δ_skewness_a', 'PLP0_ΔΔ_skewness_a']

(b) Boruta

Figura 6.7: Lista di feature estratte dall’algoritmo Boruta vocalizzo /a/ unito alla frase migliore su Italiano1+Privato1

[illegible]

(a) Crosscorrelazione

BORUTA_ZSCORE	['DFA_COEFF__f1', 'HNR_mean__f1', 'SPEC_CENT_GRAVITY_mean__f1', 'MFCC2_mean__f1', 'PLP0_mean__f1', 'HNR_median__f1', 'SPEC_SKEWNESS_median__f1', 'MFCC2_median__f1', 'F0_std__f1', 'SPEC_FLUX_median_a', 'PLP2_median_a', 'SPEC_SLOPE_std_a', 'PLP0_skewness_a', 'PLP0_A_skewness_a', 'DFA_COEFF__f2', 'SPEC_DECREASE_mean__f2', 'MFCC7_A_mean__f2', 'PLP8_mean__f2', 'HNR_median__f2', 'PLP8_ΔA_std__f2', 'PLP10_A_kurtosis__f2']
BORUTA_MINMAX	['DFA_COEFF__f1', 'HNR_mean__f1', 'MFCC4_mean__f1', 'SPEC_SKEWNESS_median__f1', 'MFCC2_median__f1', 'PLP2_A_skewness__f1', 'PLP0_A_mean_a', 'SPEC_SLOPE_std_a', 'SPEC_DECREASE_std__f2']

(b) Boruta

Figura 6.8: Lista di feature estratte dall’algoritmo boruta vocalizzo /a/ unito alle 2 frasi migliori su Italiano1+Privato1

CROSSCORR_ZSCORE	['HNR_mean_a','MFCC2_ΔΔ_std_a','MFCC5_Δ_mean_a','MFCC6_kurtosis_a', 'PLP10_median_a','PLP8_std_a','PLP8_Δ_std_a']
CROSSCORR_MINMAX	['FO_std_a','FORM_BAND1_mean_a','FORM_BAND1_median_a','MFCC10_Δ_std_a','MFCC10_ΔΔ_std_a','MFCC13_Δ_std_a','MFCC13_ΔΔ_std_a','MFCC1_Δ_std_a', 'MFCC1_ΔΔ_std_a','MFCC2_Δ_std_a','MFCC2_ΔΔ_std_a','MFCC3_mean_a','MFCC3_median_a','MFCC3_Δ_std_a','MFCC3_ΔΔ_std_a','MFCC4_Δ_std_a', 'MFCC4_ΔΔ_std_a','PLP1_ΔΔ_std_a','PLP2_ΔΔ_std_a','PLP4_ΔΔ_std_a','SHIM_APQ11_a','SHIM_APQ3_a','SHIM_DB_a','SHIM_a']
BORUTA_ZSCORE	['JIT_L_a','PLP2_ΔΔ_mean_a','FORM1_std_a','MFCC11_ΔΔ_std_a','PLP8_std_a','FORM_BAND3_kurtosis_a']
BORUTA_MINMAX	['SHIM_APQ11_a','PLP2_ΔΔ_mean_a','PLP10_median_a','MFCC4_ΔΔ_std_a','MFCC13_Δ_std_a','GNE_kurtosis_a','FORM_BAND3_kurtosis_a','FORM1_skewness_a', 'MFCC2_ΔΔ_skewness_a','PLP11_skewness_a']

Figura 6.9: Lista di feature vocalizzo /a/ su Colombiano

CROSSCORR_ZSCORE	['GNE_mean_frased','GNE_median_frased','MFCC1_kurtosis_frased','MFCC2_Δ_std_frased','MFCC2_ΔΔ_std_frased','MFCC3_std_frased','MFCC3_Δ_std_frased','MFCC3_ΔΔ_std_frased','MFCC5_Δ_std_frased','MFCC6_Δ_std_frased','MFCC7_Δ_std_frased', 'PLP1_Δ_mean_frased','PLP3_std_frased','PLP3_Δ_std_frased','PLP3_ΔΔ_std_frased','PLP5_Δ_std_frased','PLP5_ΔΔ_std_frased','PLP6_Δ_std_frased','PLP6_ΔΔ_std_frased','SPEC_CENT_GRAVITY_mean_frased','SPEC_CENT_GRAVITY_median_frased', 'SPEC_ROLLOFF_mean_frased','SPEC_ROLLOFF_std_frased','SPEC_SLOPE_median_frased']
CROSSCORR_MINMAX	['FORM_BAND3_median_frased','GNE_mean_frased','GNE_median_frased','GNE_skewness_frased','GNE_std_frased','MFCC1_Δ_std_frased','MFCC1_ΔΔ_std_frased','MFCC2_Δ_std_frased','MFCC2_ΔΔ_std_frased','MFCC3_std_frased','MFCC3_Δ_std_frased', 'MFCC3_ΔΔ_std_frased','MFCC5_Δ_median_frased','MFCC6_std_frased','MFCC6_Δ_std_frased','MFCC7_Δ_std_frased','PLP0_Δ_std_frased','PLP0_ΔΔ_std_frased','PLP1_kurtosis_frased','PLP1_std_frased','PLP1_Δ_std_frased', 'PLP1_ΔΔ_std_frased','PLP3_std_frased','PLP3_Δ_std_frased','PLP3_ΔΔ_std_frased','PLP4_mean_frased','PLP6_median_frased','PLP6_ΔΔ_std_frased','PLP7_kurtosis_frased','SPEC_CENT_GRAVITY_mean_frased','SPEC_KURTOSIS_median_frased', 'SPEC_ROLLOFF_mean_frased','SPEC_ROLLOFF_std_frased','SPEC_SKEWNESS_mean_frased','SPEC_SKEWNESS_median_frased']
BORUTA_ZSCORE	['JIT_PPOS_frased','GNE_mean_frased','PLP4_mean_frased','PLP1_Δ_mean_frased','GNE_median_frased','MFCC3_median_frased','MFCC5_Δ_median_frased','MFCC3_std_frased','MFCC6_std_frased','MFCC6_Δ_std_frased','PLP2_Δ_std_frased', 'PLP3_Δ_std_frased','PLP2_ΔΔ_std_frased','MFCC8_Δ_skewness_frased']
BORUTA_MINMAX	['GNE_mean_frased','FORM1_mean_frased','PLP4_mean_frased','PLP8_Δ_mean_frased','GNE_median_frased','MFCC3_median_frased','MFCC5_Δ_median_frased','MFCC13_Δ_median_frased','MFCC3_std_frased','MFCC2_Δ_std_frased', 'MFCC6_Δ_std_frased','PLP3_Δ_std_frased','MFCC1_ΔΔ_kurtosis_frased','MFCC2_ΔΔ_skewness_frased']

Figura 6.10: Lista di feature frase migliore su Colombiano

Figura 6.13: Lista di feature vocalizzo /e/ su Italiano 2

Figura 6.14: Lista di feature frase migliore su Italiano 2

Classif	Cross-corr + z-score						Cross-corr + minmax					
	Val	Test	Spec.	Sens.	AUC	F1	Val	Test	Spec.	Sens.	AUC	F1
KNN	80%	75%	78%	73%	76%	79%	70%	71%	67%	73%	70%	76%
LSVM	84%	71%	89%	60%	74%	72%	55%	63%	0%	100%	50%	77%
GPr	81%	75%	67%	80%	73%	80%	74%	83%	89%	80%	84%	86%
DT	70%	54%	56%	53%	54%	59%	64%	71%	67%	73%	70%	76%
RF	80%	71%	67%	73%	70%	76%	76%	79%	78%	80%	79%	83%
NN	77%	67%	56%	73%	64%	73%	74%	83%	89%	80%	84%	86%
AdaB	70%	63%	67%	60%	63%	67%	72%	75%	67%	80%	73%	80%
NB	80%	75%	78%	73%	76%	79%	74%	75%	89%	67%	78%	77%
LDA	70%	58%	78%	47%	62%	58%	66%	75%	89%	67%	78%	77%
XGB	74%	63%	56%	67%	61%	69%	59%	79%	67%	87%	77%	84%

(a) Crosscorrelazione

Classif	Boruta + z-score						Boruta + minmax					
	Val	Test	Spec.	Sens.	AUC	F1	Val	Test	Spec.	Sens.	AUC	F1
KNN	69%	67%	45%	85%	65%	73%	74%	67%	88%	56%	72%	69%
LSVM	76%	71%	55%	85%	70%	76%	62%	67%	0%	100%	50%	80%
GPr	74%	71%	55%	85%	70%	76%	80%	67%	88%	56%	72%	69%
DT	76%	67%	36%	92%	64%	75%	65%	54%	38%	63%	50%	65%
RF	85%	67%	36%	92%	64%	75%	78%	63%	88%	50%	69%	64%
NN	74%	71%	55%	85%	70%	76%	82%	67%	88%	56%	72%	69%
AdaB	74%	63%	36%	85%	60%	71%	78%	71%	63%	75%	69%	77%
NB	66%	50%	0%	92%	46%	67%	62%	71%	25%	94%	59%	81%
LDA	72%	75%	64%	85%	74%	79%	80%	58%	75%	50%	63%	62%
XGB	78%	63%	27%	92%	60%	73%	76%	63%	88%	50%	69%	64%

(b) Boruta

Tabella 6.8: Risultati delle metriche di valutazione ottenuti su 10 Classificatori, vocalizzo /a/ su Italiano1+Privato1

Classif	Cross-corr + z-score						Cross-corr + minmax					
	Val	Test	Spec.	Sens.	AUC	F1	Val	Test	Spec.	Sens.	AUC	F1
KNN	80%	75%	78%	73%	76%	79%	78%	79%	78%	80%	79%	83%
LSVM	84%	71%	89%	60%	74%	72%	55%	58%	0%	93%	47%	74%
GPr	81%	75%	67%	80%	73%	80%	83%	67%	67%	67%	67%	71%
DT	72%	54%	56%	53%	54%	59%	72%	54%	56%	53%	54%	59%
RF	68%	58%	56%	60%	58%	64%	76%	63%	78%	53%	66%	64%
NN	77%	67%	56%	73%	64%	73%	80%	67%	67%	67%	67%	71%
AdaB	68%	63%	67%	60%	63%	67%	70%	63%	67%	60%	63%	67%
NB	80%	75%	78%	73%	76%	79%	80%	75%	78%	73%	76%	79%
LDA	70%	58%	78%	47%	62%	58%	70%	58%	78%	47%	62%	58%
XGB	74%	63%	56%	67%	61%	69%	74%	63%	56%	67%	61%	69%

(a) Crosscorrelazione

Classif	Boruta + z-score						Boruta + minmax					
	Val	Test	Spec.	Sens.	AUC	F1	Val	Test	Spec.	Sens.	AUC	F1
KNN	87%	67%	67%	67%	67%	71%	85%	67%	67%	67%	67%	71%
LSVM	83%	63%	44%	73%	59%	71%	53%	63%	0%	100%	50%	77%
GPr	82%	58%	44%	67%	56%	67%	82%	63%	44%	73%	59%	71%
DT	84%	71%	56%	80%	68%	77%	76%	58%	44%	67%	56%	67%
RF	74%	83%	78%	87%	82%	87%	78%	88%	78%	93%	86%	90%
NN	85%	58%	44%	67%	56%	67%	83%	58%	33%	73%	53%	69%
AdaB	76%	75%	67%	80%	73%	80%	71%	71%	67%	73%	70%	76%
NB	57%	67%	11%	100%	56%	79%	57%	67%	11%	100%	56%	79%
LDA	82%	58%	44%	67%	56%	67%	82%	58%	44%	67%	56%	67%
XGB	76%	67%	56%	73%	64%	73%	76%	67%	56%	73%	64%	73%

(b) Boruta

Tabella 6.9: Risultati delle metriche di valutazione ottenuti su 10 Classificatori, frase migliore su Italiano1+Privato1

Classif	Cross-corr + z-score						Cross-corr + minmax					
	Val	Test	Spec.	Sens.	AUC	F1	Val	Test	Spec.	Sens.	AUC	F1
KNN	79%	71%	50%	86%	68%	77%	89%	79%	89%	73%	81%	81%
LSVM	79%	71%	50%	86%	68%	77%	84%	71%	89%	60%	74%	72%
GPr	78%	71%	40%	93%	66%	79%	85%	75%	89%	67%	78%	77%
DT	66%	67%	40%	86%	63%	75%	73%	63%	78%	53%	66%	64%
RF	74%	79%	90%	71%	81%	80%	81%	71%	78%	67%	72%	74%
NN	81%	67%	50%	79%	64%	73%	87%	75%	89%	67%	78%	77%
AdaB	72%	75%	60%	86%	73%	80%	76%	71%	78%	67%	72%	74%
NB	80%	83%	90%	79%	84%	85%	87%	71%	89%	60%	74%	72%
LDA	74%	50%	50%	50%	50%	54%	78%	50%	44%	53%	49%	57%
XGB	72%	71%	40%	93%	66%	79%	79%	79%	89%	73%	81%	81%

(a) Crosscorrelazione

Classif	Boruta + z-score						Boruta + minmax					
	Val	Test	Spec.	Sens.	AUC	F1	Val	Test	Spec.	Sens.	AUC	F1
KNN	93%	67%	63%	69%	66%	73%	93%	83%	78%	87%	82%	87%
LSVM	94%	75%	75%	75%	75%	80%	53%	67%	11%	100%	56%	79%
GPr	95%	75%	75%	75%	75%	80%	96%	83%	78%	87%	82%	87%
DT	80%	79%	100%	69%	84%	81%	64%	92%	100%	87%	93%	93%
RF	93%	79%	88%	75%	81%	83%	91%	79%	78%	80%	79%	83%
NN	95%	79%	75%	81%	78%	84%	96%	83%	78%	87%	82%	87%
AdaB	78%	63%	63%	63%	63%	69%	89%	83%	78%	87%	82%	87%
NB	87%	79%	88%	75%	81%	83%	91%	88%	89%	87%	88%	90%
LDA	96%	67%	63%	69%	66%	73%	93%	75%	67%	80%	73%	80%
XGB	87%	79%	100%	69%	84%	81%	85%	83%	78%	87%	82%	87%

(b) Boruta

Tabella 6.10: Risultati delle metriche di valutazione ottenuti su 10 Classificatori, vocalizzo /a/ unito alla frase migliore su Italiano1+Privato1

Classif	Cross-corr + z-score						Cross-corr + minmax					
	Val	Test	Spec.	Sens.	AUC	F1	Val	Test	Spec.	Sens.	AUC	F1
KNN	83%	75%	90%	64%	77%	75%	83%	67%	56%	73%	64%	73%
LSVM	83%	71%	80%	64%	72%	72%	84%	75%	100%	60%	80%	75%
GPr	83%	75%	70%	79%	74%	79%	81%	75%	78%	73%	76%	79%
DT	72%	79%	70%	86%	78%	83%	70%	63%	89%	47%	68%	61%
RF	81%	75%	70%	79%	74%	79%	74%	58%	56%	60%	58%	64%
NN	83%	71%	80%	64%	72%	72%	81%	67%	56%	73%	64%	73%
AdaB	74%	79%	80%	79%	79%	81%	81%	71%	78%	67%	72%	74%
NB	87%	75%	80%	71%	76%	77%	84%	75%	89%	67%	78%	77%
LDA	78%	83%	100%	71%	86%	83%	77%	58%	56%	60%	58%	64%
XGB	72%	75%	70%	79%	74%	79%	81%	75%	78%	73%	76%	79%

(a) Crosscorrelazione

Classif	Boruta + z-score						Boruta + minmax					
	Val	Test	Spec.	Sens.	AUC	F1	Val	Test	Spec.	Sens.	AUC	F1
KNN	93%	67%	67%	67%	67%	71%	85%	83%	75%	92%	83%	85%
LSVM	93%	75%	89%	67%	78%	77%	59%	50%	0%	100%	50%	67%
GPr	95%	71%	78%	67%	72%	74%	89%	79%	75%	83%	79%	80%
DT	80%	79%	56%	93%	74%	85%	83%	75%	50%	100%	75%	80%
RF	93%	67%	67%	67%	67%	71%	83%	75%	75%	75%	75%	75%
NN	95%	71%	78%	67%	72%	74%	85%	92%	92%	92%	92%	92%
AdaB	85%	75%	78%	73%	76%	79%	87%	63%	42%	83%	63%	69%
NB	89%	71%	56%	80%	68%	77%	77%	67%	33%	100%	67%	75%
LDA	87%	67%	78%	60%	69%	69%	79%	88%	83%	92%	88%	88%
XGB	83%	83%	67%	93%	80%	88%	93%	79%	58%	100%	79%	83%

(b) Boruta

Tabella 6.11: Risultati delle metriche di valutazione ottenuti su 10 Classificatori, vocalizzo /a/ unito alle 2 frasi migliori su Italiano1+Privato1

Classif	Cross-corr + z-score						Cross-corr + minmax					
	Val	Test	Spec.	Sens.	AUC	F1	Val	Test	Spec.	Sens.	AUC	F1
KNN	63%	47%	57%	38%	47%	43%	69%	67%	82%	58%	70%	69%
LSVM	69%	60%	79%	44%	61%	54%	54%	37%	100%	0%	50%	0%
GPr	69%	60%	71%	50%	61%	57%	71%	57%	64%	53%	58%	61%
DT	69%	43%	57%	31%	44%	37%	56%	67%	82%	58%	70%	69%
RF	67%	60%	57%	63%	60%	63%	67%	60%	91%	42%	67%	57%
NN	64%	60%	64%	56%	60%	60%	70%	60%	55%	63%	59%	67%
AdaB	67%	53%	50%	56%	53%	56%	64%	60%	64%	58%	61%	65%
NB	70%	60%	64%	56%	60%	60%	71%	53%	64%	47%	56%	56%
LDA	71%	60%	71%	50%	61%	57%	63%	53%	73%	42%	57%	53%
XGB	73%	60%	64%	56%	60%	60%	63%	63%	64%	63%	63%	69%

(a) Crosscorrelazione

Classif	Boruta + z-score						Boruta + minmax					
	Val	Test	Spec.	Sens.	AUC	F1	Val	Test	Spec.	Sens.	AUC	F1
KNN	73%	63%	77%	53%	65%	62%	77%	50%	40%	70%	55%	48%
LSVM	71%	63%	77%	53%	65%	62%	57%	33%	0%	100%	50%	50%
GPr	74%	60%	69%	53%	61%	60%	83%	47%	35%	70%	53%	47%
DT	79%	43%	46%	41%	44%	45%	67%	40%	35%	50%	43%	36%
RF	84%	63%	77%	53%	65%	62%	79%	50%	40%	70%	55%	48%
NN	76%	57%	69%	47%	58%	55%	80%	47%	35%	70%	53%	47%
AdaB	80%	53%	54%	53%	53%	56%	79%	50%	40%	70%	55%	48%
NB	70%	57%	77%	41%	59%	52%	74%	60%	55%	70%	63%	54%
LDA	73%	67%	77%	59%	68%	67%	74%	50%	40%	70%	55%	48%
XGB	83%	53%	62%	47%	54%	53%	83%	43%	30%	70%	50%	45%

(b) Boruta

Tabella 6.12: Risultati delle metriche di valutazione ottenuti su 10 Classificatori, vocalizzo /a/ su Colombiano

Classif	Cross-corr + z-score						Cross-corr + minmax					
	Val	Test	Spec.	Sens.	AUC	F1	Val	Test	Spec.	Sens.	AUC	F1
KNN	73%	70%	60%	90%	75%	67%	79%	87%	92%	83%	88%	88%
LSVM	77%	77%	70%	90%	80%	72%	64%	70%	100%	50%	75%	67%
GPr	83%	77%	70%	90%	80%	72%	84%	73%	75%	72%	74%	76%
DT	63%	70%	60%	90%	75%	67%	74%	53%	33%	67%	50%	63%
RF	71%	80%	75%	90%	83%	75%	67%	70%	67%	72%	69%	74%
NN	71%	80%	75%	90%	83%	75%	86%	73%	75%	72%	74%	76%
AdaB	74%	73%	65%	90%	78%	69%	81%	70%	58%	78%	68%	76%
NB	81%	83%	80%	90%	85%	78%	81%	77%	83%	72%	78%	79%
LDA	66%	60%	55%	70%	63%	54%	74%	60%	50%	67%	58%	67%
XGB	67%	73%	65%	90%	78%	69%	80%	67%	58%	72%	65%	72%

(a) Crosscorrelazione

Classif	Boruta + z-score						Boruta + minmax					
	Val	Test	Spec.	Sens.	AUC	F1	Val	Test	Spec.	Sens.	AUC	F1
KNN	81%	80%	67%	93%	80%	82%	86%	83%	94%	71%	83%	80%
LSVM	83%	90%	80%	100%	90%	91%	59%	47%	0%	100%	50%	64%
GPr	89%	93%	87%	100%	93%	94%	81%	83%	88%	79%	83%	81%
DT	73%	70%	60%	80%	70%	73%	76%	67%	56%	79%	67%	69%
RF	81%	77%	67%	87%	77%	79%	90%	87%	94%	79%	86%	85%
NN	87%	93%	87%	100%	93%	94%	80%	83%	88%	79%	83%	81%
AdaB	83%	77%	60%	93%	77%	80%	84%	80%	75%	86%	80%	80%
NB	83%	90%	80%	100%	90%	91%	84%	87%	94%	79%	86%	85%
LDA	89%	83%	67%	100%	83%	86%	76%	77%	75%	79%	77%	76%
XGB	84%	70%	53%	87%	70%	74%	83%	63%	56%	71%	64%	65%

