



UNIVERSITÀ  
DEGLI STUDI  
DI PADOVA



DIPARTIMENTO  
DI INGEGNERIA  
DELL'INFORMAZIONE

UNIVERSITA' DEGLI STUDI DI PADOVA  
DIPARTIMENTO DI INGEGNERIA DELL'INFORMAZIONE  
CORSO DI LAUREA MAGISTRALE IN BIOINGEGNERIA

# **Screening e valutazione otoneurologica del parlato usando Frequency Following Response e metodi di machine learning**

Relatore: **Prof. Giovanni Sparacino**

Laureanda: **Laura Pertegato**  
Matricola n.: **2095572**

Correlatore: **Ing. Giulio Dehesh**

**ANNO ACCADEMICO 2025 – 2026**

**Data di laurea 09/04/2026**



# Indice

<b>1. Frequency Following Response: aspetti teorici, metodologici ed applicativi.....</b>	<b>1</b>
1.1 Il sistema uditivo .....	1
1.2 Potenziali evocati uditivi e Frequency Following Response (FFR).....	3
1.2.1 Origine neurofisiologica delle FFR.....	3
1.2.2 Morfologia delle FFR.....	4
1.2.3 Proprietà funzionali .....	6
1.3 Protocollo standard di acquisizione di FFR.....	7
1.3.1 Stimolazione .....	7
1.3.2 Registrazione .....	8
1.3.3 Pre-processing .....	8
1.4 Applicazione clinica e diagnostica delle FFR .....	9
1.5 Estrazione delle FFR .....	10
<b>2. Estrazione ed analisi delle FFR .....</b>	<b>11</b>
2.1 Metodi di letteratura per l'estrazione delle FFR.....	11
2.1.1 Media convenzionale (CA).....	11
2.1.2 Filtro alla Wiener a Posteriori (APWF).....	11
2.1.3 SubBand Filter (SB) .....	13
2.1.4 Metodi a confronto .....	14
2.2 Indicatori proposti in letteratura per la quantizzazione delle FFR .....	15
2.2.1 Response consistency (RC) .....	15
2.2.2 Phase consistency (PC).....	16
2.2.3 Cross-correlazione .....	18
2.2.3.1 Cross-correlazione stimolo-risposta .....	19
2.2.3.2 Cross-correlazione risposta-risposta.....	19
2.2.4 Pitch Tracking .....	20

2.2.4.1	Frequency Error (FE) .....	21
2.2.4.2	Slope Error (SE).....	22
2.2.4.3	Pitch Strength (PS).....	23
2.3	Riassunto su stato dell'arte e problemi aperti .....	23
2.4	Scopo della tesi e presentazione.....	24
<b>3.</b>	<b>Metodi di machine learning per screening e valutazione otoneurologica.....</b>	<b>26</b>
3.1	Regressione Logistica (LR) .....	26
3.1.1	Modello .....	26
3.1.2	Vantaggi.....	28
3.1.3	Limiti.....	28
3.2	Decision Tree (DT) .....	29
3.2.1	Costruzione dell'albero decisionale .....	30
3.2.2	Vantaggi.....	31
3.2.3	Limiti.....	31
3.3	Random Forest (RF).....	31
3.3.1	Costruzione e predizione della RF .....	32
3.3.2	Vantaggi.....	33
3.3.3	Limiti.....	33
<b>3.4</b>	<b>Principal Component Analysis (PCA).....</b>	<b>34</b>
3.4.1	Formulazione matematica .....	34
3.4.2	Vantaggi.....	36
3.4.3	Limiti.....	36
<b>4.</b>	<b>Materiali e Metodi.....</b>	<b>38</b>
4.1	Dataset reale 1: soggetti normoudenti e patologici.....	38
4.1.1	Condizioni sperimentali .....	38
4.1.2	Stimolazione.....	39
4.1.3	Strumentazione per l'acquisizione .....	39
4.1.4	Pre-processing.....	41

4.2	Dataset reale 2: soggetti portatori di impianto cocleare (IC).....	42
4.2.1	Condizioni sperimentali utilizzate .....	42
4.2.2	Stimolazione .....	42
4.2.3	Registrazione .....	43
4.3	Addestramento ed analisi dei modelli .....	43
4.3.1	Estrazione delle caratteristiche e assegnazione classe.....	43
4.3.2	Addestramento e valutazione metodi di classificazione.....	44
4.3.2.1	Addestramento.....	44
4.3.2.2	Metriche di valutazione .....	45
4.3.3	Formazione e valutazione indici di capacità uditiva .....	47
4.3.3.1	Costruzione indici.....	47
4.3.3.2	Metriche di valutazione .....	48
<b>5.</b>	<b>Risultati sulla rilevazione della presenza/assenza delle FFR.....</b>	<b>49</b>
5.1	Dataset normoudenti e patologici .....	49
5.2	Dataset soggetti portatori IC.....	55
<b>6.</b>	<b>Risultati sugli indici di capacità uditiva .....</b>	<b>61</b>
6.1	Dataset normoudenti e patologici .....	61
6.2	Dataset soggetti portatori di impianto cocleare .....	67
<b>7.</b>	<b>Conclusioni e sviluppi futuri.....</b>	<b>70</b>
7.1	Bilancio del lavoro svolto.....	70
7.2	Margini di miglioramento per sviluppi futuri.....	71
	<b>Appendice A .....</b>	<b>73</b>
	<b>Appendice B .....</b>	<b>84</b>
	<b>Bibliografia.....</b>	<b>104</b>



# Abstract

Nell'ambito audiologico, le Frequency Following Responses (FFRs) sono risposte cerebrali sottocorticali in grado di emulare le caratteristiche spettrali e temporali del segnale di stimolazione uditiva. La loro registrazione riveste un ruolo fondamentale nell'individuazione di patologie neurologiche e di disturbi della funzione uditiva. Tuttavia, il basso rapporto segnale/rumore (SNR) richiede generalmente un elevato numero di registrazioni (sweep), limitandone l'applicabilità in contesti clinici e in popolazioni difficili da valutare.

La presente tesi si propone di applicare metodi di machine learning per identificare la presenza delle FFR e valutare la capacità uditiva dei soggetti sottoposti all'analisi audiometrica. In particolare, l'obiettivo principale è quello di ottenere una valutazione affidabile utilizzando un numero ridotto di sweep, al fine di diminuire il tempo di acquisizione e rendere il protocollo più efficiente e clinicamente applicabile.

A tal fine, i segnali FFR sono stati sottoposti a tecniche avanzate di post-processing, tra cui Conventional Averaging (CA), A Posteriori Wiener Filter (APWF) e SubBand filtering (SB), con l'obiettivo di migliorare la qualità del segnale e incrementare il SNR. Successivamente, sono stati estratti specifici indicatori diagnostici relativi alla fedeltà, alla coerenza e alla codifica del segnale, utilizzati come input per modelli di classificazione, basati su Logistic Regression, Decision Tree o Random Forest, finalizzati a definire la presenza/assenza delle FFR. Inoltre, mediante l'applicazione della Principal Component Analysis (PCA), sono stati sviluppati indici sintetici di capacità uditiva, utili per una valutazione multidimensionale delle prestazioni e per il monitoraggio nei percorsi riabilitativi.

Le metodologie sviluppate sono state applicate a dataset comprendenti soggetti normoudenti (48 tracciati) e patologici (1 tracciato), nonché individui portatori di impianto cocleare (14 tracciati). I risultati ottenuti dimostrano come sia possibile raggiungere prestazioni di classificazione soddisfacenti anche in presenza di un numero ridotto di registrazioni (dell'ordine del 10% - 25 % del totale), evidenziando il potenziale dell'approccio proposto per applicazioni di screening rapido in contesti clinici di primo livello, come farmacie, ambulatori di medicina generale e negozio di audioprotesi. Nel complesso, l'integrazione tra tecniche di elaborazione del segnale e modelli di machine learning rappresenta una strategia promettente per una valutazione oggettiva, efficiente e accessibile della funzione uditiva.