(b) Boruta

Tabella 6.13: Risultati delle metriche di valutazione ottenuti su 10 Classificatori, frase migliore su Colombiano

Classif	Cross-corr + z-score .5						Cross-corr + minmax					
	Val	Test	Spec.	Sens.	AUC	F1	Val	Test	Spec.	Sens.	AUC	F1
KNN	87%	57%	67%	50%	58%	58%	81%	70%	57%	81%	69%	74%
LSVM	86%	60%	75%	50%	63%	60%	80%	70%	64%	75%	70%	73%
GPr	87%	57%	58%	56%	57%	61%	79%	70%	50%	88%	69%	76%
DT	79%	70%	67%	72%	69%	74%	70%	70%	64%	75%	70%	73%
RF	71%	60%	58%	61%	60%	65%	77%	63%	43%	81%	62%	70%
NN	89%	60%	58%	61%	60%	65%	83%	67%	50%	81%	66%	72%
AdaB	76%	70%	75%	67%	71%	73%	76%	67%	57%	75%	66%	71%
NB	86%	63%	67%	61%	64%	67%	80%	67%	50%	81%	66%	72%
LDA	70%	50%	67%	39%	53%	48%	76%	40%	36%	44%	40%	44%
XGB	76%	67%	75%	61%	68%	69%	71%	70%	64%	75%	70%	73%

(a) Crosscorrelazione

Classif	Boruta + z-score						Boruta + minmax					
	Val	Test	Spec.	Sens.	AUC	F1	Val	Test	Spec.	Sens.	AUC	F1
KNN	90%	77%	76%	77%	77%	74%	81%	80%	71%	92%	81%	80%
LSVM	89%	80%	76%	85%	81%	79%	53%	43%	0%	100%	50%	60%
GPr	91%	73%	71%	77%	74%	71%	86%	73%	65%	85%	75%	73%
DT	69%	63%	59%	69%	64%	62%	73%	70%	65%	77%	71%	69%
RF	83%	73%	76%	69%	73%	69%	80%	60%	53%	69%	61%	60%
NN	91%	73%	71%	77%	74%	71%	87%	77%	65%	92%	79%	77%
AdaB	87%	87%	76%	100%	88%	87%	84%	70%	65%	77%	71%	69%
NB	86%	77%	88%	62%	75%	70%	86%	80%	71%	92%	81%	80%
LDA	94%	70%	71%	69%	70%	67%	81%	70%	71%	69%	70%	67%
XGB	90%	73%	65%	85%	75%	73%	86%	73%	65%	85%	75%	73%

(b) Boruta

Tabella 6.14: Risultati delle metriche di valutazione ottenuti su 10 Classificatori, vocalizzo /a/ unito alla frase migliore su Colombiano

Classif	Cross-corr + z-score						Cross-corr + minmax					
	Val	Test	Spec.	Sens.	AUC	F1	Val	Test	Spec.	Sens.	AUC	F1
KNN	81%	73%	77%	71%	74%	75%	81%	80%	93%	67%	80%	77%
LSVM	84%	77%	77%	76%	77%	79%	87%	77%	80%	73%	77%	76%
GPr	71%	70%	77%	65%	71%	71%	83%	77%	87%	67%	77%	74%
DT	71%	77%	77%	76%	77%	79%	67%	73%	87%	60%	73%	69%
RF	76%	67%	54%	76%	65%	72%	79%	80%	93%	67%	80%	77%
NN	76%	73%	69%	76%	73%	76%	83%	77%	87%	67%	77%	74%
AdaB	81%	77%	77%	76%	77%	79%	74%	70%	87%	53%	70%	64%
NB	83%	80%	85%	76%	81%	81%	86%	77%	87%	67%	77%	74%
LDA	79%	57%	62%	53%	57%	58%	76%	63%	80%	47%	63%	56%
XGB	76%	73%	77%	71%	74%	75%	76%	73%	93%	53%	73%	67%

(a) Crosscorrelazione

Classif	Boruta + z-score						Boruta + minmax					
	Val	Test	Spec.	Sens.	AUC	F1	Val	Test	Spec.	Sens.	AUC	F1
KNN	79%	80%	82%	77%	80%	77%	86%	63%	53%	73%	63%	67%
LSVM	86%	77%	76%	77%	77%	74%	86%	73%	67%	80%	73%	75%
GPr	90%	73%	71%	77%	74%	71%	80%	63%	47%	80%	63%	69%
DT	84%	57%	53%	62%	57%	55%	74%	67%	67%	67%	67%	67%
RF	83%	80%	82%	77%	80%	77%	83%	67%	53%	80%	67%	71%
NN	90%	73%	71%	77%	74%	71%	84%	67%	53%	80%	67%	71%
AdaB	86%	67%	59%	77%	68%	67%	81%	60%	33%	87%	60%	68%
NB	80%	90%	88%	92%	90%	89%	84%	70%	60%	80%	70%	73%
LDA	90%	70%	71%	69%	70%	67%	81%	63%	47%	80%	63%	69%
XGB	86%	73%	65%	85%	75%	73%	79%	63%	47%	80%	63%	69%

(b) Boruta

Tabella 6.15: Risultati delle metriche di valutazione ottenuti su 10 Classificatori, vocalizzo /a/ unito alle 2 frasi migliori su Colombiano

Classif	Cross-corr + z-score						Cross-corr + minmax					
	Val	Test	Spec.	Sens.	AUC	F1	Val	Test	Spec.	Sens.	AUC	F1
KNN	75%	87%	92%	83%	88%	88%	83%	65%	75%	53%	64%	59%
LSVM	76%	87%	100%	78%	89%	88%	69%	61%	88%	33%	60%	45%
GPr	75%	84%	100%	72%	86%	84%	82%	68%	81%	53%	67%	62%
DT	68%	71%	100%	50%	75%	67%	78%	74%	81%	67%	74%	71%
RF	79%	87%	100%	78%	89%	88%	79%	81%	94%	67%	80%	77%
NN	76%	81%	92%	72%	82%	81%	82%	68%	81%	53%	67%	62%
AdaB	77%	94%	100%	89%	94%	94%	78%	71%	81%	60%	71%	67%
NB	78%	84%	92%	78%	85%	85%	79%	61%	63%	60%	61%	60%
LDA	67%	84%	92%	78%	85%	85%	61%	74%	88%	60%	74%	69%
XGB	85%	90%	100%	83%	92%	91%	83%	81%	88%	73%	80%	79%

(a) Crosscorrelazione

Classif	Boruta + z-score						Boruta + minmax					
	Val	Test	Spec.	Sens.	AUC	F1	Val	Test	Spec.	Sens.	AUC	F1
KNN	81%	74%	63%	92%	77%	73%	88%	71%	80%	63%	71%	69%
LSVM	78%	81%	89%	67%	78%	73%	58%	52%	100%	6%	53%	12%
GPr	84%	74%	63%	92%	77%	73%	85%	65%	80%	50%	65%	59%
DT	84%	74%	58%	100%	79%	75%	85%	58%	87%	31%	59%	43%
RF	84%	71%	53%	100%	76%	73%	88%	68%	80%	56%	68%	64%
NN	82%	71%	63%	83%	73%	69%	80%	74%	93%	56%	75%	69%
AdaB	82%	74%	58%	100%	79%	75%	89%	58%	80%	38%	59%	48%
NB	79%	84%	89%	75%	82%	78%	79%	71%	93%	50%	72%	64%
LDA	84%	77%	74%	83%	79%	74%	83%	77%	87%	69%	78%	76%
XGB	84%	74%	58%	100%	79%	75%	86%	58%	87%	31%	59%	43%

(b) Boruta

Tabella 6.16: Risultati delle metriche di valutazione ottenuti su 10 Classificatori, vocalizzo /e/ su Italiano 2

Classif	Cross-corr + z-score						Cross-corr + minmax					
	Val	Test	Spec.	Sens.	AUC	F1	Val	Test	Spec.	Sens.	AUC	F1
KNN	82%	71%	100%	47%	74%	64%	85%	81%	93%	71%	82%	80%
LSVM	79%	74%	100%	53%	76%	69%	60%	45%	100%	0%	50%	0%
GPr	69%	84%	100%	71%	85%	83%	83%	77%	71%	82%	77%	80%
DT	59%	68%	71%	65%	68%	69%	65%	65%	64%	65%	64%	67%
RF	79%	65%	86%	47%	66%	59%	72%	48%	64%	35%	50%	43%
NN	74%	84%	100%	71%	85%	83%	83%	74%	71%	76%	74%	76%
AdaB	72%	74%	93%	59%	76%	71%	72%	74%	64%	82%	73%	78%
NB	85%	68%	86%	53%	69%	64%	82%	74%	71%	76%	74%	76%
LDA	51%	74%	86%	65%	75%	73%	73%	65%	71%	59%	65%	65%
XGB	67%	68%	86%	53%	69%	64%	75%	77%	71%	82%	77%	80%

(a) Crosscorrelazione

Classif	Boruta + z-score						Boruta + minmax					
	Val	Test	Spec.	Sens.	AUC	F1	Val	Test	Spec.	Sens.	AUC	F1
KNN	87%	81%	79%	83%	81%	77%	85%	81%	80%	82%	81%	75%
LSVM	89%	87%	84%	92%	88%	85%	53%	71%	100%	18%	59%	31%
GPr	89%	77%	74%	83%	79%	74%	88%	77%	75%	82%	78%	72%
DT	72%	71%	74%	67%	70%	64%	72%	71%	70%	73%	71%	64%
RF	88%	81%	79%	83%	81%	77%	82%	74%	75%	73%	74%	67%
NN	89%	81%	79%	83%	81%	77%	88%	84%	85%	82%	83%	78%
AdaB	79%	84%	79%	92%	85%	81%	84%	71%	75%	64%	69%	61%
NB	87%	87%	84%	92%	88%	85%	92%	74%	70%	82%	76%	69%
LDA	89%	81%	74%	92%	83%	79%	82%	74%	75%	73%	74%	67%
XGB	76%	81%	79%	83%	81%	77%	81%	77%	75%	82%	78%	72%

(b) Boruta

Tabella 6.17: Risultati delle metriche di valutazione ottenuti su 10 Classificatori, frase migliore su Italiano 2

Classif	Cross-corr + z-score .5						Cross-corr + minmax					
	Val	Test	Spec.	Sens.	AUC	F1	Val	Test	Spec.	Sens.	AUC	F1
KNN	89%	81%	76%	86%	81%	80%	85%	77%	83%	69%	76%	72%
LSVM	92%	71%	65%	79%	72%	71%	84%	81%	94%	62%	78%	73%
GPr	88%	77%	71%	86%	78%	77%	85%	81%	78%	85%	81%	79%
DT	83%	77%	76%	79%	78%	76%	83%	81%	83%	77%	80%	77%
RF	78%	81%	76%	86%	81%	80%	82%	74%	89%	54%	71%	64%
NN	92%	74%	65%	86%	75%	75%	85%	77%	72%	85%	78%	76%
AdaB	82%	84%	76%	93%	85%	84%	85%	87%	83%	92%	88%	86%
NB	92%	71%	59%	86%	72%	73%	86%	77%	72%	85%	78%	76%
LDA	81%	71%	76%	64%	70%	67%	72%	74%	72%	77%	75%	71%
XGB	89%	71%	76%	64%	70%	67%	89%	90%	89%	92%	91%	89%

(a) Crosscorrelazione

Classif	Boruta + z-score						Boruta + minmax					
	Val	Test	Spec.	Sens.	AUC	F1	Val	Test	Spec.	Sens.	AUC	F1
KNN	85%	87%	93%	82%	88%	88%	92%	81%	79%	82%	80%	82%
LSVM	79%	87%	93%	82%	88%	88%	60%	55%	100%	18%	59%	30%
GPr	79%	94%	93%	94%	93%	94%	94%	74%	71%	76%	74%	76%
DT	80%	71%	71%	71%	71%	73%	86%	81%	86%	76%	81%	81%
RF	89%	90%	86%	94%	90%	91%	92%	84%	79%	88%	83%	86%
NN	80%	90%	93%	88%	91%	91%	89%	77%	71%	82%	77%	80%
AdaB	82%	90%	86%	94%	90%	91%	92%	77%	79%	76%	78%	79%
NB	80%	81%	93%	71%	82%	80%	88%	81%	79%	82%	80%	82%
LDA	79%	87%	93%	82%	88%	88%	89%	71%	71%	71%	71%	73%
XGB	89%	97%	93%	100%	96%	97%	90%	84%	79%	88%	83%	86%

(b) Boruta

Tabella 6.18: Risultati delle metriche di valutazione ottenuti su 10 Classificatori, vocalizzo /e/ unito alla frase migliore su Italiano 2

Classif	Cross-corr + z-score						Cross-corr + minmax					
	Val	Test	Spec.	Sens.	AUC	F1	Val	Test	Spec.	Sens.	AUC	F1
KNN	89%	90%	88%	93%	90%	90%	90%	81%	93%	69%	81%	79%
LSVM	88%	74%	88%	60%	74%	69%	90%	81%	87%	75%	81%	80%
GPr	83%	84%	88%	80%	84%	83%	93%	81%	87%	75%	81%	80%
DT	80%	71%	75%	67%	71%	69%	81%	81%	87%	75%	81%	80%
RF	86%	71%	75%	67%	71%	69%	79%	71%	87%	56%	71%	67%
NN	90%	77%	88%	67%	77%	74%	92%	84%	87%	81%	84%	84%
AdaB	87%	77%	81%	73%	77%	76%	88%	90%	93%	88%	90%	90%
NB	88%	71%	69%	73%	71%	71%	90%	77%	80%	75%	78%	77%
LDA	86%	81%	81%	80%	81%	80%	89%	68%	87%	50%	68%	62%
XGB	83%	87%	94%	80%	87%	86%	86%	84%	87%	81%	84%	84%

(a) Crosscorrelazione

Classif	Boruta + z-score						Boruta + minmax					
	Val	Test	Spec.	Sens.	AUC	F1	Val	Test	Spec.	Sens.	AUC	F1
KNN	86%	87%	94%	79%	86%	85%	82%	97%	100%	92%	96%	96%
LSVM	89%	87%	94%	79%	86%	85%	56%	71%	100%	31%	65%	47%
GPr	92%	81%	82%	79%	80%	79%	90%	90%	94%	85%	90%	88%
DT	76%	68%	59%	79%	69%	69%	82%	94%	94%	92%	93%	92%
RF	85%	90%	100%	79%	89%	88%	92%	97%	100%	92%	96%	96%
NN	93%	77%	76%	79%	78%	76%	93%	90%	94%	85%	90%	88%
AdaB	85%	87%	88%	86%	87%	86%	92%	94%	100%	85%	92%	92%
NB	88%	84%	88%	79%	83%	81%	89%	90%	100%	77%	88%	87%
LDA	92%	77%	82%	71%	77%	74%	91%	81%	89%	69%	79%	75%
XGB	80%	87%	82%	93%	88%	87%	92%	97%	100%	92%	96%	96%

(b) Boruta

Tabella 6.19: Risultati delle metriche di valutazione ottenuti su 10 Classificatori, vocalizzo /e/ unito alle 2 frasi migliori su Italiano 2

6.5 Machine learning per l'analisi multilingua

In questa sezione si riportano i risultati e le feature (Tabella 6.21 e Tabella 6.20) dell'applicazione dell'algoritmo di classificazione al dataset formato dall'unione di Italiano 1+ Privato 1+ Colombiano + Ceco + Privato 2 sul vocalizzo /a/. In grassetto sono evidenziati le prestazioni migliori.

BORUTA_ZSCORE	BORUTA_MINMAX
SPEC_CENT_GRAVITY_mean_a	SPEC_CENT_GRAVITY_mean_a
STE_median_a	FORM1_std_a
SPEC_FLUX_median_a	
MFCC7_DD_median_a	
PLP0_median_a	
FORM1_std_a	

Tabella 6.20: feature selezionate dall'algoritmo Boruta applicato al vocalizzo /a/ dataset unito multilingua

Classif	Boruta + z-score						Boruta + minmax					
	Val	Test	Spec.	Sens.	AUC	F1	Val	Test	Spec.	Sens.	AUC	F1
KNN	68%	61%	54%	71%	65%	67%	72%	65%	74%	55%	60%	65%
LSVM	74%	51%	37%	71%	65%	65%	53%	49%	0%	100%	50%	71%
GPr	76%	62%	54%	74%	69%	69%	72%	68%	82%	53%	65%	67%
DT	71%	62%	63%	61%	59%	62%	71%	65%	69%	61%	68%	63%
RF	72%	58%	57%	61%	70%	71%	73%	65%	69%	61%	63%	68%
NN	73%	61%	52%	74%	66%	68%	73%	69%	79%	58%	64%	65%
AdaB	76%	61%	59%	65%	61%	62%	75%	66%	77%	55%	60%	63%
NB	69%	52%	46%	61%	67%	64%	53%	51%	3%	100%	62%	58%
LDA	75%	57%	48%	71%	65%	64%	74%	65%	79%	50%	65%	66%
XGB	72%	65%	63%	68%	68%	69%	68%	62%	72%	53%	62%	64%

Tabella 6.21: Risultati delle metriche di valutazione ottenuti su 10 Classificatori, vocalizzo /a/ del dataset unito multilingua

6.6 Analisi statistica

6.6.1 Analisi su ogni dataset

La Tabella 6.22 riporta i risultati del test di Mann-Whitney U e la variazione dei parametri statistici dei pazienti affetti da MP rispetto ai CS. Settando i valori di $p < 0,05$ a 1 e i valori di $p > 0,05$ a 0, si è calcolato il numero di volte in cui una feature si attiva (1) sui dataset (ovvero, risulta essere significativa). In grassetto sono state evidenziate le somme più alte.