# Elenco degli acronimi utilizzati

ABR	Auditory Brainstem Response
ALR	Auditory Late Response
APWF	A Posteriori Wiener Filter
CA	Conventional Averaging
dB	Decibel
dB SPL	Decibel Sound Pressure Level
DBS	Deep Brain Stimulation
DT	Decision Tree
EEG	Elettroencefalogramma
ENV	Envelope
EMR	Empirical Risk Minimization
FE	Frequency Error
5-FOLD CV	5-Fold Cross-Validation
FFR	Frequency Following Response
fMRI	Risonanza magnetica funzionale
FN	False Negative
F0	Frequenza Fondamentale
FP	False Positive
GI	Global Index
GITPL	Global Inter-Trial Phase Locking
IC	Impianto Cocleare
ISI	Inter-Stimulus Interval
LR	Logistic Regression
MEG	Magnetoencefalografia
MMN	Mismatch Response
PC	Phase Consistency
PC1	Prima Componente Principale
PC2	Seconda Componente Principale
PCA	Principal Component Analysis
PS	Pitch Strength
RC	Response Consistency
RF	Random Forest

SB	SubBand Filter
SE	Slope Error
SNR	Signal to Noise Ratio
SVD	Singular Value Decomposition
TP	True Positive
TN	True Negative
TFS	Temporal Fine Structure





# Capitolo 1

## Frequency Following Response: aspetti teorici, metodologici ed applicativi

### 1.1 Il sistema uditivo

Il sistema uditivo umano è un apparato sensoriale complesso specializzato nel mantenimento dell'equilibrio e nella ricezione, trasmissione e interpretazione delle onde sonore. Dal punto di vista anatomico, tale sistema si può dividere in due sezioni principali: una sezione periferica, responsabile della captazione e trasmissione del suono e una sezione centrale, addetta all'elaborazione ed interpretazione delle informazioni raccolte.

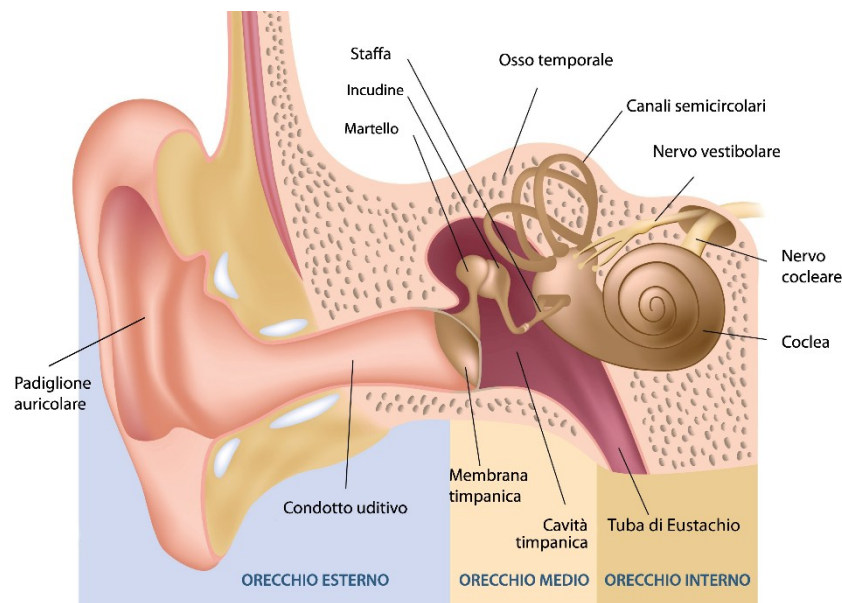


Figura 1: Anatomia sezione periferica del sistema uditivo tratto da [1]

Come illustrato in Figura 1, la sezione periferica può essere ulteriormente divisa in tre segmenti principali noti come orecchio esterno, orecchio medio e orecchio interno.

L'orecchio esterno comprende il padiglione auricolare, fondamentale per la captazione e la localizzazione del suono, e il condotto uditivo, responsabile della trasmissione delle onde sonore alla membrana timpanica.