Tabella 6.22: feature con valore di p riferito alla classe ottenuto con il test di Mann-Whitney, per i 5 dataset

feature	ITA.1	PRIV.1	COL.	CECO	PRIV.2	SOM.
JIT_L_a	0	1	1	0	1	3
JIT_LAb_a	0	0	1	0	0	1
JIT_RAP_a	0	0	1	0	0	1
JIT_DDP_a	0	0	1	0	0	1
JIT_PPQ5_a	0	0	1	0	0	1
SHIM_a	1	0	1	0	0	2
SHIM_DB_a	1	0	1	0	0	2
SHIM_APQ3_a	1	0	1	0	0	2
SHIM_APQ5_a	1	0	1	0	0	2
SHIM_APQ11_a	1	0	1	0	0	2
INT_SD_a	0	0	0	0	1	1
INT_AVG_a	1	0	0	1	0	2
DFA_COEFF_a	0	0	0	0	0	0
F0_mean_a	0	0	0	0	0	0
HNR_mean_a	1	0	1	0	1	3
STE_mean_a	1	0	0	1	0	2
CPP_mean_a	1	0	1	0	0	2
GNE_mean_a	0	0	1	0	0	1
SPEC_CENT_GRAVITY_mean_a	1	1	0	1	0	3
SPEC_STD_mean_a	0	0	0	0	0	0
SPEC_SKEWNESS_mean_a	1	0	0	0	0	1
SPEC_KURTOSIS_mean_a	1	0	0	0	0	1
SPEC_FLUX_mean_a	1	0	0	0	0	1
SPEC_SLOPE_mean_a	0	0	0	0	0	0
SPEC_DECREASE_mean_a	0	0	0	0	0	0
SPEC_ROLLOFF_mean_a	1	0	0	0	0	1
FORM1_mean_a	0	0	1	1	0	2
FORM2_mean_a	0	0	0	0	0	0
FORM3_mean_a	0	0	0	0	1	1
FORM_BAND1_mean_a	0	0	1	0	1	2
FORM_BAND2_mean_a	0	0	0	0	0	0
FORM_BAND3_mean_a	0	0	0	0	0	0
MFCC1_mean_a	1	0	0	0	0	1
MFCC2_mean_a	1	1	0	1	0	3
MFCC3_mean_a	1	0	1	0	1	3
MFCC4_mean_a	1	0	0	1	0	2
MFCC5_mean_a	0	0	0	0	0	0
MFCC6_mean_a	0	0	0	0	0	0
MFCC7_mean_a	1	0	0	0	0	1

(Continua alla pagina successiva)

(Continua dalla pagina precedente)

MFCC8_mean_a	0	0	0	0	0	0
MFCC9_mean_a	0	0	0	0	0	0
MFCC10_mean_a	0	0	0	0	0	0
MFCC11_mean_a	0	0	0	0	0	0
MFCC12_mean_a	0	0	0	0	0	0
MFCC13_mean_a	0	0	0	0	0	0
MFCC1_D_mean_a	0	1	0	0	0	1
MFCC1_DD_mean_a	0	0	0	0	0	0
MFCC2_D_mean_a	0	0	1	0	0	1
MFCC2_DD_mean_a	0	0	0	0	0	0
MFCC3_D_mean_a	0	0	0	0	0	0
MFCC3_DD_mean_a	0	0	0	0	0	0
MFCC4_D_mean_a	1	0	0	0	0	1
MFCC4_DD_mean_a	0	0	0	0	0	0
MFCC5_D_mean_a	0	0	1	0	0	1
MFCC5_DD_mean_a	0	0	0	0	0	0
MFCC6_D_mean_a	0	0	0	0	0	0
MFCC6_DD_mean_a	0	0	0	0	0	0
MFCC7_D_mean_a	0	0	0	0	0	0
MFCC7_DD_mean_a	0	1	0	0	0	1
MFCC8_D_mean_a	1	0	1	0	0	2
MFCC8_DD_mean_a	0	0	0	0	0	0
MFCC9_D_mean_a	1	0	0	0	0	1
MFCC9_DD_mean_a	0	0	0	1	0	1
MFCC10_D_mean_a	0	0	0	0	0	0
MFCC10_DD_mean_a	0	0	0	0	0	0
MFCC11_D_mean_a	1	0	0	0	0	1
MFCC11_DD_mean_a	0	0	0	0	0	0
MFCC12_D_mean_a	0	0	0	0	0	0
MFCC12_DD_mean_a	0	0	0	0	0	0
MFCC13_D_mean_a	0	1	0	0	0	1
MFCC13_DD_mean_a	0	0	0	0	0	0
PLP0_mean_a	1	0	0	0	1	2
PLP1_mean_a	0	0	1	0	0	1
PLP2_mean_a	1	0	0	1	1	3
PLP3_mean_a	0	0	0	0	0	0
PLP4_mean_a	0	0	0	1	0	1
PLP5_mean_a	0	0	1	0	0	1
PLP6_mean_a	0	0	0	1	0	1
PLP7_mean_a	1	0	0	0	0	1
PLP8_mean_a	0	0	0	0	0	0
PLP9_mean_a	0	1	1	0	0	2
PLP10_mean_a	1	0	1	0	0	2

(Continua alla pagina successiva)

(Continua dalla pagina precedente)

PLP11_mean_a	0	0	0	0	0	0
PLP12_mean_a	0	0	0	0	1	1
PLP0_D_mean_a	1	0	0	1	0	2
PLP1_D_mean_a	1	0	0	1	1	3
PLP2_D_mean_a	0	0	0	0	0	0
PLP3_D_mean_a	1	0	0	0	0	1
PLP4_D_mean_a	0	0	0	0	0	0
PLP5_D_mean_a	1	0	0	0	0	1
PLP6_D_mean_a	0	0	0	0	0	0
PLP7_D_mean_a	0	0	0	1	1	2
PLP8_D_mean_a	1	0	0	0	0	1
PLP9_D_mean_a	0	0	0	0	0	0
PLP10_D_mean_a	0	1	0	0	1	2
PLP11_D_mean_a	0	0	0	0	0	0
PLP12_D_mean_a	0	0	0	0	0	0
PLP0_DD_mean_a	1	0	0	0	1	2
PLP1_DD_mean_a	0	1	0	0	1	2
PLP2_DD_mean_a	0	0	1	0	0	1
PLP3_DD_mean_a	0	0	0	0	0	0
PLP4_DD_mean_a	0	0	0	0	0	0
PLP5_DD_mean_a	0	0	0	0	0	0
PLP6_DD_mean_a	0	0	0	0	0	0
PLP7_DD_mean_a	0	0	0	0	1	1
PLP8_DD_mean_a	1	0	0	0	0	1
PLP9_DD_mean_a	0	0	0	0	0	0
PLP10_DD_mean_a	0	0	0	0	0	0
PLP11_DD_mean_a	0	0	0	0	0	0
PLP12_DD_mean_a	0	0	0	0	0	0
F0_median_a	0	0	0	0	0	0
HNR_median_a	1	0	1	0	0	2
STE_median_a	1	0	0	1	0	2
CPP_median_a	1	0	1	0	0	2
GNE_median_a	1	0	1	0	0	2
SPEC_CENT_GRAVITY_median_a	1	0	0	1	0	2
SPEC_STD_median_a	0	0	0	0	0	0
SPEC_SKEWNESS_median_a	1	0	0	0	0	1
SPEC_KURTOSIS_median_a	1	0	0	0	0	1
SPEC_FLUX_median_a	1	0	0	1	0	2
SPEC_SLOPE_median_a	0	1	0	0	0	1
SPEC_DECREASE_median_a	0	0	0	0	0	0
SPEC_ROLLOFF_median_a	0	0	0	0	0	0
FORM1_median_a	0	0	0	1	0	1
FORM2_median_a	0	0	0	0	0	0

(Continua alla pagina successiva)

(Continua dalla pagina precedente)

FORM3_median_a	0	0	0	0	1	1
FORM_BAND1_median_a	0	0	1	0	1	2
FORM_BAND2_median_a	0	0	0	0	0	0
FORM_BAND3_median_a	0	0	0	0	0	0
MFCC1_median_a	1	0	0	0	0	1
MFCC2_median_a	1	1	0	1	0	3
MFCC3_median_a	1	0	1	0	0	2
MFCC4_median_a	1	0	0	1	0	2
MFCC5_median_a	0	0	0	0	0	0
MFCC6_median_a	0	0	0	0	0	0
MFCC7_median_a	1	0	0	0	0	1
MFCC8_median_a	0	0	0	0	0	0
MFCC9_median_a	0	0	0	0	0	0
MFCC10_median_a	0	0	0	0	0	0
MFCC11_median_a	0	0	0	0	0	0
MFCC12_median_a	0	0	0	0	0	0
MFCC13_median_a	0	0	0	0	0	0
MFCC1_D_median_a	0	0	0	0	0	0
MFCC1_DD_median_a	0	0	0	0	0	0
MFCC2_D_median_a	0	0	0	0	0	0
MFCC2_DD_median_a	0	0	0	0	0	0
MFCC3_D_median_a	0	0	0	0	0	0
MFCC3_DD_median_a	0	0	0	0	0	0
MFCC4_D_median_a	0	0	0	0	0	0
MFCC4_DD_median_a	0	0	0	0	0	0
MFCC5_D_median_a	0	1	0	0	0	1
MFCC5_DD_median_a	1	0	0	0	0	1
MFCC6_D_median_a	0	0	0	0	0	0
MFCC6_DD_median_a	1	0	0	0	0	1
MFCC7_D_median_a	0	0	0	0	0	0
MFCC7_DD_median_a	0	1	0	0	0	1
MFCC8_D_median_a	0	0	0	0	0	0
MFCC8_DD_median_a	0	0	0	0	0	0
MFCC9_D_median_a	0	0	0	0	0	0
MFCC9_DD_median_a	0	0	0	0	0	0
MFCC10_D_median_a	0	0	0	0	0	0
MFCC10_DD_median_a	0	0	0	0	0	0
MFCC11_D_median_a	0	0	0	1	0	1
MFCC11_DD_median_a	0	0	0	0	0	0
MFCC12_D_median_a	0	0	0	1	0	1
MFCC12_DD_median_a	0	0	0	0	0	0
MFCC13_D_median_a	0	0	0	0	0	0
MFCC7_D_median_a	0	0	0	0	0	0

(Continua alla pagina successiva)

(Continua dalla pagina precedente)

MFCC7_DD_median_a	0	1	0	0	0	1
MFCC8_D_median_a	0	0	0	0	0	0
MFCC8_DD_median_a	0	0	0	0	0	0
MFCC9_D_median_a	0	0	0	0	0	0
MFCC9_DD_median_a	0	0	0	0	0	0
MFCC10_D_median_a	0	0	0	0	0	0
MFCC10_DD_median_a	0	0	0	0	0	0
MFCC11_D_median_a	0	0	0	1	0	1
MFCC11_DD_median_a	0	0	0	0	0	0
MFCC12_D_median_a	0	0	0	1	0	1
MFCC12_DD_median_a	0	0	0	0	0	0
MFCC13_D_median_a	0	0	0	0	0	0
MFCC13_DD_median_a	0	0	0	0	0	0
PLP0_median_a	1	0	1	0	0	2
PLP1_median_a	0	0	1	0	0	1
PLP2_median_a	0	1	0	0	0	1
PLP3_median_a	0	0	0	1	0	1
PLP4_median_a	0	0	0	0	0	0
PLP5_median_a	0	0	1	0	0	1
PLP6_median_a	0	0	1	1	0	2
PLP7_median_a	0	0	0	0	0	0
PLP8_median_a	0	1	0	1	0	2
PLP9_median_a	0	0	0	0	0	0
PLP10_median_a	0	0	1	0	0	1
PLP11_median_a	0	0	0	0	0	0
PLP12_median_a	0	0	0	0	0	0
PLP0_D_median_a	0	0	0	1	1	2
PLP1_D_median_a	1	0	0	0	1	2
PLP2_D_median_a	0	0	0	0	0	0
PLP3_D_median_a	0	0	0	0	0	0
PLP4_D_median_a	0	0	0	0	1	1
PLP5_D_median_a	0	0	0	0	0	0
PLP6_D_median_a	0	0	0	0	0	0
PLP7_D_median_a	0	0	0	0	0	0
PLP8_D_median_a	0	0	1	0	0	1
PLP9_D_median_a	0	0	0	0	0	0
PLP10_D_median_a	0	0	0	0	0	0
PLP11_D_median_a	0	0	0	0	0	0
PLP12_D_median_a	1	0	0	0	0	1
F0_median_a	0	0	0	0	0	0
HNR_median_a	1	0	1	0	0	2
STE_median_a	1	0	0	1	0	2
CPP_median_a	1	0	1	0	0	2

(Continua alla pagina successiva)

(Continua dalla pagina precedente)

GNE_median_a	1	0	1	0	0	2
SPEC_CENT_GRAVITY_median_a	1	0	0	1	0	2
SPEC_STD_median_a	0	0	0	0	0	0
SPEC_SKEWNESS_median_a	1	0	0	0	0	1
SPEC_KURTOSIS_median_a	1	0	0	0	0	1
SPEC_FLUX_median_a	1	0	0	1	0	2
SPEC_SLOPE_median_a	0	1	0	0	0	1
SPEC_DECREASE_median_a	0	0	0	0	0	0
SPEC_ROLLOFF_median_a	0	0	0	0	0	0
PLP0_DD_median_a	1	0	0	0	1	2
PLP1_DD_median_a	0	0	0	0	0	0
PLP2_DD_median_a	0	0	0	0	0	0
PLP3_DD_median_a	0	0	0	0	0	0
PLP4_DD_median_a	1	0	0	0	0	1
PLP5_DD_median_a	0	0	0	0	0	0
PLP6_DD_median_a	0	0	0	0	0	0
PLP7_DD_median_a	0	0	0	0	0	0
PLP8_DD_median_a	0	0	0	0	0	0
PLP9_DD_median_a	0	0	0	0	0	0
PLP10_DD_median_a	0	0	0	0	0	0
PLP11_DD_median_a	1	0	0	0	0	1
PLP12_DD_median_a	0	0	0	0	0	0
F0_std_a	0	0	1	0	1	2
HNR_std_a	1	0	1	0	1	3
STE_std_a	1	0	0	0	0	1
CPP_std_a	0	0	0	1	0	1
GNE_std_a	0	0	1	0	1	2
SPEC_CENT_GRAVITY_std_a	1	0	1	0	1	3
SPEC_STD_std_a	1	0	1	0	0	2
SPEC_SKEWNESS_std_a	0	0	0	0	1	1
SPEC_KURTOSIS_std_a	1	0	0	0	0	1
SPEC_FLUX_std_a	1	0	0	0	0	1
SPEC_SLOPE_std_a	1	0	0	1	0	2
SPEC_DECREASE_std_a	0	0	0	0	0	0
SPEC_ROLLOFF_std_a	0	0	0	1	0	1
FORM1_std_a	0	0	1	1	1	3
FORM2_std_a	0	0	0	1	1	2
FORM3_std_a	0	0	0	0	1	1
FORM_BAND1_std_a	1	0	1	0	1	3
FORM_BAND2_std_a	1	0	0	0	1	2
FORM_BAND3_std_a	1	0	0	0	1	2
MFCC1_std_a	0	0	0	0	1	1
MFCC2_std_a	0	0	1	0	1	2

(Continua alla pagina successiva)

(Continua dalla pagina precedente)

MFCC3_std_a	1	0	0	0	1	2
MFCC4_std_a	0	0	1	0	1	2
MFCC5_std_a	0	0	0	0	0	0
MFCC6_std_a	0	0	0	0	1	1
MFCC7_std_a	0	0	0	0	0	0
MFCC8_std_a	0	0	1	0	0	1
MFCC9_std_a	0	0	0	0	0	0
MFCC10_std_a	0	0	1	0	0	1
MFCC11_std_a	0	0	0	0	0	0
MFCC12_std_a	0	0	0	0	0	0
MFCC13_std_a	0	0	0	0	0	0
MFCC1_D_std_a	1	0	0	0	0	1
MFCC1_DD_std_a	1	0	1	0	0	2
MFCC2_D_std_a	0	0	1	0	0	1
MFCC2_DD_std_a	0	0	1	0	0	1
MFCC3_D_std_a	1	0	1	0	0	2
MFCC3_DD_std_a	1	0	1	0	0	2
MFCC4_D_std_a	0	0	1	0	0	1
MFCC4_DD_std_a	0	0	1	0	0	1
MFCC5_D_std_a	0	0	1	0	0	1
MFCC5_DD_std_a	0	0	1	0	0	1
MFCC6_D_std_a	0	0	1	0	0	1
MFCC6_DD_std_a	0	0	1	0	0	1
MFCC7_D_std_a	0	0	1	0	0	1
MFCC7_DD_std_a	0	0	1	0	0	1
MFCC8_D_std_a	0	0	1	0	0	1
MFCC8_DD_std_a	0	0	1	0	0	1
MFCC9_D_std_a	0	0	1	0	0	1
MFCC9_DD_std_a	0	0	1	0	0	1
MFCC10_D_std_a	0	0	1	0	0	1
MFCC10_DD_std_a	0	0	1	0	0	1
MFCC11_D_std_a	0	0	1	0	0	1
MFCC11_DD_std_a	0	0	1	0	0	1
MFCC12_D_std_a	0	0	1	0	0	1
MFCC12_DD_std_a	0	0	1	0	0	1
MFCC13_D_std_a	0	0	1	0	0	1
MFCC13_DD_std_a	0	0	1	0	0	1
PLP0_std_a	1	0	0	0	1	2
PLP1_std_a	0	0	0	0	1	1
PLP2_std_a	1	0	0	0	1	2
PLP3_std_a	1	0	0	0	0	1
PLP4_std_a	1	0	1	0	0	2
PLP5_std_a	1	0	0	0	0	1

(Continua alla pagina successiva)

(Continua dalla pagina precedente)

PLP6_std_a	1	0	0	0	0	1
PLP7_std_a	1	0	0	0	1	2
PLP8_std_a	1	0	1	0	0	2
PLP9_std_a	1	0	1	0	0	2
PLP10_std_a	1	0	0	0	1	2
PLP11_std_a	1	0	0	0	0	1
PLP12_std_a	1	0	0	0	0	1
PLP0_D_std_a	1	0	0	0	1	2
PLP1_D_std_a	1	0	0	1	1	3
PLP2_D_std_a	1	0	0	0	1	2
PLP3_D_std_a	1	0	0	0	0	1
PLP4_D_std_a	1	0	1	0	0	2
PLP5_D_std_a	1	0	1	0	0	2
PLP6_D_std_a	1	0	1	0	0	2
PLP7_D_std_a	1	0	0	0	0	1
PLP8_D_std_a	1	0	1	0	0	2
PLP9_D_std_a	1	0	1	0	0	2
PLP10_D_std_a	1	0	0	0	1	2
PLP11_D_std_a	1	0	1	0	0	2
PLP12_D_std_a	1	0	0	0	0	1
PLP0_DD_std_a	1	0	0	0	1	2
PLP1_DD_std_a	1	0	1	1	0	3
PLP2_DD_std_a	1	0	1	0	1	3
PLP3_DD_std_a	1	0	0	0	0	1
PLP4_DD_std_a	1	0	1	0	0	2
PLP5_DD_std_a	1	0	1	0	0	2
PLP6_DD_std_a	1	0	1	0	0	2
PLP7_DD_std_a	1	0	0	0	0	1
PLP8_DD_std_a	1	0	1	0	0	2
PLP9_DD_std_a	1	0	1	0	0	2
PLP10_DD_std_a	1	0	0	0	1	2
PLP11_DD_std_a	1	0	1	0	0	2
PLP12_DD_std_a	1	0	1	0	0	2
F0_kurtosis_a	0	0	0	0	0	0
HNR_kurtosis_a	1	0	0	0	1	2
STE_kurtosis_a	0	0	1	1	1	3
CPP_kurtosis_a	0	0	0	1	0	1
GNE_kurtosis_a	1	0	1	0	0	2
SPEC_CENT_GRAVITY_kurtosis_a	0	0	0	0	0	0
SPEC_STD_kurtosis_a	1	0	0	0	1	2
SPEC_SKEWNESS_kurtosis_a	1	0	0	0	1	2
SPEC_KURTOSIS_kurtosis_a	0	0	0	0	1	1
SPEC_FLUX_kurtosis_a	0	0	1	1	1	3

(Continua alla pagina successiva)

(Continua dalla pagina precedente)

SPEC_SLOPE_kurtosis_a	0	1	0	0	0	1
SPEC_DECREASE_kurtosis_a	0	0	0	0	0	0
SPEC_ROLLOFF_kurtosis_a	0	0	1	1	1	3
FORM1_kurtosis_a	0	1	0	0	1	2
FORM2_kurtosis_a	0	0	0	0	1	1
FORM3_kurtosis_a	0	0	0	0	1	1
FORM_BAND1_kurtosis_a	0	0	0	0	1	1
FORM_BAND2_kurtosis_a	0	0	1	0	1	2
FORM_BAND3_kurtosis_a	0	0	1	0	1	2
MFCC1_kurtosis_a	1	0	0	0	1	2
MFCC2_kurtosis_a	0	0	0	0	0	0
MFCC3_kurtosis_a	0	0	0	0	0	0
MFCC4_kurtosis_a	1	0	0	0	0	1
MFCC5_kurtosis_a	0	0	0	0	1	1
MFCC6_kurtosis_a	0	0	0	0	1	1
MFCC7_kurtosis_a	0	0	0	0	0	0
MFCC8_kurtosis_a	0	0	0	0	1	1
MFCC9_kurtosis_a	0	0	0	0	1	1
MFCC10_kurtosis_a	0	0	0	0	1	1
MFCC11_kurtosis_a	0	0	0	0	0	0
MFCC12_kurtosis_a	0	0	0	0	0	0
MFCC13_kurtosis_a	0	0	0	0	1	1
MFCC1_D_kurtosis_a	0	0	0	0	1	1
MFCC1_DD_kurtosis_a	0	0	0	0	1	1
MFCC2_D_kurtosis_a	0	0	1	0	1	2
MFCC2_DD_kurtosis_a	0	0	0	0	0	0
MFCC3_D_kurtosis_a	0	0	1	0	0	1
MFCC3_DD_kurtosis_a	0	0	0	0	1	1
MFCC4_D_kurtosis_a	0	0	1	0	0	1
MFCC4_DD_kurtosis_a	0	0	0	0	0	0
MFCC5_D_kurtosis_a	0	0	0	0	1	1
MFCC5_DD_kurtosis_a	0	0	0	0	1	1
MFCC6_D_kurtosis_a	0	0	0	0	0	0
MFCC6_DD_kurtosis_a	0	0	0	0	1	1
MFCC7_D_kurtosis_a	0	0	0	0	0	0
MFCC7_DD_kurtosis_a	0	0	0	1	0	1
MFCC8_D_kurtosis_a	0	0	0	0	0	0
MFCC8_DD_kurtosis_a	0	0	0	1	0	1
MFCC9_D_kurtosis_a	0	0	0	0	1	1
MFCC9_DD_kurtosis_a	0	0	0	0	1	1
MFCC10_D_kurtosis_a	0	0	0	0	1	1
MFCC10_DD_kurtosis_a	0	0	0	0	1	1
MFCC11_D_kurtosis_a	0	0	0	0	1	1

(Continua alla pagina successiva)

(Continua dalla pagina precedente)