La regione compresa tra la membrana timpanica e il sistema cocleare rappresenta l'orecchio medio. In quest' area sono presenti tre ossicini denominati martello, incudine e staffa. Queste strutture ossee sono fondamentali per propagare le vibrazioni, generate dalle onde sonore nella membrana timpanica, fino all'orecchio interno.

Nel complesso, queste due sezioni periferiche, oltre a favorire la propagazione del suono verso l'interno, fungono anche da amplificatori. Infatti, il suono arriva all'orecchio interno amplificato all'incirca di 30 dB.[2]

Al termine della sezione periferica è situato l'orecchio interno, costituito da due sistemi distinti: il sistema vestibolare e il sistema cocleare.

Nel sistema cocleare sono posti la coclea, l'organo dei Corti e il nervo cocleare. La coclea è una struttura ossea a forma spirale che opera sia come analizzatore di frequenze e che da trasduttore sensoriale. L'organo dei Corti invece converte l'energia del suono in segnale neurale mediante le cellule ciliate e, tali segnali vengono successivamente trasmessi alla corteccia cerebrale mediante il nervo cocleare.[2]

Al raggiungimento della corteccia cerebrale, l'informazione sonora, dopo aver attraversato il talamo, viene distribuita in diverse aree corticali. In particolare, grazie all'organizzazione gerarchica della corteccia uditiva, le caratteristiche sonore vengono elaborate separatamente dalla corteccia uditiva primaria (A1) e dalle aree associative. La corteccia uditiva primaria è responsabile dell'analisi delle caratteristiche di base, come frequenza e intensità, mentre le aree associative elaborano proprietà più complesse, quali pattern fonetici, ritmo, melodia e altri stimoli multisensoriali. [3]

Oltre alla presenza di un'organizzazione gerarchica, è importante sottolineare che l'informazione uditiva attraversa la corteccia tramite percorsi paralleli e interoperanti. Questi percorsi consentono sia la localizzazione spaziale del suono sia il suo riconoscimento, determinando un'elaborazione dinamica e bidirezionale. Le informazioni percettive, in questo modo, vengono integrate con esperienze pregresse, contesto e processi cognitivi di alto livello. Grazie a queste caratteristiche, la corteccia cerebrale è in grado di effettuare un'elaborazione multisensoriale sofisticata, combinando informazioni di base con dati spaziali, temporali e linguistici, garantendo così una percezione uditiva accurata e funzionale, essenziale sia per la comprensione del linguaggio sia per l'interazione con l'ambiente [3].

## **1.2 Potenziali evocati uditivi e Frequency Following Response**

### **(FFR)**

Il linguaggio rappresenta un segnale acustico complesso che impone numerose richieste al sistema uditivo. Affinché un sistema uditivo possa essere definito sano, esso deve essere altamente sensibile a segnali caratterizzati da rapide variazioni spettrali, da un basso rapporto segnale/rumore (SNR) e da elevate velocità di stimolazione.

Tali richieste vengono soddisfatte attraverso l'attivazione sincrona di pattern biologici distribuiti su ampie popolazioni neurali, che si estendono dal tronco encefalico alla corteccia cerebrale.

Lo studio di questa attività neuronale sincronizzata può essere condotto mediante diversi potenziali evocati uditivi. Questi potenziali riflettono l'attività neurale a differenti livelli lungo la via uditiva e rappresentano uno strumento fondamentale per valutare eventuali alterazioni della tempistica biologica, che possono essere alla base di deficit nella codifica e nella percezione del linguaggio.[4] Tra i principali potenziali evocati uditivi si annoverano Auditory Brainstem Response (ABR), i picchi P1/N1/P2/N2, la Mismatch Response (MMN) e la Frequency Following Response (FFR). Tra questi, la FFR riveste un ruolo di notevole importanza per l'identificazione di rappresentazioni tipiche ed atipiche dei segnali uditivi e, per questo motivo, verrà analizzata nei paragrafi successivi.