MFCC11_DD_kurtosis_a	0	0	0	0	1	1
MFCC12_D_kurtosis_a	0	0	0	0	1	1
MFCC12_DD_kurtosis_a	0	0	0	0	1	1
MFCC13_D_kurtosis_a	0	0	0	0	1	1
MFCC13_DD_kurtosis_a	0	0	0	0	1	1
PLP0_kurtosis_a	0	0	0	0	0	0
PLP1_kurtosis_a	1	0	0	0	1	2
PLP2_kurtosis_a	0	0	0	0	1	1
PLP3_kurtosis_a	0	0	0	0	0	0
PLP4_kurtosis_a	1	0	0	0	1	2
PLP5_kurtosis_a	0	0	1	0	1	2
PLP6_kurtosis_a	0	0	0	0	0	0
PLP7_kurtosis_a	0	0	0	0	1	1
PLP8_kurtosis_a	0	0	0	0	1	1
PLP9_kurtosis_a	0	0	0	1	1	2
PLP10_kurtosis_a	0	0	0	0	0	0
PLP11_kurtosis_a	0	0	0	0	1	1
PLP12_kurtosis_a	0	0	0	0	0	0
PLP0_D_kurtosis_a	0	0	0	0	1	1
PLP1_D_kurtosis_a	1	0	0	0	1	2
PLP2_D_kurtosis_a	0	0	0	1	1	2
PLP3_D_kurtosis_a	0	0	0	0	1	1
PLP4_D_kurtosis_a	0	0	0	0	0	0
PLP5_D_kurtosis_a	0	0	0	0	0	0
PLP6_D_kurtosis_a	0	0	0	0	0	0
PLP7_D_kurtosis_a	0	0	0	0	1	1
PLP8_D_kurtosis_a	0	0	0	0	0	0
PLP9_D_kurtosis_a	0	0	0	0	1	1
PLP10_D_kurtosis_a	0	0	0	1	0	1
PLP11_D_kurtosis_a	0	0	0	1	1	2
PLP12_D_kurtosis_a	0	0	0	0	0	0
PLP0_DD_kurtosis_a	0	0	0	0	1	1
PLP1_DD_kurtosis_a	0	0	0	0	1	1
PLP2_DD_kurtosis_a	0	0	0	1	1	2
PLP3_DD_kurtosis_a	0	0	0	0	1	1
PLP4_DD_kurtosis_a	0	0	0	0	0	0
PLP5_DD_kurtosis_a	0	0	0	0	0	0
PLP6_DD_kurtosis_a	0	0	0	1	0	1
PLP7_DD_kurtosis_a	0	0	0	0	1	1
PLP8_DD_kurtosis_a	0	0	0	0	0	0
PLP9_DD_kurtosis_a	0	0	0	0	1	1
PLP10_DD_kurtosis_a	0	0	0	1	0	1
PLP11_DD_kurtosis_a	0	0	0	0	1	1

(Continua alla pagina successiva)

(Continua dalla pagina precedente)

PLP12_DD_kurtosis_a	0	0	0	1	0	1
F0_skewness_a	0	0	0	0	0	0
HNR_skewness_a	1	0	1	0	1	3
STE_skewness_a	0	0	1	1	1	3
CPP_skewness_a	1	0	0	1	1	3
GNE_skewness_a	1	0	1	0	0	2
SPEC_CENT_GRAVITY_skewness_a	0	0	0	0	0	0
SPEC_STD_skewness_a	1	0	0	0	1	2
SPEC_SKEWNESS_skewness_a	0	0	0	0	1	1
SPEC_KURTOSIS_skewness_a	0	0	0	0	1	1
SPEC_FLUX_skewness_a	0	0	1	1	1	3
SPEC_SLOPE_skewness_a	0	0	0	0	0	0
SPEC_DECREASE_skewness_a	0	0	0	0	0	0
SPEC_ROLLOFF_skewness_a	0	0	1	1	1	3
FORM1_skewness_a	0	0	1	1	0	2
FORM2_skewness_a	0	0	0	0	1	1
FORM3_skewness_a	1	0	0	0	0	1
FORM_BAND1_skewness_a	0	0	0	0	1	1
FORM_BAND2_skewness_a	0	0	1	0	1	2
FORM_BAND3_skewness_a	0	0	1	0	1	2
MFCC1_skewness_a	1	0	0	0	1	2
MFCC2_skewness_a	0	0	0	0	0	0
MFCC3_skewness_a	0	1	0	0	1	2
MFCC4_skewness_a	0	0	0	0	1	1
MFCC5_skewness_a	1	0	0	0	0	1
MFCC6_skewness_a	1	0	0	0	0	1
MFCC7_skewness_a	0	0	0	0	0	0
MFCC8_skewness_a	0	0	0	0	1	1
MFCC9_skewness_a	0	0	0	0	0	0
MFCC10_skewness_a	0	0	0	0	0	0
MFCC11_skewness_a	1	0	0	0	1	2
MFCC12_skewness_a	0	0	0	0	0	0
MFCC13_skewness_a	0	0	0	0	0	0
MFCC1_D_skewness_a	0	0	0	0	0	0
MFCC1_DD_skewness_a	1	0	0	0	0	1
MFCC2_D_skewness_a	1	0	0	0	1	2
MFCC2_DD_skewness_a	0	0	0	1	0	1
MFCC3_D_skewness_a	0	0	1	0	0	1
MFCC3_DD_skewness_a	1	0	0	0	0	1
MFCC4_D_skewness_a	0	0	0	0	0	0
MFCC4_DD_skewness_a	0	0	0	0	0	0
MFCC5_D_skewness_a	0	0	0	0	0	0
MFCC5_DD_skewness_a	0	0	0	0	0	0

(Continua alla pagina successiva)

(Continua dalla pagina precedente)

MFCC6_D_skewness_a	0	0	0	0	0	0
MFCC6_DD_skewness_a	0	1	0	0	0	1
MFCC7_D_skewness_a	0	0	0	0	0	0
MFCC7_DD_skewness_a	0	0	0	0	0	0
MFCC8_D_skewness_a	1	0	0	0	1	2
MFCC8_DD_skewness_a	0	0	0	0	0	0
MFCC9_D_skewness_a	0	0	0	0	0	0
MFCC9_DD_skewness_a	0	0	0	0	0	0
MFCC10_D_skewness_a	0	0	0	0	1	1
MFCC10_DD_skewness_a	0	0	0	0	0	0
MFCC11_D_skewness_a	1	0	0	1	1	3
MFCC11_DD_skewness_a	0	0	0	0	0	0
MFCC12_D_skewness_a	0	0	0	0	0	0
MFCC12_DD_skewness_a	0	0	0	0	0	0
MFCC13_D_skewness_a	0	0	0	0	0	0
MFCC13_DD_skewness_a	0	0	0	0	0	0
PLP0_skewness_a	1	0	0	0	1	2
PLP1_skewness_a	1	0	0	0	0	1
PLP2_skewness_a	1	0	0	0	0	1
PLP3_skewness_a	1	0	0	0	1	2
PLP4_skewness_a	0	0	0	0	0	0
PLP5_skewness_a	0	0	0	0	0	0
PLP6_skewness_a	1	0	0	0	0	1
PLP7_skewness_a	0	0	0	0	0	0
PLP8_skewness_a	0	0	0	0	0	0
PLP9_skewness_a	0	0	1	0	0	1
PLP10_skewness_a	0	0	0	0	0	0
PLP11_skewness_a	0	0	1	0	0	1
PLP12_skewness_a	0	0	0	0	1	1
PLP0_D_skewness_a	1	0	0	0	1	2
PLP1_D_skewness_a	1	0	0	0	0	1
PLP2_D_skewness_a	0	0	0	0	1	1
PLP3_D_skewness_a	0	1	0	0	0	1
PLP4_D_skewness_a	0	0	0	0	0	0
PLP5_D_skewness_a	0	0	0	0	0	0
PLP6_D_skewness_a	0	0	0	0	0	0
PLP7_D_skewness_a	0	0	0	0	0	0
PLP8_D_skewness_a	0	0	0	0	0	0
PLP9_D_skewness_a	0	0	0	0	0	0
PLP10_D_skewness_a	0	0	0	0	0	0
PLP11_D_skewness_a	0	0	1	0	0	1
PLP12_D_skewness_a	0	0	0	0	0	0
PLP0_DD_skewness_a	1	0	1	0	0	2

(Continua alla pagina successiva)

(Continua dalla pagina precedente)

PLP1_DD_skewness_a	0	0	0	0	0	0
PLP2_DD_skewness_a	1	0	1	0	0	2
PLP3_DD_skewness_a	0	0	0	0	1	1
PLP4_DD_skewness_a	0	0	0	0	0	0
PLP5_DD_skewness_a	0	0	0	0	0	0
PLP6_DD_skewness_a	1	0	0	0	0	1
PLP7_DD_skewness_a	0	0	0	1	0	1
PLP8_DD_skewness_a	0	0	0	0	0	0
PLP9_DD_skewness_a	0	0	0	0	0	0
PLP10_DD_skewness_a	0	0	0	0	0	0
PLP11_DD_skewness_a	0	0	0	0	0	0
PLP12_DD_skewness_a	0	0	0	0	0	0

Tabella 6.22: feature con valore di p riferito alla classe ottenuto con il test di Mann-Whitney, per i 5 dataset

6.6.2 Analisi sui 5 dataset uniti

In Tabella 6.23 sono rappresentate le feature con valore di $p < 0,05$ riferito alla classe ottenuto con il test di Mann-Whitney. In essa sono anche riportati i valori di p, separatamente per pazienti con MP e controlli sani, riferiti al genere (ottenuti con il test di Mann-Whitney), al dataset (ottenuti con il test di Kruskal-Wallis) e alla lingua (ottenuti con il test di Kruskal-Wallis).

In Tabella 6.24 sono rappresentate le feature con valore di $p > 0,05$ riferito alla classe ottenuto con il test di Mann-Whitney. In essa sono anche riportati i valori di p, separatamente per pazienti con MP e controlli sani, riferiti al genere (ottenuti con il test di Mann-Whitney), al dataset (ottenuti con il test di Kruskal-Wallis) e alla lingua (ottenuti con il test di Kruskal-Wallis).

Tabella 6.23: feature statisticamente discriminanti in termini di classe

feature	Controlli sani				Pazienti MP		
	Clas.	Gen	Datas.	Ling	Gen	Datas.	Ling
MFCC4_mean_a	<0,01	0,469	0,446	0,239	0,266	0,393	0,211
MFCC4_median_a	<0,01	0,363	0,492	0,254	0,291	0,472	0,279
MFCC7_DD_median_a	0,01	0,463	0,287	0,793	0,065	0,739	0,718
MFCC13_D_median_a	0,043	0,73	0,654	0,812	0,477	0,818	0,579
MFCC13_DD_median_a	0,021	0,549	0,603	0,354	0,208	0,363	0,526
MFCC7_DD_skewness_a	0,029	0,417	0,506	0,389	0,122	0,451	0,65
PLP2_DD_skewness_a	<0,01	0,854	0,706	0,852	0,153	0,019	0,898

(Continua alla pagina successiva)

(Continua dalla pagina precedente)

MFCC1_D_mean_a	0,024	0,805	0,098	0,186	0,5	<0,01	0,032
MFCC5_D_mean_a	0,01	0,866	0,65	0,702	0,77	<0,01	<0,01
MFCC5_skewness_a	0,032	0,952	0,804	0,678	0,034	0,116	0,129
PLP3_mean_a	0,013	0,137	0,177	0,808	0,003	0,012	0,015
PLP5_D_mean_a	0,025	0,797	0,083	0,017	0,254	0,063	0,03
MFCC6_skewness_a	0,039	0,669	0,024	0,455	0,075	0,089	0,501
MFCC6_DD_skewness_a	0,018	0,986	0,04	0,069	0,865	0,69	0,454
PLP5_DD_std_a	0,038	0,457	<0,01	0,065	0,71	<0,01	0,728
PLP12_DD_std_a	0,014	0,532	<0,01	0,095	0,287	<0,01	0,153
SPEC_FLUX_mean_a	0,004	0,078	0,003	0,127	0,569	0,001	<0,01
PLP5_mean_a	0,036	0,613	0,023	0,006	0,497	0,144	0,434
MFCC5_kurtosis_a	0,04	0,396	0,045	0,02	0,734	0,051	0,046
PLP0_median_a	0,005	0,807	<0,01	<0,01	0,85	<0,01	0,083
PLP1_median_a	0,045	0,426	<0,01	0,011	0,17	0,005	0,051
STE_std_a	<0,01	0,598	<0,01	<0,01	0,293	<0,01	0,066
SPEC_SLOPE_std_a	0,027	0,293	<0,01	<0,01	0,806	0,028	0,284
MFCC11_D_std_a	<0,01	0,588	<0,01	0,003	0,285	<0,01	0,084
PLP8_D_std_a	0,011	0,738	<0,01	0,005	0,346	<0,01	0,254
PLP9_D_std_a	0,047	0,777	<0,01	0,033	0,112	<0,01	0,067
PLP8_DD_std_a	0,007	0,606	<0,01	0,005	0,418	<0,01	0,276
PLP9_DD_std_a	0,032	0,695	<0,01	0,028	0,121	<0,01	0,296
INT_AVG_a	0,011	0,206	<0,01	<0,01	0,295	<0,01	<0,01
STE_mean_a	0,009	0,208	<0,01	<0,01	0,298	<0,01	<0,01
MFCC3_mean_a	<0,01	0,405	0,01	0,004	0,836	<0,01	<0,01
PLP1_mean_a	0,005	0,328	<0,01	0,019	0,129	<0,01	<0,01
PLP2_D_mean_a	0,022	0,248	<0,01	<0,01	0,512	<0,01	<0,01
PLP2_DD_mean_a	0,001	0,809	<0,01	<0,01	0,637	0,04	0,009
STE_median_a	0,006	0,244	<0,01	<0,01	0,471	<0,01	<0,01
SPEC_FLUX_median_a	0,012	0,492	<0,01	<0,01	0,477	<0,01	<0,01
MFCC3_median_a	<0,01	0,391	0,011	0,005	0,85	<0,01	<0,01
HNR_std_a	0,028	0,118	<0,01	<0,01	0,076	<0,01	<0,01
CPP_std_a	0,041	0,143	<0,01	<0,01	0,506	<0,01	<0,01
SPEC_SKEWNESS_std_a	0,042	0,498	<0,01	<0,01	0,731	<0,01	<0,01
SPEC_KURTOSIS_std_a	0,026	0,374	<0,01	<0,01	0,522	<0,01	<0,01
SPEC_FLUX_std_a	0,011	0,057	<0,01	<0,01	0,346	<0,01	<0,01
MFCC6_std_a	0,03	0,94	<0,01	0,001	0,112	<0,01	<0,01
MFCC8_std_a	0,009	0,986	<0,01	<0,01	0,112	<0,01	<0,01
MFCC10_std_a	0,008	0,128	<0,01	<0,01	0,08	<0,01	<0,01
MFCC11_std_a	0,022	0,143	<0,01	0,016	0,163	<0,01	<0,01
MFCC2_DD_std_a	0,012	0,505	<0,01	<0,01	0,806	<0,01	<0,01
MFCC3_DD_std_a	0,042	0,642	<0,01	<0,01	0,631	<0,01	<0,01
MFCC4_D_std_a	0,02	0,703	<0,01	<0,01	0,334	<0,01	<0,01
MFCC4_DD_std_a	0,01	0,642	<0,01	<0,01	0,361	<0,01	<0,01

(Continua alla pagina successiva)

(Continua dalla pagina precedente)

MFCC5_D_std_a	0,034	0,801	<0,01	<0,01	0,179	<0,01	<0,01
MFCC5_DD_std_a	0,028	0,45	<0,01	<0,01	0,149	<0,01	<0,01
MFCC6_D_std_a	0,003	0,642	<0,01	<0,01	0,512	<0,01	<0,01
MFCC6_DD_std_a	0,003	0,595	<0,01	<0,01	0,44	<0,01	<0,01
MFCC7_D_std_a	0,043	0,417	<0,01	<0,01	0,11	<0,01	<0,01
MFCC7_DD_std_a	0,02	0,214	<0,01	<0,01	0,126	<0,01	<0,01
MFCC8_D_std_a	0,01	0,132	<0,01	<0,01	0,076	<0,01	<0,01
MFCC8_DD_std_a	0,005	0,055	<0,01	<0,01	0,091	<0,01	<0,01
MFCC9_D_std_a	0,006	0,669	<0,01	0,001	0,334	<0,01	0,006
MFCC9_DD_std_a	0,005	0,463	<0,01	<0,01	0,336	<0,01	0,001
MFCC10_D_std_a	0,001	0,624	<0,01	<0,01	0,067	<0,01	0,028
MFCC10_DD_std_a	<0,01	0,368	<0,01	<0,01	0,073	<0,01	0,002
MFCC11_DD_std_a	<0,01	0,284	<0,01	<0,01	0,175	<0,01	0,04
MFCC12_D_std_a	0,001	0,952	<0,01	0,001	0,442	<0,01	0,025
MFCC12_DD_std_a	<0,01	0,535	<0,01	<0,01	0,325	<0,01	0,008
MFCC13_D_std_a	<0,01	0,797	<0,01	<0,01	0,327	<0,01	0,011
MFCC13_DD_std_a	<0,01	0,42	<0,01	<0,01	0,189	<0,01	0,007
PLP8_std_a	0,016	0,882	<0,01	<0,01	0,193	<0,01	0,033
GNE_kurtosis_a	0,019	0,382	<0,01	<0,01	0,799	<0,01	<0,01
SPEC_STD_kurtosis_a	0,03	0,502	<0,01	<0,01	0,289	<0,01	<0,01
SPEC_SKEWNESS_kurtosis_a	0,007	0,15	<0,01	<0,01	0,749	<0,01	<0,01
SPEC_KURTOSIS_kurtosis_a	0,001	0,071	<0,01	0,008	0,474	0,011	0,002
FORM_BAND2_kurtosis_a	0,013	0,588	<0,01	<0,01	0,334	<0,01	<0,01
MFCC2_D_kurtosis_a	0,035	0,082	0,002	<0,01	0,932	<0,01	<0,01
PLP11_D_kurtosis_a	0,034	0,054	0,044	0,01	0,358	0,002	<0,01
GNE_skewness_a	0,016	0,365	<0,01	<0,01	0,792	<0,01	<0,01
FORM_BAND2_skewness_a	0,013	0,688	<0,01	<0,01	0,302	<0,01	<0,01
PLP0_skewness_a	0,006	0,411	<0,01	<0,01	0,756	<0,01	<0,01
PLP0_DD_skewness_a	<0,01	0,793	<0,01	<0,01	0,588	<0,01	<0,01
GNE_mean_a	0,035	0,089	<0,01	<0,01	0,002	<0,01	<0,01
GNE_std_a	0,003	0,234	<0,01	<0,01	0,025	<0,01	<0,01
FORM2_std_a	0,001	0,919	<0,01	<0,01	<0,01	<0,01	<0,01
JIT_L_a	<0,01	0,008	0,583	0,401	0,285	0,214	0,866
JIT_LAb_a	0,004	<0,01	0,522	0,397	0,079	0,432	0,985
FORM_BAND1_median_a	<0,01	0,038	0,461	0,832	0,065	0,869	0,866
PLP6_median_a	0,005	0,012	0,927	0,695	0,21	0,262	0,093
SPEC_SKEWNESS_mean_a	0,01	<0,01	0,242	0,382	0,179	<0,01	0,008
SPEC_SKEWNESS_median_a	0,026	0,001	0,458	0,544	0,101	<0,01	0,007
JIT_RAP_a	<0,01	<0,01	0,09	0,019	0,585	0,347	0,407
JIT_DDP_a	<0,01	<0,01	0,09	0,019	0,585	0,347	0,407
JIT_PPQ5_a	<0,01	<0,01	0,146	0,04	0,129	0,24	0,525
FORM_BAND1_mean_a	<0,01	0,045	0,036	0,492	0,101	0,659	0,39
F0_std_a	0,036	<0,01	0,008	0,167	0,053	0,165	0,55

(Continua alla pagina successiva)

(Continua dalla pagina precedente)

MFCC2_mean_a	<0,01	<0,01	0,016	0,065	0,013	<0,01	<0,01
MFCC2_median_a	<0,01	<0,01	0,013	0,056	0,011	<0,01	<0,01
SPEC_KURTOSIS_skewness_a	0,011	0,028	<0,01	0,019	0,763	0,08	0,039
SPEC_CENT_GRAVITY_mean_a	<0,01	0,003	<0,01	<0,01	0,318	<0,01	<0,01
SPEC_CENT_GRAVITY_median_a	<0,01	0,021	<0,01	<0,01	0,26	<0,01	<0,01
FORM1_std_a	<0,01	0,032	<0,01	0,019	0,54	<0,01	<0,01
PLP11_DD_std_a	0,019	0,025	<0,01	0,002	0,322	<0,01	0,006
CPP_kurtosis_a	0,024	0,001	<0,01	<0,01	0,727	<0,01	<0,01
FORM_BAND3_kurtosis_a	0,025	0,042	<0,01	<0,01	0,278	0,034	0,011
CPP_skewness_a	0,001	0,022	<0,01	<0,01	0,111	<0,01	<0,01
FORM_BAND3_skewness_a	0,021	0,044	<0,01	<0,01	0,192	0,045	0,021
FORM1_mean_a	0,034	<0,01	<0,01	<0,01	0,007	<0,01	<0,01
FORM1_median_a	0,035	<0,01	<0,01	<0,01	0,003	<0,01	<0,01

Tabella 6.23: feature che non presentano alcuna differenza statisticamente significativa in termini di classe

Tabella 6.24: feature che non presentano differenze statisticamente significative in termini di classe

feature	Controlli sani				Pazienti MP		
	Clas.	Gen	Datas.	Ling	Gen	Datas.	Ling
SPEC_SLOPE_mean_a	0,834	0,24	0,792	0,916	0,127	0,325	0,117
SPEC_DECREASE_mean_a	0,266	0,399	0,339	0,121	0,724	0,541	0,265
MFCC5_mean_a	0,158	0,532	0,215	0,524	0,073	0,686	0,48
MFCC13_mean_a	0,168	0,23	0,185	0,136	0,869	0,464	0,208
MFCC1_DD_mean_a	0,789	0,136	0,098	0,379	0,767	0,074	0,129
MFCC5_DD_mean_a	0,897	0,969	0,584	0,476	0,749	0,894	0,656
MFCC6_DD_mean_a	0,09	0,248	0,445	0,192	0,727	0,552	0,502
MFCC7_D_mean_a	0,137	0,981	0,318	0,201	0,556	0,366	0,233
MFCC7_DD_mean_a	0,257	0,222	0,199	0,362	0,627	0,325	0,838
MFCC9_DD_mean_a	0,687	0,957	0,645	0,57	0,097	0,562	0,317
MFCC10_D_mean_a	0,978	0,216	0,81	0,912	0,42	0,2	0,338
MFCC10_DD_mean_a	0,739	0,899	0,956	0,783	0,675	0,372	0,416
MFCC11_D_mean_a	0,667	0,672	0,319	0,1	0,434	0,361	0,889
MFCC11_DD_mean_a	0,129	0,238	0,888	0,849	0,658	0,474	0,238
MFCC12_D_mean_a	0,718	0,072	0,862	0,64	0,378	0,453	0,315
MFCC12_DD_mean_a	0,487	0,932	0,348	0,695	0,474	0,467	0,236
PLP8_D_mean_a	0,395	0,363	0,504	0,421	0,203	0,157	0,081
PLP9_D_mean_a	0,506	0,168	0,222	0,088	0,189	0,911	0,618
PLP10_D_mean_a	0,11	0,284	0,082	0,501	0,114	0,227	0,604
PLP5_DD_mean_a	0,419	0,382	0,141	0,851	0,118	0,973	0,941
PLP8_DD_mean_a	0,199	0,99	0,451	0,341	0,598	0,451	0,184
PLP9_DD_mean_a	0,761	0,084	0,978	0,969	0,12	0,689	0,379
PLP10_DD_mean_a	0,26	0,232	0,187	0,155	0,235	0,811	0,901
PLP11_DD_mean_a	0,922	0,182	0,782	0,438	0,181	0,265	0,191
PLP12_DD_mean_a	0,797	0,272	0,978	0,965	0,738	0,439	0,504
MFCC5_median_a	0,104	0,489	0,234	0,489	0,07	0,623	0,419
MFCC13_median_a	0,165	0,25	0,184	0,154	0,843	0,449	0,195
MFCC2_DD_median_a	0,27	0,834	0,727	0,442	0,203	0,57	0,309
MFCC3_DD_median_a	0,618	0,75	0,505	0,192	0,468	0,692	0,401
MFCC5_D_median_a	0,26	0,793	0,405	0,817	0,781	0,117	0,106
MFCC6_D_median_a	0,895	0,21	0,64	0,832	0,083	0,67	0,324
MFCC6_DD_median_a	0,248	0,502	0,429	0,463	0,442	0,768	0,587
MFCC7_D_median_a	0,619	0,088	0,271	0,484	0,876	0,277	0,148
MFCC8_D_median_a	0,827	0,891	0,289	0,569	0,81	0,652	0,497
MFCC9_D_median_a	0,507	0,882	0,362	0,939	0,862	0,943	0,698
MFCC9_DD_median_a	0,734	0,758	0,355	0,345	0,059	0,913	0,892
MFCC10_DD_median_a	0,696	0,22	0,19	0,106	0,91	0,436	0,218
MFCC11_DD_median_a	0,486	0,447	0,803	0,592	0,749	0,761	0,803

(Continua alla pagina successiva)

(Continua dalla pagina precedente)

MFCC12_D_median_a	0,316	0,134	0,218	0,1	0,921	0,194	0,059
MFCC12_DD_median_a	0,283	0,376	0,783	0,453	0,451	0,442	0,299
PLP5_median_a	0,074	0,765	0,307	0,095	0,35	0,06	0,109
PLP9_median_a	0,503	0,151	0,21	0,062	0,309	0,254	0,465
PLP12_median_a	0,135	0,442	0,726	0,504	0,895	0,998	0,98
PLP5_D_median_a	0,852	0,691	0,363	0,191	0,051	0,324	0,685
PLP6_D_median_a	0,837	0,323	0,482	0,723	0,546	0,878	0,832
PLP7_D_median_a	0,085	0,203	0,633	0,299	0,181	0,35	0,465
PLP8_D_median_a	0,136	0,205	0,274	0,466	0,682	0,232	0,078
PLP10_D_median_a	0,476	0,165	0,387	0,313	0,239	0,463	0,411
PLP1_DD_median_a	0,654	0,489	0,643	0,398	0,601	0,992	0,89
PLP4_DD_median_a	0,257	0,065	0,98	0,898	0,477	0,251	0,241
PLP5_DD_median_a	0,664	0,834	0,34	0,18	0,503	0,28	0,119
PLP6_DD_median_a	0,269	0,485	0,927	0,725	0,788	0,711	0,346
PLP8_DD_median_a	0,447	0,275	0,539	0,945	0,256	0,273	0,128
PLP9_DD_median_a	0,516	0,051	0,594	0,252	0,213	0,859	0,831
PLP10_DD_median_a	0,96	0,99	0,706	0,405	0,48	0,976	0,848
PLP11_DD_median_a	0,275	0,722	0,328	0,292	0,581	0,595	0,533
PLP12_DD_median_a	0,581	0,817	0,552	0,255	0,862	0,747	0,923
MFCC2_kurtosis_a	0,218	0,528	0,506	0,994	0,682	0,086	0,255
MFCC4_kurtosis_a	0,528	0,588	0,065	0,271	0,832	0,114	0,337
MFCC12_kurtosis_a	0,492	0,365	0,153	0,247	0,531	0,548	0,283
MFCC8_D_kurtosis_a	0,771	0,111	0,055	0,108	0,869	0,183	0,146
MFCC13_D_kurtosis_a	0,167	0,057	0,726	0,447	0,756	0,078	0,091
SPEC_SLOPE_skewness_a	0,49	0,838	0,959	0,857	0,208	0,517	0,271
MFCC4_skewness_a	0,333	0,515	0,373	0,333	0,611	0,065	0,39
MFCC9_skewness_a	0,059	0,825	0,303	0,113	0,423	0,835	0,912
MFCC10_skewness_a	0,858	0,205	0,864	0,745	0,474	0,257	0,548
MFCC12_skewness_a	0,763	0,646	0,343	0,32	0,208	0,488	0,417
MFCC5_D_skewness_a	0,319	0,069	0,187	0,092	0,06	0,991	0,908
MFCC5_DD_skewness_a	0,141	0,994	0,621	0,982	0,145	0,776	0,851
MFCC6_D_skewness_a	0,854	0,253	0,568	0,843	0,477	0,353	0,147
MFCC7_D_skewness_a	0,091	0,077	0,95	0,817	0,738	0,877	0,564
MFCC8_DD_skewness_a	0,226	0,17	0,288	0,157	0,459	0,309	0,65
MFCC9_D_skewness_a	0,834	0,887	0,87	0,775	0,289	0,365	0,661
MFCC9_DD_skewness_a	0,584	0,85	0,21	0,07	0,166	0,185	0,054
MFCC10_DD_skewness_a	0,324	0,639	0,159	0,262	0,627	0,413	0,476
MFCC11_DD_skewness_a	0,763	0,932	0,7	0,506	0,865	0,422	0,295
MFCC12_D_skewness_a	0,306	0,277	0,979	0,864	0,692	0,185	0,152
MFCC12_DD_skewness_a	0,94	0,707	0,184	0,109	0,661	0,549	0,303
MFCC13_D_skewness_a	0,396	0,672	0,412	0,613	0,685	0,839	0,557
MFCC13_DD_skewness_a	0,347	0,688	0,485	0,464	0,699	0,765	0,641
PLP3_skewness_a	0,202	0,973	0,309	0,74	0,752	0,256	0,937

(Continua alla pagina successiva)

(Continua dalla pagina precedente)

PLP6_skewness_a	0,137	0,758	0,941	0,906	0,491	0,433	0,465
PLP8_skewness_a	0,833	0,858	0,127	0,23	0,598	0,728	0,467
PLP9_skewness_a	0,176	0,066	0,328	0,144	0,09	0,37	0,203
PLP10_skewness_a	0,251	0,961	0,192	0,566	0,914	0,058	0,12
PLP11_skewness_a	0,068	0,595	0,264	0,089	0,252	0,432	0,245
PLP4_D_skewness_a	0,152	0,07	0,602	0,32	0,668	0,96	0,8
PLP6_D_skewness_a	0,611	0,172	0,733	0,482	0,208	0,177	0,285
PLP8_D_skewness_a	0,643	0,887	0,414	0,264	0,763	0,66	0,308
PLP9_D_skewness_a	0,173	0,87	0,108	0,288	0,644	0,442	0,978
PLP10_D_skewness_a	0,96	0,726	0,173	0,316	0,966	0,965	0,867
PLP3_DD_skewness_a	0,843	0,924	0,168	0,228	0,575	0,156	0,381
PLP5_DD_skewness_a	0,998	0,588	0,293	0,385	0,873	0,447	0,168
PLP8_DD_skewness_a	0,384	0,676	0,499	0,63	0,491	0,698	0,473
PLP9_DD_skewness_a	0,704	0,447	0,987	0,929	0,869	0,344	0,355
PLP10_DD_skewness_a	0,783	0,411	0,213	0,203	0,149	0,501	0,316
PLP11_DD_skewness_a	0,187	0,388	0,718	0,59	0,962	0,817	0,494
PLP12_DD_skewness_a	0,915	0,878	0,495	0,477	0,341	0,814	0,526
MFCC9_D_mean_a	0,543	0,277	0,137	0,124	0,906	0,053	0,019
PLP12_mean_a	0,187	0,928	0,247	0,313	0,278	0,144	0,037
MFCC10_D_median_a	0,276	0,915	0,627	0,373	0,396	0,152	0,048
MFCC9_kurtosis_a	0,595	0,124	0,083	0,48	0,451	0,054	0,047
SPEC_DECREASE_skewness_a	0,39	0,522	0,963	0,774	0,959	0,091	0,039
MFCC8_skewness_a	0,095	0,878	0,593	0,856	0,48	0,108	0,035
MFCC13_skewness_a	0,494	0,177	0,363	0,19	0,594	0,151	0,045
MFCC4_DD_skewness_a	0,527	0,374	0,314	0,13	0,944	0,102	0,029
PLP5_skewness_a	0,7	0,453	0,151	0,077	0,888	0,144	0,043
PLP5_D_skewness_a	0,571	0,154	0,314	0,096	0,627	0,126	0,034
PLP7_D_skewness_a	0,757	0,064	0,86	0,951	0,412	0,082	0,04
PLP12_D_skewness_a	0,348	0,866	0,239	0,377	0,235	0,159	0,041
PLP8_mean_a	0,732	0,101	0,358	0,267	0,404	0,026	0,191
MFCC7_kurtosis_a	0,783	0,105	0,052	0,148	0,955	0,008	0,071
MFCC10_D_skewness_a	0,772	0,539	0,213	0,377	0,512	0,018	0,795
PLP10_mean_a	0,44	0,159	0,132	0,076	0,929	0,003	0,002
PLP7_D_mean_a	0,721	0,067	0,193	0,122	0,109	<0,01	<0,01
FORM1_kurtosis_a	0,459	0,508	0,449	0,168	0,88	0,004	0,006
FORM3_kurtosis_a	0,192	0,928	0,203	0,107	0,531	0,003	0,014
MFCC3_kurtosis_a	0,947	0,153	0,647	0,442	0,318	0,049	0,013
MFCC10_kurtosis_a	0,524	0,944	0,103	0,409	0,836	0,001	0,005
MFCC13_kurtosis_a	0,595	0,469	0,388	0,564	0,07	0,037	0,034
MFCC5_D_kurtosis_a	0,451	0,114	0,201	0,059	0,788	0,002	<0,01
MFCC6_DD_kurtosis_a	0,725	0,411	0,058	0,063	0,396	0,001	<0,01
MFCC7_DD_kurtosis_a	0,162	0,117	0,206	0,108	0,412	<0,01	<0,01
MFCC11_skewness_a	0,309	0,077	0,105	0,281	0,899	0,002	<0,01

(Continua alla pagina successiva)

(Continua dalla pagina precedente)

MFCC3_DD_skewness_a	0,266	0,472	0,63	0,945	0,147	0,006	0,005
MFCC8_D_skewness_a	0,437	0,376	0,457	0,449	0,665	<0,01	<0,01
PLP7_skewness_a	0,681	0,206	0,502	0,728	0,21	0,006	0,002
PLP12_skewness_a	0,85	0,306	0,091	0,429	0,591	0,009	0,003
PLP3_D_skewness_a	0,064	0,73	0,521	0,417	0,77	0,013	0,004
PLP7_DD_skewness_a	0,172	0,308	0,659	0,749	0,105	0,036	0,028
MFCC8_DD_mean_a	0,453	0,642	0,204	0,105	0,009	0,74	0,558
PLP11_D_mean_a	0,282	0,326	0,273	0,133	<0,01	0,776	0,823
MFCC3_D_median_a	0,381	0,218	0,852	0,524	0,025	0,228	0,07
MFCC5_DD_median_a	0,247	0,472	0,321	0,275	0,004	0,866	0,573
PLP11_D_median_a	0,304	0,874	1,000	0,992	<0,01	0,39	0,903
PLP12_D_median_a	0,064	0,734	0,157	0,126	0,032	0,498	0,225
PLP3_DD_median_a	0,705	0,104	0,349	0,807	<0,01	0,684	0,674
MFCC7_skewness_a	0,563	0,344	0,142	0,069	0,041	0,921	0,956
PLP6_DD_skewness_a	0,337	0,489	0,25	0,851	0,023	0,353	0,209
PLP6_DD_mean_a	0,88	0,447	0,89	0,657	0,017	0,093	0,02
MFCC13_D_mean_a	0,192	0,079	0,159	0,123	0,024	0,043	0,079
PLP3_median_a	0,35	0,129	0,096	0,875	<0,01	0,024	0,013
MFCC3_DD_mean_a	0,823	0,498	0,081	0,035	0,175	0,438	0,28
MFCC8_D_mean_a	0,738	0,148	0,061	0,014	0,289	0,218	0,234
PLP4_mean_a	0,257	0,472	0,056	0,014	0,585	0,137	0,29
MFCC8_DD_median_a	0,517	0,657	0,114	0,033	0,588	0,677	0,807
PLP4_median_a	0,781	0,357	0,063	0,012	0,962	0,242	0,149
PLP11_median_a	0,937	0,68	0,092	0,039	0,983	0,554	0,573
PLP3_D_median_a	0,149	0,289	0,119	0,047	0,506	0,054	0,091
F0_skewness_a	0,829	0,253	0,05	0,017	0,094	0,275	0,088
PLP11_D_skewness_a	0,627	0,298	0,116	0,038	0,08	0,363	0,158
MFCC5_DD_kurtosis_a	0,374	0,665	0,052	0,015	0,921	<0,01	<0,01
MFCC7_D_kurtosis_a	0,241	0,068	0,05	0,023	0,285	0,005	0,003
PLP7_DD_mean_a	0,563	0,715	0,082	0,037	<0,01	0,33	0,147
FORM_BAND2_mean_a	0,635	0,508	0,007	0,516	0,713	0,354	0,389
MFCC4_DD_mean_a	0,442	0,222	0,042	0,102	0,724	0,151	0,076
PLP11_mean_a	0,913	0,765	0,049	0,06	0,671	0,245	0,068
PLP9_D_median_a	0,117	0,809	0,037	0,058	0,884	0,08	0,124
MFCC1_DD_skewness_a	0,865	0,75	0,005	0,091	0,09	0,081	0,773
PLP3_D_std_a	0,491	0,617	<0,01	0,147	0,774	<0,01	0,837
PLP5_D_std_a	0,092	0,62	<0,01	0,14	0,5	<0,01	0,445
PLP12_D_std_a	0,074	0,628	<0,01	0,058	0,388	<0,01	0,126
PLP3_DD_std_a	0,275	0,813	<0,01	0,077	0,825	<0,01	0,909
MFCC11_D_skewness_a	0,846	0,552	0,026	0,051	0,3	0,012	0,513
MFCC8_mean_a	0,863	0,777	0,044	0,11	0,459	<0,01	<0,01
MFCC8_median_a	0,795	0,754	0,046	0,097	0,426	0,001	0,001
PLP4_DD_skewness_a	0,775	0,726	0,008	0,841	0,383	0,029	0,02

(Continua alla pagina successiva)

(Continua dalla pagina precedente)

MFCC3_D_mean_a	0,682	0,602	0,005	<0,01	0,058	0,2	0,069
PLP9_mean_a	0,158	0,115	<0,01	<0,01	0,601	0,451	0,921
PLP12_D_mean_a	0,962	0,85	0,011	0,003	0,624	0,409	0,192
SPEC_DECREASE_median_a	0,777	0,769	0,011	0,003	0,153	0,415	0,207
FORM_BAND2_median_a	0,795	0,588	0,003	0,003	0,654	0,613	0,305
MFCC1_D_median_a	0,259	0,998	0,03	0,026	0,462	0,422	0,193
MFCC4_DD_median_a	0,713	0,192	0,006	0,026	0,955	0,75	0,518
MFCC11_D_median_a	0,485	0,32	0,03	0,005	0,803	0,148	0,479
PLP2_median_a	0,466	0,868	0,006	0,025	0,086	0,114	0,144
PLP10_median_a	0,097	0,093	0,023	0,042	0,836	0,244	0,077
MFCC11_kurtosis_a	0,764	0,151	0,012	0,003	0,858	0,131	0,129
MFCC2_skewness_a	0,276	0,718	0,014	0,019	0,313	0,609	0,328
MFCC2_DD_skewness_a	0,261	0,159	0,028	0,011	0,217	0,148	0,607
MFCC4_D_skewness_a	0,621	0,684	0,013	0,003	0,239	0,459	0,252
PLP1_D_skewness_a	0,073	0,718	0,031	0,015	0,429	0,151	0,468
PLP3_DD_mean_a	0,202	0,426	0,009	0,002	0,459	0,06	0,02
HNR_median_a	0,166	0,091	0,005	0,002	0,447	<0,01	0,778
PLP2_std_a	0,475	0,46	<0,01	0,008	0,228	<0,01	0,838
PLP3_std_a	0,785	0,469	<0,01	0,01	0,741	<0,01	0,206
PLP4_std_a	0,654	0,919	<0,01	0,01	0,144	<0,01	0,352
PLP6_std_a	0,632	0,65	<0,01	0,002	0,266	<0,01	0,079
PLP2_D_std_a	0,345	0,508	<0,01	0,015	0,281	<0,01	0,738
PLP4_D_std_a	0,364	0,512	<0,01	0,007	0,17	<0,01	0,87
PLP6_D_std_a	0,267	0,617	<0,01	0,015	0,133	<0,01	0,277
PLP7_D_std_a	0,537	0,606	<0,01	0,013	0,562	<0,01	0,439
PLP2_DD_std_a	0,05	0,472	<0,01	<0,01	0,268	<0,01	0,36
PLP4_DD_std_a	0,231	0,313	<0,01	<0,01	0,189	<0,01	0,551
PLP6_DD_std_a	0,12	0,62	<0,01	0,008	0,097	<0,01	0,401
PLP7_DD_std_a	0,309	0,761	<0,01	0,012	0,594	<0,01	0,539
PLP10_DD_std_a	0,166	0,114	<0,01	0,028	0,738	<0,01	0,088
MFCC3_skewness_a	0,63	0,688	0,049	0,033	0,876	0,016	0,071
MFCC2_D_skewness_a	0,958	0,834	<0,01	<0,01	0,788	0,005	0,054
SHIM_a	0,534	0,341	<0,01	<0,01	0,404	<0,01	<0,01
SHIM_DB_a	0,612	0,357	<0,01	<0,01	0,418	<0,01	0,002
SHIM_APQ3_a	0,726	0,466	<0,01	<0,01	0,468	<0,01	<0,01
SHIM_APQ5_a	0,814	0,266	<0,01	<0,01	0,244	<0,01	<0,01
SHIM_APQ11_a	0,369	0,291	<0,01	<0,01	0,215	<0,01	0,013
DFA_COEFF_a	0,107	0,115	<0,01	<0,01	0,745	<0,01	<0,01
FORM_BAND3_mean_a	0,529	0,068	0,001	<0,01	0,41	0,001	<0,01
MFCC2_D_mean_a	0,336	0,192	<0,01	<0,01	0,329	<0,01	<0,01
MFCC4_D_mean_a	0,744	0,055	<0,01	<0,01	0,559	<0,01	<0,01
PLP0_mean_a	0,362	0,793	<0,01	<0,01	0,745	<0,01	<0,01
PLP2_mean_a	0,121	0,57	0,006	0,001	0,243	0,005	0,014

(Continua alla pagina successiva)

(Continua dalla pagina precedente)

PLP0_D_mean_a	0,585	0,214	<0,01	<0,01	0,202	<0,01	<0,01
PLP1_D_mean_a	0,874	0,118	<0,01	<0,01	0,756	<0,01	<0,01
PLP3_D_mean_a	0,165	0,858	0,001	<0,01	0,731	<0,01	<0,01
PLP6_D_mean_a	0,591	0,672	0,044	0,014	0,578	0,004	<0,01
PLP0_DD_mean_a	0,113	0,162	<0,01	<0,01	0,1	<0,01	<0,01
PLP1_DD_mean_a	0,395	0,072	0,003	0,022	0,88	<0,01	<0,01
PLP4_DD_mean_a	0,283	0,121	0,017	0,049	0,119	<0,01	<0,01
SPEC_SLOPE_median_a	0,055	0,336	<0,01	<0,01	0,795	<0,01	<0,01
FORM_BAND3_median_a	0,408	0,532	0,009	0,003	0,895	0,004	0,002
MFCC2_D_median_a	0,777	0,214	0,042	0,017	0,578	0,011	0,006
PLP0_D_median_a	0,19	0,85	0,026	0,017	0,959	<0,01	<0,01
PLP1_D_median_a	0,51	0,108	0,01	0,002	0,477	<0,01	<0,01
PLP2_D_median_a	0,156	0,94	0,009	0,001	0,072	0,009	0,008
PLP0_DD_median_a	0,081	0,376	<0,01	<0,01	0,839	<0,01	<0,01
PLP2_DD_median_a	0,414	0,393	0,008	0,001	0,434	0,005	0,028
SPEC_STD_std_a	0,349	0,203	<0,01	<0,01	0,122	<0,01	<0,01
SPEC_DECREASE_std_a	0,178	0,785	<0,01	<0,01	0,097	<0,01	<0,01
FORM3_std_a	0,101	0,546	<0,01	<0,01	0,383	<0,01	<0,01
FORM_BAND1_std_a	0,218	0,102	<0,01	<0,01	0,81	<0,01	<0,01
FORM_BAND2_std_a	0,304	0,944	<0,01	<0,01	0,215	<0,01	<0,01
MFCC1_std_a	0,145	0,084	<0,01	<0,01	0,839	<0,01	<0,01
MFCC2_std_a	0,127	0,291	<0,01	<0,01	0,471	<0,01	0,003
MFCC3_std_a	0,312	0,242	<0,01	<0,01	0,985	<0,01	<0,01
MFCC4_std_a	0,183	0,628	<0,01	<0,01	0,348	<0,01	0,005
MFCC5_std_a	0,231	0,738	<0,01	<0,01	0,059	<0,01	<0,01
MFCC9_std_a	0,403	0,628	<0,01	0,001	0,85	<0,01	0,003
MFCC1_D_std_a	0,336	0,838	<0,01	<0,01	0,862	<0,01	<0,01
MFCC1_DD_std_a	0,114	0,68	<0,01	<0,01	0,959	<0,01	<0,01
MFCC2_D_std_a	0,056	0,809	<0,01	<0,01	0,951	<0,01	<0,01
MFCC3_D_std_a	0,081	0,738	<0,01	<0,01	0,703	<0,01	<0,01
PLP0_std_a	0,244	0,866	<0,01	<0,01	0,167	<0,01	<0,01
PLP1_std_a	0,479	0,081	<0,01	<0,01	0,239	<0,01	<0,01
PLP5_std_a	0,473	0,777	<0,01	0,016	0,307	<0,01	0,043
PLP7_std_a	0,845	0,512	<0,01	0,004	0,621	<0,01	0,035
PLP9_std_a	0,294	0,834	<0,01	0,003	0,109	<0,01	0,002
PLP10_std_a	0,747	0,301	<0,01	0,001	0,512	<0,01	0,002
PLP12_std_a	0,38	0,642	<0,01	0,001	0,565	<0,01	0,013
PLP0_D_std_a	0,887	0,801	<0,01	<0,01	0,962	<0,01	<0,01
PLP1_D_std_a	0,695	0,129	<0,01	<0,01	0,182	<0,01	<0,01
PLP10_D_std_a	0,283	0,183	<0,01	0,019	0,981	<0,01	0,017
PLP0_DD_std_a	0,621	0,734	<0,01	<0,01	0,951	<0,01	<0,01
PLP1_DD_std_a	0,358	0,096	<0,01	<0,01	0,21	<0,01	<0,01
F0_kurtosis_a	0,81	0,878	0,014	0,006	0,884	0,027	0,01

(Continua alla pagina successiva)

(Continua dalla pagina precedente)

HNR_kurtosis_a	0,448	0,793	<0,01	<0,01	0,565	<0,01	<0,01
STE_kurtosis_a	0,299	0,192	<0,01	<0,01	0,661	<0,01	<0,01
SPEC_CENT_GRAVITY_kurtosis_a	0,259	0,444	<0,01	<0,01	0,404	<0,01	<0,01
SPEC_FLUX_kurtosis_a	0,315	0,187	<0,01	<0,01	0,661	<0,01	<0,01
SPEC_SLOPE_kurtosis_a	0,665	0,253	<0,01	<0,01	0,97	<0,01	<0,01
SPEC_DECREASE_kurtosis_a	0,337	0,061	<0,01	<0,01	0,311	<0,01	<0,01
SPEC_ROLLOFF_kurtosis_a	0,329	0,19	<0,01	<0,01	0,696	<0,01	<0,01
FORM2_kurtosis_a	0,849	0,57	0,032	0,008	0,724	0,006	0,042
FORM_BAND1_kurtosis_a	0,958	0,396	<0,01	<0,01	0,202	<0,01	<0,01
MFCC1_kurtosis_a	0,991	0,75	<0,01	<0,01	0,727	<0,01	<0,01
MFCC6_kurtosis_a	0,694	0,919	<0,01	0,024	0,832	<0,01	<0,01
MFCC8_kurtosis_a	0,963	0,178	0,017	0,01	0,888	0,013	0,006
MFCC1_D_kurtosis_a	0,484	0,62	<0,01	<0,01	0,668	<0,01	<0,01
MFCC1_DD_kurtosis_a	0,695	0,16	<0,01	<0,01	0,966	<0,01	<0,01
MFCC2_DD_kurtosis_a	0,63	0,068	<0,01	<0,01	0,763	<0,01	<0,01
MFCC3_D_kurtosis_a	0,189	0,408	0,003	<0,01	0,781	<0,01	<0,01
MFCC3_DD_kurtosis_a	0,746	0,642	0,002	<0,01	0,981	<0,01	<0,01
MFCC4_D_kurtosis_a	0,073	0,272	0,044	0,009	0,465	<0,01	<0,01
MFCC4_DD_kurtosis_a	0,311	0,259	0,004	<0,01	0,391	<0,01	<0,01
MFCC6_D_kurtosis_a	0,342	0,175	0,035	0,01	0,624	0,006	0,004
PLP0_kurtosis_a	0,445	0,858	<0,01	<0,01	0,71	<0,01	<0,01
PLP1_kurtosis_a	0,819	0,584	<0,01	<0,01	0,243	<0,01	<0,01
PLP2_kurtosis_a	0,297	0,097	0,005	0,001	0,187	<0,01	<0,01
PLP3_kurtosis_a	0,775	0,083	0,004	<0,01	0,985	<0,01	<0,01
PLP4_kurtosis_a	0,672	0,911	0,036	0,01	0,383	<0,01	<0,01
PLP5_kurtosis_a	0,154	0,75	0,041	0,036	0,217	<0,01	<0,01
PLP6_kurtosis_a	0,923	0,118	0,01	0,001	0,396	<0,01	<0,01
PLP7_kurtosis_a	0,608	0,128	<0,01	<0,01	0,741	<0,01	<0,01
PLP8_kurtosis_a	0,662	0,18	0,003	0,006	0,07	0,001	0,009
PLP9_kurtosis_a	0,429	0,051	0,041	0,009	0,283	0,013	0,003
PLP10_kurtosis_a	0,822	0,156	<0,01	<0,01	0,055	<0,01	<0,01
PLP11_kurtosis_a	0,497	0,882	<0,01	<0,01	0,876	<0,01	<0,01
PLP0_D_kurtosis_a	0,275	0,874	<0,01	<0,01	0,404	<0,01	<0,01
PLP1_D_kurtosis_a	0,841	0,264	<0,01	<0,01	0,448	<0,01	<0,01
PLP2_D_kurtosis_a	0,149	0,303	0,039	0,011	0,117	<0,01	<0,01
PLP4_D_kurtosis_a	0,781	0,121	0,006	<0,01	0,81	<0,01	<0,01
PLP5_D_kurtosis_a	0,204	0,291	0,006	0,002	0,08	0,003	0,005
PLP7_D_kurtosis_a	0,521	0,182	0,002	<0,01	0,614	<0,01	<0,01
PLP10_D_kurtosis_a	0,124	0,22	<0,01	<0,01	0,064	0,004	<0,01
PLP0_DD_kurtosis_a	0,2	0,83	<0,01	<0,01	0,205	<0,01	<0,01
PLP1_DD_kurtosis_a	0,638	0,134	<0,01	<0,01	0,509	<0,01	<0,01
PLP2_DD_kurtosis_a	0,324	0,159	0,003	<0,01	0,32	<0,01	<0,01
PLP4_DD_kurtosis_a	0,904	0,129	<0,01	<0,01	0,836	<0,01	<0,01

(Continua alla pagina successiva)

(Continua dalla pagina precedente)

PLP5_DD_kurtosis_a	0,981	0,726	<0,01	<0,01	0,134	0,006	0,01
PLP6_DD_kurtosis_a	0,54	0,068	0,026	0,01	0,675	0,002	0,001
PLP7_DD_kurtosis_a	0,858	0,919	<0,01	<0,01	0,138	<0,01	<0,01
PLP8_DD_kurtosis_a	0,865	0,18	0,014	0,003	0,378	0,002	0,006
PLP10_DD_kurtosis_a	0,091	0,253	0,003	<0,01	0,141	0,002	<0,01
PLP11_DD_kurtosis_a	0,241	0,081	0,024	0,004	0,148	<0,01	<0,01
PLP12_DD_kurtosis_a	0,616	0,074	0,004	<0,01	0,068	0,002	<0,01
HNR_skewness_a	0,559	0,584	<0,01	<0,01	0,774	<0,01	<0,01
STE_skewness_a	0,291	0,192	<0,01	<0,01	0,654	<0,01	<0,01
SPEC_CENT_GRAVITY_skewness_a	0,284	0,711	<0,01	<0,01	0,985	<0,01	<0,01
SPEC_STD_skewness_a	0,399	0,388	<0,01	<0,01	0,105	<0,01	<0,01
SPEC_SKEWNESS_skewness_a	0,065	0,695	<0,01	<0,01	0,734	<0,01	<0,01
SPEC_FLUX_skewness_a	0,315	0,187	<0,01	<0,01	0,661	<0,01	<0,01
SPEC_ROLLOFF_skewness_a	0,331	0,187	<0,01	<0,01	0,703	<0,01	<0,01
FORM1_skewness_a	0,058	0,236	<0,01	<0,01	0,08	0,005	0,011
FORM2_skewness_a	0,309	0,306	<0,01	<0,01	0,407	0,002	0,001
FORM_BAND1_skewness_a	0,688	0,532	<0,01	<0,01	0,169	<0,01	<0,01
MFCC1_skewness_a	0,781	0,715	<0,01	<0,01	0,407	<0,01	<0,01
MFCC1_D_skewness_a	0,33	0,657	0,012	0,004	0,611	0,015	0,005
MFCC3_D_skewness_a	0,133	0,695	<0,01	<0,01	0,884	0,035	0,034
PLP1_skewness_a	0,276	0,479	0,006	0,002	0,77	<0,01	<0,01
PLP2_skewness_a	0,32	0,365	<0,01	<0,01	0,929	<0,01	<0,01
PLP0_D_skewness_a	0,113	0,277	<0,01	<0,01	0,553	<0,01	<0,01
PLP2_D_skewness_a	0,267	0,354	0,004	0,002	0,09	0,001	0,003
PLP1_DD_skewness_a	0,225	0,318	<0,01	<0,01	0,491	<0,01	<0,01
FORM3_median_a	0,308	0,078	0,019	0,011	0,008	0,178	0,079
MFCC1_DD_median_a	0,838	0,915	0,03	0,006	0,018	0,075	0,187
INT_SD_a	0,058	0,253	<0,01	<0,01	0,039	<0,01	<0,01
HNR_mean_a	0,092	0,609	<0,01	<0,01	0,011	<0,01	<0,01
CPP_mean_a	0,338	0,277	<0,01	<0,01	0,028	<0,01	<0,01
FORM3_mean_a	0,142	0,178	<0,01	<0,01	0,023	0,01	0,002
CPP_median_a	0,532	0,201	<0,01	<0,01	0,049	<0,01	<0,01
GNE_median_a	0,162	0,068	<0,01	<0,01	0,002	<0,01	<0,01
MFCC7_std_a	0,214	0,83	<0,01	<0,01	0,012	<0,01	<0,01
FORM3_skewness_a	0,114	0,084	0,021	0,031	0,003	0,01	0,045
MFCC6_D_mean_a	0,673	0,007	0,635	0,806	0,16	0,731	0,776
MFCC13_DD_mean_a	0,874	0,045	0,762	0,987	0,195	0,184	0,547
PLP7_mean_a	0,056	0,002	0,923	0,812	0,139	0,494	0,291
PLP7_median_a	0,093	0,02	0,366	0,646	0,631	0,938	0,919
PLP7_DD_median_a	0,431	0,039	0,432	0,196	0,281	0,876	0,574
PLP4_skewness_a	0,053	0,047	0,673	0,712	0,992	0,543	0,93
MFCC11_mean_a	0,842	<0,01	0,092	0,068	0,073	0,007	0,001
PLP6_mean_a	0,126	0,037	0,915	0,904	0,371	0,013	0,003

(Continua alla pagina successiva)

(Continua dalla pagina precedente)

MFCC11_median_a	0,826	<0,01	0,096	0,092	0,107	0,008	0,001
PLP8_median_a	0,286	0,046	0,544	0,434	0,525	0,01	0,012
PLP4_D_median_a	0,313	0,045	0,385	0,232	0,239	0,011	0,002
MFCC8_DD_kurtosis_a	0,963	0,043	0,059	0,189	0,327	0,008	0,005
MFCC9_DD_kurtosis_a	0,951	<0,01	0,25	0,173	0,437	<0,01	<0,01
MFCC10_D_kurtosis_a	0,747	0,002	0,974	0,851	0,921	0,003	0,008
MFCC10_DD_kurtosis_a	0,611	0,006	0,592	0,32	0,644	<0,01	<0,01
MFCC11_DD_kurtosis_a	0,634	0,025	0,183	0,064	0,344	0,021	0,006
MFCC12_D_kurtosis_a	0,995	0,004	0,637	0,439	0,992	0,004	0,004
MFCC12_DD_kurtosis_a	0,282	<0,01	0,396	0,273	0,759	<0,01	<0,01
MFCC13_DD_kurtosis_a	0,22	0,02	0,863	0,922	0,644	0,005	0,003
F0_mean_a	0,577	<0,01	0,355	0,372	<0,01	0,503	0,645
MFCC6_mean_a	0,911	<0,01	0,292	0,328	<0,01	0,223	0,225
MFCC10_mean_a	0,117	<0,01	0,602	0,44	<0,01	0,962	0,945
F0_median_a	0,683	<0,01	0,317	0,285	<0,01	0,656	0,721
MFCC6_median_a	0,797	<0,01	0,409	0,411	<0,01	0,195	0,196
MFCC10_median_a	0,136	<0,01	0,586	0,42	<0,01	0,947	0,945
MFCC7_mean_a	0,743	<0,01	0,242	0,354	<0,01	<0,01	0,004
MFCC7_median_a	0,679	<0,01	0,234	0,321	<0,01	<0,01	0,007
MFCC4_D_median_a	0,725	0,005	0,054	0,015	0,738	<0,01	0,002
MFCC9_D_kurtosis_a	0,488	0,007	0,114	0,039	0,298	0,03	0,018
MFCC11_D_kurtosis_a	0,941	0,016	0,059	0,016	0,556	0,003	<0,01
PLP6_D_kurtosis_a	0,708	0,016	0,106	0,031	0,531	0,001	<0,01
MFCC12_mean_a	0,459	0,001	0,058	0,014	<0,01	0,548	0,349
MFCC12_median_a	0,448	0,003	0,06	0,017	<0,01	0,594	0,386
SPEC_STD_mean_a	0,292	0,004	<0,01	<0,01	0,989	<0,01	<0,01
SPEC_KURTOSIS_mean_a	0,141	0,004	0,001	<0,01	0,231	<0,01	<0,01
SPEC_ROLLOFF_mean_a	0,138	0,001	<0,01	<0,01	0,497	<0,01	<0,01
FORM2_mean_a	0,097	<0,01	<0,01	<0,01	0,059	<0,01	<0,01
MFCC1_mean_a	0,112	0,024	<0,01	<0,01	0,423	<0,01	<0,01
MFCC2_DD_mean_a	0,333	0,031	0,035	0,007	0,921	0,005	<0,01
PLP4_D_mean_a	0,236	0,006	<0,01	<0,01	0,494	<0,01	<0,01
SPEC_STD_median_a	0,164	0,007	<0,01	<0,01	0,503	<0,01	<0,01
SPEC_KURTOSIS_median_a	0,223	0,004	0,008	0,002	0,13	<0,01	<0,01
SPEC_ROLLOFF_median_a	0,507	<0,01	<0,01	<0,01	0,113	<0,01	<0,01
MFCC1_median_a	0,132	0,038	<0,01	<0,01	0,396	<0,01	<0,01
SPEC_CENT_GRAVITY_std_a	0,597	0,018	<0,01	<0,01	0,678	<0,01	<0,01
SPEC_ROLLOFF_std_a	0,941	0,005	<0,01	<0,01	0,756	<0,01	<0,01
FORM_BAND3_std_a	0,575	0,006	<0,01	<0,01	0,122	<0,01	0,001
MFCC13_std_a	0,094	0,007	<0,01	0,036	0,442	<0,01	<0,01
PLP11_std_a	0,388	0,022	<0,01	<0,01	0,373	<0,01	<0,01
PLP11_D_std_a	0,052	0,033	<0,01	<0,01	0,332	<0,01	0,003
PLP12_kurtosis_a	0,96	0,019	<0,01	<0,01	0,843	0,007	<0,01

(Continua alla pagina successiva)

(Continua dalla pagina precedente)

PLP3_D_kurtosis_a	0,796	0,034	0,031	0,006	0,895	<0,01	<0,01
PLP8_D_kurtosis_a	0,471	0,017	0,042	0,01	0,388	0,015	0,027
PLP9_D_kurtosis_a	0,545	0,003	0,003	<0,01	0,955	<0,01	<0,01
PLP12_D_kurtosis_a	0,66	0,008	0,027	0,006	0,244	0,004	<0,01
PLP3_DD_kurtosis_a	0,787	0,033	0,003	<0,01	0,951	<0,01	<0,01
PLP9_DD_kurtosis_a	0,42	0,015	<0,01	<0,01	0,685	<0,01	<0,01
MFCC9_mean_a	0,063	<0,01	0,018	0,006	<0,01	0,004	0,004
FORM2_median_a	0,182	<0,01	<0,01	<0,01	0,004	<0,01	<0,01
MFCC9_median_a	0,058	<0,01	0,027	0,008	<0,01	0,005	0,006
MFCC12_std_a	0,085	0,024	<0,01	<0,01	0,006	<0,01	<0,01

Tabella 6.24: feature che non presentano alcuna differenza statisticamente significativa in termini di classe

6.6.3 Set di feature

Dall'analisi statistica si individuano i 3 gruppi di feature:

1. Gruppo 1: feature con valore di $p < 0,05$ per la classe e valore di $p > 0,05$ per lingua, genere e dataset, individuate dall'analisi statistica sul dataset unito (Tabella 6.25 (a));
2. Gruppo 2: feature con valore di $p < 0,05$ per la classe e per il genere (sia soggetti MP che CS) e valore di $p > 0,05$ per lingua e dataset, individuate dall'analisi statistica sul dataset unito (Tabella 6.25 (b));
3. Gruppo 3: feature che si attivano in più dataset individuate dall'analisi statistica sui singoli dataset (Tabella 6.25 (c));

Set di feature		
MFCC4_mean_a	MFCC4_median_a	MFCC7_DD_median_a
MFCC13_D_median_a	MFCC13_DD_median_a	MFCC7_DD_skewness_a

(a) Gruppo 1

Set di feature		
JIT_L_a	JIT_LAb_a	FORM_BAND1_median_a
PLP6_median_a	MFCC5_skewness_a	

(b) Gruppo 2

Set di feature	
SPEC_ROLLOFF_skewness_a	HNR_mean_a
MFCC11_D_skewness_a	SPEC_CENT_GRAVITY_mean_a
MFCC3_mean_a	MFCC2_mean_a
PLP1_D_mean_a	PLP2_mean_a
HNR_std_a	MFCC2_median_a
SPEC_ROLLOFF_kurtosis_a	PLP1_D_std_a
HNR_skewness_a	PLP1_DD_std_a
STE_skewness_a	PLP2_DD_std_a
STE_kurtosis_a	SPEC_FLUX_kurtosis_a
SPEC_CENT_GRAVITY_std_a	FORM1_std_a
CPP_skewness_a	SPEC_FLUX_skewness_a
FORM_BAND1_std_a	

(c) Gruppo 3

Tabella 6.25: 3 gruppi di feature

In Tabella 6.26 sono raggruppati i 3 gruppi di feature individuati dall'analisi statistica, in cui si contano 34 feature in totale.

Set di feature		
MFCC4_mean_a	MFCC4_median_a	MFCC7_DD_median_a
MFCC13_D_median_a	MFCC13_DD_median_a	MFCC7_DD_skewness_a
JIT_L_a	JIT_LAb_a	FORM_BAND1_median_a
PLP6_median_a	MFCC5_skewness_a	HNR_mean_a
SPEC_CENT_GRAVITY_mean_a	MFCC2_mean_a	MFCC3_mean_a
MFCC2_median_a	HNR_std_a	SPEC_CENT_GRAVITY_std_a
FORM1_std_a	PLP1_D_std_a	PLP1_DD_std_a
PLP2_DD_std_a	STE_kurtosis_a	HNR_skewness_a
STE_skewness_a	CPP_skewness_a	SPEC_FLUX_skewness_a
PLP2_mean_a	FORM_BAND1_std_a	SPEC_FLUX_kurtosis_a
SPEC_ROLLOFF_skewness_a	MFCC11_D_skewness_a	SPEC_ROLLOFF_kurtosis_a
PLP1_D_mean_a		

Tabella 6.26: Set completo di feature robusto individuato dall'analisi statistica

6.6.4 Feature robuste e Machine learning

Si riportano i risultati dei 4 esperimenti effettuati utilizzando l'algoritmo di classificazione e i gruppi di feature individuati dall'analisi statistica:

1. in Tabella 6.27 si riportano i risultati ricavati utilizzando le feature del Gruppo 1 (Tabella 6.25 (a));
2. in Tabella 6.28 si riportano i risultati ricavati utilizzando le feature del Gruppo 1 + Gruppo 2;
3. in Tabella 6.29 si riportano i risultati ricavati utilizzando la somma delle feature del Gruppo 1 + Gruppo 2 + Gruppo 3;
4. in Tabella 6.30 si riportano i risultati ricavati utilizzando la somma delle feature del Gruppo 1 + Gruppo 2 + Gruppo 3 a cui è stato applicato l'algoritmo di feature selection Boruta.

feature	classif.	valid	test	MCC	Spec	Sens	AUC	F1.
MFCC4_mean_a	KNN	59%	53%	11%	36%	74%	55%	59%
MFCC4_median_a	LSVM	55%	45%	0%	0%	100%	50%	63%
MFCC7_DD_median_a	GPr	67%	61%	26%	48%	77%	62%	64%
MFCC13_D_median_a	DT	56%	55%	9%	55%	54%	55%	52%
MFCC13_DD_median_a	RF	62%	56%	15%	43%	71%	57%	60%
MFCC7_DD_skewness_a	NN	62%	58%	19%	48%	71%	60%	61%
	AdaB	54%	51%	2%	45%	57%	51%	51%
	NB	61%	48%	-1%	31%	69%	50%	55%
	LDA	63%	58%	18%	52%	66%	59%	59%
	XGB	62%	44%	-11%	38%	51%	45%	46%

Tabella 6.27: Risultati ottenuti utilizzando le feature del Gruppo 1

feature	classif.	valid	test	MCC	Spec	Sens	AUC	F1.
MFCC4_mean_a	KNN	60%	57%	17%	47%	71%	59%	59%
MFCC4_median_a	LSVM	55%	44%	0%	0%	100%	50%	61%
MFCC7_DD_median_a	GPr	67%	65%	29%	67%	62%	65%	61%
MFCC13_D_median_a	DT	59%	60%	19%	60%	59%	60%	56%
MFCC13_DD_median_a	RF	65%	56%	15%	44%	71%	57%	59%
MFCC7_DD_skewness_a	NN	69%	62%	24%	65%	59%	62%	58%
JIT_L_a	AdaB	60%	61%	23%	56%	68%	62%	61%
JIT_LAb_a	NB	63%	61%	20%	72%	47%	60%	52%
FORM_BAND1_median_a	LDA	67%	68%	34%	74%	59%	67%	62%
	XGB	65%	61%	25%	51%	74%	62%	63%

Tabella 6.28: Risultati ottenuti utilizzando le feature del Gruppo 2

classif.	valid	test	MCC	Spec.	Sens.	AUC	F1.
NN	70%	65%	29%	67%	62%	65%	61%
LSVM	55%	44%	0%	0%	100%	50%	61%
GPr	70%	73%	45%	77%	68%	72%	69%
DT	59%	70%	44%	58%	85%	72%	72%
RF	68%	62%	29%	49%	79%	64%	65%
NN	69%	69%	37%	72%	65%	68%	65%
AdaB	66%	73%	46%	70%	76%	73%	71%
NB	62%	69%	36%	79%	56%	67%	61%
LDA	69%	69%	38%	67%	71%	69%	67%
XGB	66%	73%	47%	67%	79%	73%	72%

Tabella 6.29: Risultati ottenuti utilizzando tutte le feature del set completo (Tabella 6.26)

34 feature con Boruta	classif.	valid	test	MCC	Spec	Sens	AUC	F1.
MFCC7_DD_median_a	KNN	71%	62%	25%	65%	60%	62%	62%
JIT_L_a	LSVM	52%	52%	0%	0%	100%	50%	68%
SPEC_CENT_GRAVITY_mean_a	GPr	70%	64%	28%	70%	58%	64%	62%
HNR_std_a	DT	73%	65%	30%	65%	65%	65%	66%
FORM1_std_a	RF	68%	56%	11%	49%	63%	56%	60%
	NN	70%	61%	22%	65%	58%	61%	61%
registrazione	AdaB	72%	60%	20%	62%	58%	60%	60%
genere	NB	67%	62%	29%	84%	43%	63%	54%
lingua	LDA	72%	64%	28%	68%	60%	64%	63%
	XGB	78%	61%	22%	57%	65%	61%	63%

Tabella 6.30: Risultati ottenuti utilizzando tutte le feature della Tabella 6.26, applicando Boruta e aggiungendo le feature *registrazione*, *lingua*, *genere*

Capitolo 7

Discussioni

Gli obiettivi dello studio affrontato nella presente tesi sono stati 2: trovare il compito vocale che permetta di ottenere la corretta discriminazione tra MP e CS mantenendo minimo lo sforzo per il paziente e trovare un subset di feature che risulti robusto nell'analisi cross-corpora. La ricerca è stata condotta utilizzando le registrazioni dei 2 compiti vocali (vocalizzi e frasi libere) eseguiti dai 176 soggetti di controllo e dai 178 pazienti con MP, acquisiti in 3 diverse lingue.

In questo capitolo si valuta quali compiti vocali consentano di centrare gli obiettivi, in base ai risultati ottenuti, e si mettono a confronto le prestazioni dei modelli migliori con i modelli presenti in letteratura.

7.1 Registrazione

Sono stati uniti due dataset registrati rispettivamente con un microfono professionale (Italiano 1) e con uno smartphone (Privato 1), per valutare se l'accuratezza sul validation aumenti, rimanga uguale o diminuisca quando si registra con dispositivi non professionali rispetto alle prestazioni ottenute con microfono professionale.

I risultati del vocalizzo /a/ sono riportati in Tabella 6.1, delle frasi per il dataset Italiano 1 in Tabella 6.2 e delle frasi del dataset Italiano1 + Privato 1 in Tabella 6.3. Sul vocalizzo, si parte da un'accuratezza su Italiano 1 del 95% e dopo l'unione, l'accuratezza massima scende al 79%. In grassetto sono evidenziati i valori di accuratezza media più alti. Sulle frasi l'accuratezza massima del 98% sul solo Italiano 1, dopo l'unione diminuisce all'86%. In grassetto è evidenziata la frase con l'accuratezza media più alta su tutte e 4 le prove. Mettendo insieme i dataset vi è una riduzione delle prestazioni:

- l'accuratezza diminuisce dal 95% al 79% nel vocalizzo,
- l'accuratezza diminuisce dal 98% all'86% nelle frasi.

Pur verificandosi una diminuzione delle prestazioni unendo i due dataset, per avere un dataset di dimensione maggiormente significativa, si è preferito considerare il dataset

unito, bilanciando opportunamente trainset e testset, per cui, nelle sezioni successive, i due dataset verranno considerati uniti. Si evince, come atteso, che l’inserimento di campioni audio registrati in condizioni non ottimali porti ad un peggioramento delle prestazioni del modello.

7.2 Scelta vocale

Con la vocale /a/ sia il dataset Italiano1 che il dataset Colombiano ottengono valori sul validation maggiori rispetto al vocalizzo /e/. Dalla Tabella 6.4 si evince che i valori di accuratezza sono 95% per il dataset Italiano 1 e di 79% per il dataset Colombiano utilizzando il vocalizzo /a/; mentre utilizzando il vocalizzo /e/, i valori di accuratezza sono 94% per il dataset Italiano 1 e di 70% per il dataset Colombiano. In grassetto sono evidenziati i valori di accuratezza media più alti. I risultati delle prestazioni con le due vocali si discostano al più del 9%, per cui non vi è un’elevata differenza ma, tenendo conto anche della disponibilità dei dataset online (la vocale /a/ è quella più utilizzata), la vocale scelta per il compito vocale della fonazione sostenuta è la /a/.

7.3 Scelta frasi

Tra le frasi che presentano le prestazioni maggiori, non si notano elevate differenze, riportando al più una differenza del 4%.

Italiano 1 + Privato 1

La frase con i valori di accuratezza sul validation complessivamente maggiori è la frase 3, con un valore massimo di 86%. La seconda frase con i valori sul validation complessivamente maggiori è la frase 4, con un valore massimo di 83% (Tabella 6.5). In grassetto sono evidenziate le frasi scelte.

Colombiano

La frase con i valori sul validation complessivamente maggiori è la frase 1, con un valore massimo di 81%, (in accordo con Orozco-Arroyave, J R et al. [20]). La seconda frase con i valori sul validation complessivamente maggiori è la frase 5, con un valore massimo di 82% (Tabella 6.6). In grassetto sono evidenziate le frasi scelte.

Italiano 2

La frase con i valori sul validation complessivamente maggiori è la frase 2, con un valore massimo di 89%. La seconda frase con i valori sul validation complessivamente maggiori è la frase 1, con un valore massimo di 87%, (Tabella 6.7). In grassetto sono evidenziate le frasi scelte.

7.4 Protocollo

Dato che l'utilizzo di 2 frasi non comporta un significativo incremento delle prestazioni (Tabella 6.10, Tabella 6.14, Tabella 6.18), si individua nel **vocalizzo unito ad una frase** il compito vocale vincente per discriminare al meglio le registrazioni dei pazienti MP dalle registrazioni dei soggetti di controllo. Purché si scelga una frase foneticamente bilanciata, non è necessario introdurre all'interno del protocollo compiti più difficili e faticosi per ottenere buone prestazioni.

Osservando le feature selezionate dagli algoritmi di feature selection Boruta e Crosscorrelazione, si nota come l'algoritmo di Cross_correlazione tenda a selezionare nettamente più feature rispetto all'algoritmo Boruta e, sommariamente, ottenga risultati peggiori.

Secondo Qiyue Wang e al. [30], il compito delle frasi è più performante del compito del vocalizzo, in quanto le feature articolatorie associate alle frasi riflettono meglio le caratteristiche fonetiche della lingua. Tra le feature in questione rientrano mean_MFCC3, mean_MFCC4, mean_MFCC3, mean_F1 (prima formante), mean_F2 (seconda formante), std_F1, mean_MFCC6. Solo sul vocalizzo /a/ Nemuel D. Pah e al. [27], con SVM come classificatore e selezionando le feature base quali Jitt(abs), Jitt(rel), Shim(abs), Shim(rel), SD(pitch), HNR, NHR, raggiungono un'accuratezza del 59% utilizzando il dataset Colombiano. Le frasi ottengono risultati migliori, ma il vocalizzo è il compito vocale più utilizzato e semplice da riprodurre.

I risultati attestano che per i dataset Italiano 1 + Privato 1 e Colombiano, il compito vocale più performante è il vocalizzo unito ad una frase mentre per il dataset Italiano 2, il compito vocale più performante è il vocalizzo unito a due frasi. Tenendo conto anche del tempo di esecuzione del compito vocale, si è scelto come prova ottimale il vocalizzo unito ad una frase.

Con particolare attenzione ai risultati ottenuti, si possono identificare:

1. Artificial Neural Network come classificatore più performante con l'utilizzo del vocalizzo unito ad una frase;
2. Boruta come algoritmo di feature selection che individua le feature utili alla discriminazione;
3. min_max come algoritmo di normalizzazione dei dataset.

Italiano 1 + Privato 1

Il compito vocale che raggiunge i valori di accuratezza migliori è l'unione del vocalizzo /a/ ad una frase. Utilizzando come classificatore l'Artificial Neural Network o Gaussian Process, l'algoritmo di feature selection Boruta e normalizzazione minimo_massimo, si raggiunge il valore di accuratezza di 96% sul validation e di 83% sul test. Sono riportati in Tabella 6.10 i risultati delle metriche di valutazione e in Figura 6.7 la lista di feature selezionate.

Colombiano

Il compito vocale che raggiunge i valori di accuratezza migliori è l'unione del vocalizzo /a/ ad una frase. Utilizzando come classificatore Gaussian Process, l'algoritmo di feature selection Boruta e normalizzazione `z_score`, si raggiunge il valore di accuratezza di 91% sul validation e di 73% sul test. Sono riportati in Tabella 6.14 i risultati delle metriche di valutazione e in Figura 6.11 la lista di feature selezionate.

Italiano 2

Il compito vocale che raggiunge i valori di accuratezza migliori è l'unione del vocalizzo /e/ a 2 frasi. Utilizzando come classificatore l'Artificial Neural Network, l'algoritmo di feature selection Boruta e normalizzazione `min_max`, si raggiunge il valore di accuratezza di 93% sul validation e di 90% sul test. Sono riportati in Tabella 6.19 i risultati delle metriche di valutazione e in Figura 6.16 la lista di feature selezionate.

7.5 Machine learning per l'analisi multilingua

Considerando l'unione dei dataset Italiano 1+ Privato 1+ Colombiano + Ceco + Privato 2 sul vocalizzo /a/, le prestazioni diminuiscono in confronto ai risultati ottenuti sui singoli dataset.

Utilizzando come classificatore Gaussian Process, l'algoritmo di feature selection Boruta e normalizzazione `z_score`, si raggiunge il valore di accuratezza di 76% sul validation e di 62% sul test. In Tabella 6.21 sono riportati i risultati delle metriche di valutazione Accuratezza (sul validation e sul test), Specificità, Sensibilità, AUC e F1_score (sul test) e in Tabella 6.20 la lista di feature selezionate.

7.6 Feature robuste e Machine learning

Dall'analisi statistica, sono stati individuati 3 gruppi di feature, Tabella 6.25, i quali, secondo i risultati dei test statistici di Mann-Whitney e Kruskal-Wallis, sono:

1. Gruppo 1: feature robuste in termini di dataset, lingua e genere (valore di $p > 0,05$), che presentano una differenza statisticamente significativa in termini di classe (valore di $p < 0,05$), Tabella 6.25 (a);
2. Gruppo 2: feature robuste in termini di dataset e lingua (valore di $p > 0,05$), che presentano una differenza statisticamente significativa in termini di classe e genere (valore di $p < 0,05$), Tabella 6.25 (b);
3. Gruppo 3: feature che si attivano in più dataset individuate dall'analisi statistica sui singoli dataset, Tabella 6.25 (c).

Di seguito, si elencano le 4 prove effettuate applicando l'algoritmo di classificazione al dataset unito e normalizzato con l'algoritmo `min_max`:

1. la prova 1 considera le feature del Gruppo 1. Il risultato migliore, Tabella 6.27, è stato ottenuto dal classificatore Gaussian Process con un'accuratezza sul validation di 67% e di 61% sul test;
2. la prova 2 considera le feature del Gruppo 1 + Gruppo 2. Il risultato migliore, Tabella 6.28, è stato ottenuto dal classificatore Neural Net con un'accuratezza sul validation di 69% e di 62% sul test;
3. la prova 3 considera le feature del Gruppo 1 + Gruppo 2 + Gruppo 3 (34 feature in totale). Il risultato migliore, Tabella 6.29, è stato ottenuto dal classificatore Gaussian Process con un'accuratezza sul validation di 70% e di 73% sul test;
4. la prova 4 considera le feature del Gruppo 1 + Gruppo 2 + Gruppo 3, applica l'algoritmo di feature selection Boruta e, alle 34 feature, somma le caratteristiche *registrazione*, *lingua*, *genere*. Il risultato migliore, Tabella 6.29, è stato ottenuto dal classificatore XGBoost con un'accuratezza sul validation del 78% e del 61% sul test.

Mettendo a confronto i risultati, la combinazione tra Boruta, le 34 feature individuate dall'analisi statistica e l'aggiunta delle feature *registrazione*, *lingua*, *genere* risulta la più performante. Il risultato ottenuto con le feature individuate dall'analisi statistica supera il risultato della prova di Machine learning. Il classificatore che ha ottenuto i risultati migliori è XGBoost.

Risultati simili si trovano nello studio condotto da Daniel Kovac et al. [31], l'unico studio nel quale viene svolta un'analisi analoga a quella condotta in questa tesi. Come compiti vocali hanno considerato il vocalizzo /a/, lettura di un testo e la ripetizione delle sillabe /pa/ /ta/ /ka/. Il database che gli autori hanno utilizzato è composto da 5 dataset in 4 lingue diverse (italiano, israeliano, inglese e spagnolo), composto da 241 MP e 265 CS. Nello studio è stata prima condotta un'analisi statistica per individuare le feature con lo stesso comportamento in tutte le lingue, poi è stato utilizzato questo set di feature per discriminare soggetti MP da CS, utilizzando un classificatore XGB. Il protocollo utilizzato nello studio ha previsto una prima fase di analisi su un singolo corpus ed una seconda fase di validazione dei risultati attraverso due tecniche: la prima è "Leave one corpus out" (in cui si utilizzano tutti i dataset meno uno per allenare il modello e il dataset rimanente per il test), la seconda è "Cross-language validation" (in cui si utilizza un dataset per allenare il modello e gli altri dataset per il test).

Prendendo in considerazione l'esperimento condotto in cui gli autori valutano le prestazioni del modello su tutti i dataset uniti, l'accuratezza risulta essere pari a $75\% \pm 6\%$, valore confrontabile con lo studio in questione. Il set di feature individuato è composto da: HRF (fattore di ricchezza armonica), NAQ (quoziente di ampiezza normalizzato medio), realF0SD (deviazione standard della frequenza fondamentale rispetto alla sua media) e jitter(PPQ). Le feature con lo stesso comportamento per lingua italiana e spagnola sono: relF1SD (deviazione standard della prima frequenza formante) e HNR (rapporto tra armoniche e rumore). I risultati sul testset non stati messi a confronto in quanto manca la presenza di un vero testset.

Nello studio di Toye et al.[71] si è dimostrata l'importanza dell'utilizzo dei coefficienti

MFCC nella discriminazione dei soggetti MP da CS. Per l'analisi sono stati utilizzati un dataset in lingua italiana (37 CS, 28 MP) e un dataset inglese (16 MP and 21 CS), considerando il compito vocale della lettura del testo. Dopo aver individuato 3 gruppi di feature diversi (gruppo 1: varianti jitter, varianti shimmer, F0 e HNR; gruppo 2: 13 coefficienti MCC; gruppo 3: unione dei 2 gruppi e selezione delle 10 feature migliori, sono stati testati i 3 gruppi su 7 diversi classificatori (KNN, DT, SVM, NB, Linear Regression (LR), GB e RF). I risultati ottenuti portano alla conclusione che la combinazione di feature acustiche classiche e coefficienti MFCC porta a risultati migliori.

7.7 Limitazioni

Per l'analisi multilingua condotta, sono state utilizzate le registrazioni solo della fonazione sostenuta della vocale /a/, poiché non tutti i dataset utilizzati possedevano anche il compito delle frasi. Il compito del vocalizzo può valutare correttamente le alterazioni del linguaggio associate alla malattia di Parkinson. Tuttavia, come si evince dai risultati nella sezione di machine learning, il compito delle frasi porta un importante contributo nell'analisi acustica. Inoltre, per ogni confronto (ex. lingua) si considerano trascurabili le caratteristiche dei soggetti che compongono i vari gruppi (ex. altre malattie presenti, dispositivi e ambienti di registrazione diversi). Benché attualmente questo sia il meglio che si possa fare, sicuramente in futuro sarà da analizzare meglio.

La composizione del dataset è un altro fattore che limita la generalizzabilità dei risultati. In studi futuri abbiamo in programma di aumentare la dimensione del dataset, anche considerando altre lingue, e validare i risultati ottenuti. L'analisi è stata attualmente svolta solo sulla classificazione binaria ma, in futuro, sarà necessario estendere i risultati alla valutazione multiclasse, includendo il grado di severità della patologia.

Capitolo 8

Conclusioni

In questa tesi sono stati analizzati due fattori fondamentali dell'analisi acustica per il riconoscimento della Malattia di Parkinson: i compiti vocali più performanti per la classificazione automatica dei soggetti in MP e in controlli sani CS, e le feature più utili e robuste per la classificazione.

Con riferimento ai compiti vocali, sono stati messi a confronto la fonazione della vocale sostenuta e la lettura di frasi, per determinare quale sia la prova che al contempo riesca a discriminare al meglio le due classi di soggetti e che tenga conto della possibilità di affaticamento dei pazienti. Considerando questi ultimi aspetti, il risultato ottenuto da questo studio risulta essere molto vantaggioso. Infatti, l'unione del vocalizzo e di una sola frase permette di ottenere delle prestazioni ottimali in termini di discriminazione e l'esecuzione dei compiti risulta breve e non eccessivamente difficoltosa. L'esito del compito vocale è stato ottenuto e validato su 3 diversi dataset, due dei quali in lingua italiana e uno in lingua spagnola, opportunamente bilanciati in termini di numerosità di MP, CS e genere.

In relazione alle feature, invece, l'analisi eseguita è stata statistica e ha individuato il subset di feature, tra le 498 feature di partenza, che porta al minimo la possibilità di essere influenzato da fattori esterni, quali lo strumento di registrazione utilizzato e la lingua natia dei soggetti. Con particolare riferimento alla lingua, il subset di feature è stato testato su un dataset molto ampio in termini di numerosità e formato dall'unione delle registrazioni in lingua italiana, spagnola e ceca. Tra le feature che compongono il subset troviamo la mediana di MFCC7, la deviazione standard della prima formante, la deviazione standard di HNR, jitter. Informazioni riguardo la registrazione, il genere e la lingua, portano ad un miglioramento delle prestazioni.

In sintesi, la tesi ha definito una soluzione che può essere clinicamente applicabile per il monitoraggio della Malattia di Parkinson, tramite analisi acustica, e contribuisce all'esplorazione delle feature acustiche, utilizzando sia la statistiche che il machine learning, che racchiudono un grande potenziale per delle possibili applicazioni di monitoraggio dei pazienti con Malattia di Parkinson.

Bibliografia

- [1] Parkinson DISEASE Werner Poewe et al. *Nature Reviews Disease Primers* (cit. alle pp. 1, 3).
- [2] Lorraine V. Kalia e Anthony E. Lang. *Parkinson's disease*. Ago. 2015. DOI: 10.1016/S0140-6736(14)61393-3 (cit. a p. 1).
- [3] Hagai Bergman e Günther Deuschl. *Pathophysiology of parkinson's disease: From clinical neurology to basic neuroscience and back*. 2002. DOI: 10.1002/mds.10140 (cit. a p. 1).
- [4] Bastiaan R. Bloem, Michael S. Okun e Christine Klein. *Parkinson's disease*. Giu. 2021. DOI: 10.1016/S0140-6736(21)00218-X (cit. a p. 1).
- [5] *Projected number of people with Parkinson disease in the most populous nations, 2005 through 2030*. 2007. URL: www.neurology.org (cit. a p. 1).
- [6] Nicola Vanacore e Letterio Morgante. *EPIDEMIOLOGIA DEI PARKINSONISMI* (cit. a p. 1).
- [7] J. Jankovic. *Parkinson's disease: Clinical features and diagnosis*. 2008. DOI: 10.1136/jnnp.2007.131045 (cit. a p. 1).
- [8] *The Unified Parkinson's Disease Rating Scale (UPDRS): Status and recommendations*. Lug. 2003. DOI: 10.1002/mds.10473 (cit. a p. 1).
- [9] Zoltan Galaz et al. «Changes in phonation and their relations with progress of Parkinson's disease». In: *Applied Sciences (Switzerland)* 8.12 (nov. 2018). ISSN: 20763417. DOI: 10.3390/app8122339 (cit. a p. 2).
- [10] Lorraine O. Ramig, Cynthia Fox e Shimon Sapir. *Speech treatment for Parkinson's disease*. Feb. 2008. DOI: 10.1586/14737175.8.2.297 (cit. a p. 2).
- [11] William O'Grady, Michael Dobrovolsky e Francis Katamba. *Contemporary linguistics*. St. Martin's, 1997 (cit. a p. 2).
- [12] treccani. In: (). URL: <https://www.treccani.it> (cit. a p. 2).

-
- [13] Andrew Ma, Kenneth K. Lau e Dominic Thyagarajan. *Voice changes in Parkinson's disease: What are they telling us?* Feb. 2020. DOI: 10.1016/j.jocn.2019.12.029 (cit. alle pp. 3, 4).
- [14] Tomi Kinnunen. «Spectral Features for Automatic Text-Independent Speaker Recognition». In: (set. 2004) (cit. a p. 3).
- [15] Petr Krýže, Tereza Tykalová, Evžen Růžička e Jan Rusz. «Effect of reading passage length on quantitative acoustic speech assessment in Czech-speaking individuals with Parkinson's disease treated with subthalamic nucleus deep brain stimulation». In: *The Journal of the Acoustical Society of America* 149.5 (mag. 2021), pp. 3366–3374. ISSN: 0001-4966. DOI: 10.1121/10.0005050 (cit. a p. 3).
- [16] Elsa Arocho-Quinones, Michael Hammer, Jonathan Bock e Peter Pahapill. «Effects of deep brain stimulation on vocal fold immobility in Parkinson's disease». In: *Surgical Neurology International* 8.1 (2017). ISSN: 21527806. DOI: 10.4103/2152-7806.200580 (cit. a p. 4).
- [17] Ana Paula Zarzur, André C. Duprat, Gilson Shinzato e Claudia A. Eckley. «Laryngeal electromyography in adults with Parkinson's disease and voice complaints». In: *Laryngoscope* 117.5 (mag. 2007), pp. 831–834. ISSN: 0023852X. DOI: 10.1097/MLG.0b013e3180333145 (cit. a p. 4).
- [18] Sergio Motta, Ugo Cesari, Mariano Paternoster, Giovanni Motta e Giuseppe Orefice. «Aerodynamic findings and Voice Handicap Index in Parkinson's disease». In: *European Archives of Oto-Rhino-Laryngology* 275.6 (giu. 2018), pp. 1569–1577. ISSN: 14344726. DOI: 10.1007/s00405-018-4967-7 (cit. a p. 4).
- [19] Pascal Hecker, Nico Steckhan, Florian Eyben, Björn W. Schuller e Bert Arnrich. *Voice Analysis for Neurological Disorder Recognition—A Systematic Review and Perspective on Emerging Trends*. Lug. 2022. DOI: 10.3389/fdgth.2022.842301 (cit. alle pp. 6, 12).
- [20] J. R. Orozco-Arroyave, F. Hönig, J. D. Arias-Londoño, J. F. Vargas-Bonilla, K. Daqrouq, S. Skodda, J. Rusz e E. Nöth. «Automatic detection of Parkinson's disease in running speech spoken in three different languages». In: *The Journal of the Acoustical Society of America* 139 (1 gen. 2016), pp. 481–500. ISSN: 0001-4966. DOI: 10.1121/1.4939739 (cit. alle pp. 6, 97).
- [21] Quoc Cuong Ngo, Mohammad Abdul Motin, Nemuel Daniel Pah, Peter Drotár, Peter Kempster e Dinesh Kumar. *Computerized analysis of speech and voice for Parkinson's disease: A systematic review*. Nov. 2022. DOI: 10.1016/j.cmpb.2022.107133 (cit. alle pp. 6, 20).
- [22] Vijay Parsa e Donald G Jamieson. *08Parsa*. 2001, pp. 327–339 (cit. a p. 7).

-
- [23] A Askenfelt e B Hammarberg. *Dept. for Speech, Music and Hearing Quarterly Progress and Status Report Speech waveform perturbation analysis*. URL: <http://www.speech.kth.se/qpsr> (cit. a p. 7).
- [24] J. Rusz, R. Cmejla, H. Ruzickova e E. Ruzicka. «Quantitative acoustic measurements for characterization of speech and voice disorders in early untreated Parkinson's disease». In: *The Journal of the Acoustical Society of America* 129 (1 gen. 2011), pp. 350–367. ISSN: 0001-4966. DOI: 10.1121/1.3514381 (cit. a p. 7).
- [25] Biswajit Karan, Sitanshu Sekhar Sahu, Juan Rafael Orozco-Arroyave e Kartik Mahto. «Non-negative matrix factorization-based time-frequency feature extraction of voice signal for Parkinson's disease prediction». In: *Computer Speech and Language* 69 (set. 2021). ISSN: 10958363. DOI: 10.1016/j.cs1.2021.101216 (cit. a p. 7).
- [26] Jan Rusz et al. «Speech Biomarkers in Rapid Eye Movement Sleep Behavior Disorder and Parkinson Disease». In: *Annals of Neurology* 90.1 (lug. 2021), pp. 62–75. ISSN: 15318249. DOI: 10.1002/ana.26085 (cit. a p. 7).
- [27] Nemuel D. Pah, Mohammad A. Motin e Dinesh K. Kumar. «Phonemes based detection of parkinson's disease for telehealth applications». In: *Scientific Reports* 12.1 (dic. 2022). ISSN: 20452322. DOI: 10.1038/s41598-022-13865-z (cit. alle pp. 7, 8, 98).
- [28] Rohit Lamba, Tarun Gulati, Anurag Jain e Pooja Rani. «A Speech-Based Hybrid Decision Support System for Early Detection of Parkinson's Disease». In: *Arabian Journal for Science and Engineering* (2022). ISSN: 21914281. DOI: 10.1007/s13369-022-07249-8 (cit. a p. 8).
- [29] Adedolapo Aishat Toyé e Suryaprakash Kompalli. «Comparative Study of Speech Analysis Methods to Predict Parkinson's Disease». In: (nov. 2021). URL: <http://arxiv.org/abs/2111.10207> (cit. a p. 8).
- [30] Qiyue Wang, Yan Fu, Baiyu Shao, Le Chang, Kang Ren, Zhonglue Chen e Yun Ling. «Early detection of Parkinson's disease from multiple signal speech: Based on Mandarin language dataset». In: *Frontiers in Aging Neuroscience* 14 (nov. 2022). ISSN: 16634365. DOI: 10.3389/fnagi.2022.1036588 (cit. alle pp. 8, 98).
- [31] Daniel Kovac, Jiri Mekyska, Vered Aharonson, Pavol Harar, Zoltan Galaz, Steven Rapcsak, Juan Rafael Orozco-Arroyave, Lubos Brabenec e Irena Rektorova. «Exploring Language-independent Digital Speech Biomarkers of Hypokinetic Dysarthria». In: (). DOI: 10.1101/2022.10.24.22281459. URL: <https://doi.org/10.1101/2022.10.24.22281459> (cit. alle pp. 8, 100).

- [32] Giovanni Dimauro, Vincenzo Di Nicola, Vitoantonio Bevilacqua, Danilo Caivano e Francesco Girardi. «Assessment of speech intelligibility in Parkinson's disease using a speech-to-text system». In: *IEEE Access* 5 (ott. 2017), pp. 22199–22208. ISSN: 21693536. DOI: 10.1109/ACCESS.2017.2762475 (cit. a p. 10).
- [33] J R Orozco-Arroyave, F Hönig, J D Arias-Londoño, J F Vargas-Bonilla, S Skodda, J Rusz e E Nöth. *Voiced/Unvoiced Transitions in Speech as a Potential Bio-Marker to Detect Parkinson's Disease*. Rapp. tecn. 2015 (cit. a p. 10).
- [34] Jan Hlavnicka, Roman Cmejla, Jiri Klempir, Evzen Ruzicka e Jan Rusz. «Acoustic tracking of pitch, modal, and subharmonic vibrations of vocal folds in Parkinson's disease and Parkinsonism». In: *IEEE Access* 7 (2019), pp. 150339–150354. ISSN: 21693536. DOI: 10.1109/ACCESS.2019.2945874 (cit. a p. 10).
- [35] Laureano Moro-Velazquez, Jorge A. Gomez-Garcia, Julian D. Arias-Londoño, Najim Dehak e Juan I. Godino-Llorente. «Advances in Parkinson's Disease detection and assessment using voice and speech: A review of the articulatory and phonatory aspects». In: *Biomedical Signal Processing and Control* 66 (apr. 2021). ISSN: 17468108. DOI: 10.1016/j.bspc.2021.102418 (cit. a p. 12).
- [36] David Weenink Paul Boersma. «Praat: doing phonetics by computer». In: (2022). URL: <https://www.praat.org> (cit. a p. 13).
- [37] «faluschi_tesi». In: () (cit. a p. 13).
- [38] Yasuhiro Tanaka, Masaki Nishio e Seiji Niimi. «Vocal acoustic characteristics of patients with parkinson's disease». In: *Folia Phoniatrica et Logopaedica* 63.5 (ago. 2011), pp. 223–230. ISSN: 10217762. DOI: 10.1159/000322059 (cit. a p. 14).
- [39] «Fonetica acustica MINISSI ROMANO RIVOIRA Manuale di Fonetica IV parte». In: () (cit. a p. 14).
- [40] J. A. Gómez-García, L. Moro-Velázquez, J. D. Arias-Londoño e J. I. Godino-Llorente. «On the design of automatic voice condition analysis systems. Part III: review of acoustic modelling strategies». In: *Biomedical Signal Processing and Control* 66 (apr. 2021). ISSN: 17468108. DOI: 10.1016/j.bspc.2020.102049 (cit. alle pp. 14–17).
- [41] João Paulo Teixeira, Carla Oliveira e Carla Lopes. «Vocal Acoustic Analysis – Jitter, Shimmer and HNR Parameters». In: *Procedia Technology* 9 (2013), pp. 1112–1122. ISSN: 22120173. DOI: 10.1016/j.protcy.2013.12.124 (cit. a p. 15).

- [42] Achraf Benba, Abdelilah Jilbab e Ahmed Hammouch. «Detecting patients with Parkinson's disease using mel frequency cepstral coefficients and support vector machines». In: *International Journal on Electrical Engineering and Informatics* 7.2 (lug. 2015), pp. 297–307. ISSN: 20875886. DOI: 10.15676/ijeei.2015.7.2.10 (cit. a p. 16).
- [43] Christopher R. Watts e Shaheen N. Awan. «Use of spectral/cepstral analyses for differentiating normal from hypofunctional voices in sustained vowel and continuous speech contexts». In: *Journal of Speech, Language, and Hearing Research* 54.6 (dic. 2011), pp. 1525–1537. ISSN: 10924388. DOI: 10.1044/1092-4388(2011/10-0209) (cit. a p. 16).
- [44] D Michaelis, T Gramss e H W Strube. *Glottal-to-Noise Excitation Ratio-a New Measure for Describing Pathological Voices*. Rapp. tecn. (cit. a p. 17).
- [45] «parselmouth». In: (). URL: <https://parselmouth.readthedocs.io/en/stable/> (cit. a p. 17).
- [46] Ravil I Muhamedyev, Kazakhstan Pushkina e Almaty Kazakhstan. *Machine learning methods: An overview 1 Introduction 2 Machine learning techniques 2.1 TYPES OF MACHINE LEARNING ALGORITHMS 2.2 SETUP OF MACHINE LEARNING SYSTEMS 2.3 FORMAL DEFINITION OF A MACHINE LEARNING PROBLEM 2.4 REGRESSIVE ALGORITHMS AND DATA CLASSIFICATION ALGORITHMS 3 Quality assessment of machine learning systems*. 2015, pp. 14–29. URL: www.cmnt.lv (cit. a p. 18).
- [47] Jiliang Tang, Salem Alelyani e Huan Liu. *Feature Selection for Classification: A Review* (cit. a p. 19).
- [48] «scikit-learn». In: (). URL: https://scikit-learn.org/stable/modules/feature_selection.html (cit. a p. 20).
- [49] Amit Kumar Patra, Ratula Ray, Azian Azamimi Abdullah e Satya Ranjan Dash. «Prediction of Parkinson's disease using Ensemble Machine Learning classification from acoustic analysis». In: vol. 1372. Institute of Physics Publishing, nov. 2019. DOI: 10.1088/1742-6596/1372/1/012041 (cit. a p. 20).
- [50] «KNN». In: (). URL: <https://scikit-learn.org/stable/modules/neighbors.html#classification> (cit. a p. 20).
- [51] «SVM». In: (). URL: <https://scikit-learn.org/stable/modules/svm.html#classification> (cit. a p. 20).
- [52] «Gaussianprocess». In: (). URL: https://scikit-learn.org/stable/modules/gaussian_process.html#gaussian-process (cit. a p. 21).
- [53] «decisiontree». In: (). URL: <https://scikit-learn.org/stable/modules/tree.html> (cit. a p. 21).

- [54] «RandomForest». In: (). URL: <https://scikit-learn.org/stable/modules/ensemble.html#forest> (cit. a p. 21).
- [55] «AdaBoost». In: (). URL: <https://scikit-learn.org/stable/modules/generated/sklearn.ensemble.AdaBoostClassifier.html> (cit. a p. 21).
- [56] «GNaivebayes». In: (). URL: https://scikit-learn.org/stable/modules/naive_bayes.html#gaussian-naive-bayes (cit. a p. 21).
- [57] «LDA». In: (). URL: https://scikit-learn.org/stable/modules/lda_qda.html#lda-qda (cit. a p. 21).
- [58] «GradientBoosting». In: (). URL: <https://scikit-learn.org/stable/modules/generated/sklearn.ensemble.GradientBoostingClassifier.html> (cit. a p. 21).
- [59] «NeuralNet». In: (). URL: <https://python-course.eu/machine-learning/neural-networks-with-scikit.php> (cit. a p. 22).
- [60] «Crossvalidation». In: (). URL: https://scikit-learn.org/stable/modules/cross_validation.html (cit. a p. 22).
- [61] Betul Erdogdu Sakar, M. Erdem Isenkul, C. Okan Sakar, Ahmet Sertbas, Fikret Gorgen, Sakir Delil, Hulya Apaydin e Olcay Kursun. «Collection and analysis of a Parkinson speech dataset with multiple types of sound recordings». In: *IEEE Journal of Biomedical and Health Informatics* 17.4 (2013), pp. 828–834. ISSN: 21682194. DOI: 10.1109/JBHI.2013.2245674 (cit. a p. 22).
- [62] Andrew E Bradley. *THE EVALUATION OF MACHINE LEARNING ALGORITHMS*. 1997, pp. 1145–1159 (cit. a p. 23).
- [63] «F1score». In: (). URL: https://scikit-learn.org/stable/modules/generated/sklearn.metrics.f1_score.html (cit. a p. 23).
- [64] S S Shapiro e M B Wilk. *Biometrika Trust An Analysis of Variance Test for Normality (Complete Samples)*. 1965, pp. 591–611 (cit. a p. 23).
- [65] H B Mann e D R Whitney. *On a Test of Whether one of Two Random Variables is Stochastically Larger than the Other*. 1947, pp. 50–60 (cit. a p. 23).
- [66] William H Kruskal e ; W Allen Wallis. *Use of Ranks in One-Criterion Variance Analysis*. 1952, pp. 583–621 (cit. a p. 23).
- [67] «The jamovi project (2022)[Computer Software]». In: (2022). URL: <https://www.jamovi.org> (cit. a p. 24).
- [68] *Mann-Whitney U test (Non-parametric equivalent to independent samples t-test)*. URL: <https://maths.shu.ac.uk/mathshelp/> (cit. a p. 24).

- [69] T. Van Hecke. «Power study of anova versus Kruskal-Wallis test». In: *Journal of Statistics and Management Systems* 15 (2-3 mag. 2012), pp. 241–247. ISSN: 0972-0510. DOI: 10.1080/09720510.2012.10701623 (cit. a p. 25).
- [70] Federica Amato, Luigi Borzì, Gabriella Olmo e Juan Rafael Orozco-Arroyave. «An algorithm for Parkinson’s disease speech classification based on isolated words analysis». In: *Health Information Science and Systems* 9 (1 dic. 2021). ISSN: 20472501. DOI: 10.1007/s13755-021-00162-8 (cit. a p. 30).
- [71] Adedolapo Aishat Toyé e Suryaprakash Kompalli. «Comparative Study of Speech Analysis Methods to Predict Parkinson’s Disease». In: *arXiv preprint arXiv:2111.10207* (2021) (cit. a p. 100).