### **1.2.1 Origine neurofisiologica delle FFR**

La FFR rappresenta una risposta neurale elettrofisiologica a segnali uditivi complessi. Il termine FFR deriva dalla sua capacità di riflettere fedelmente le proprietà acustiche dello stimolo che la induce, includendone sia le componenti periodiche che aperiodiche. Questa caratteristica la rende uno strumento fondamentale per la comprensione dei principali meccanismi di elaborazione uditiva.

Tradizionalmente, si riteneva che l'origine dell'FFR fosse prevalentemente localizzata a livello del tronco encefalico e del mesencefalo, in particolare nel collicolo inferiore. Tale ipotesi è stata supportata da studi in animali e in pazienti con lesioni focali, dove la compromissione del collicolo inferiore riduceva o aboliva il segnale FFR registrato dallo scalpo. [5][6]

Tuttavia, studi più recenti hanno dimostrato che le FFR registrate dallo scalpo possono avere origini multiple e presentare caratteristiche diverse a seconda dello stimolo acustico utilizzato,

delle modalità di registrazione della risposta (EEG, MEG, fMRI) e delle caratteristiche individuali dei soggetti. [7][8]

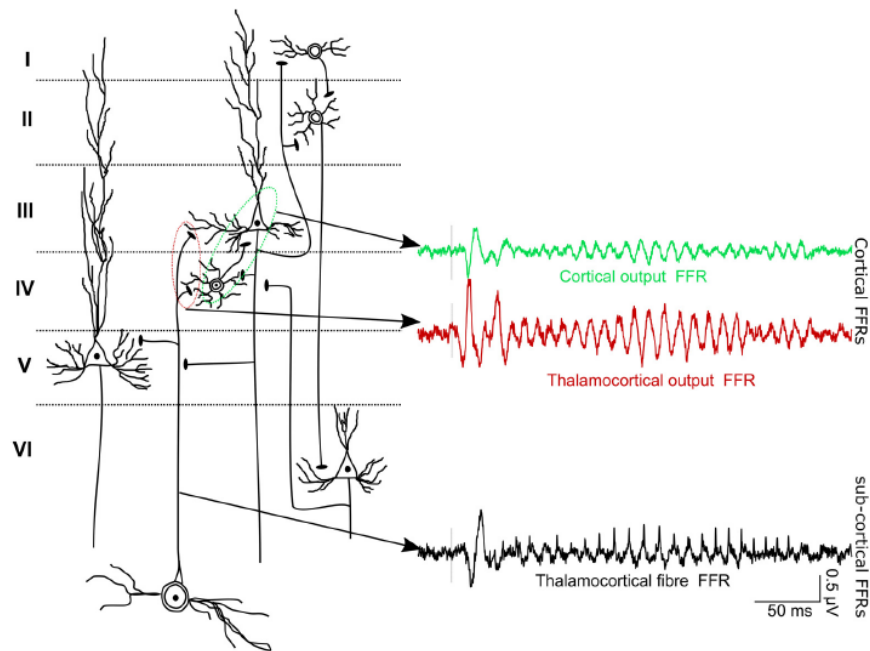


Figura 2: rappresentazione FFR corticali e subcorticali [7]

Dal punto di vista fisiologico, la FFR riflette la sincronizzazione neurale (phase-locking) a componenti periodiche e transitorie del suono. Come illustrato in Figura 2, salendo lungo la via uditiva, la capacità di phase-locking diminuisce, ma le popolazioni neuronali corticali possono comunque codificare frequenze fino a 250–500 Hz in specie animali, indicando che la codifica di frequenze complesse è distribuita su più livelli.[8]

Questa prospettiva multi-sorgente implica che la FFR non debba essere interpretata come un semplice riflesso del collicolo inferiore, ma come un indice funzionale dell'interazione tra reti corticali e sottocorticali, influenzato da plasticità, attenzione e contesto cognitivo. Di conseguenza, studi che combinano tecniche invasive e non invasive, e approcci multimodali, sono fondamentali per isolare i contributi delle singole strutture e comprendere appieno la codifica neurale dei suoni complessi.[7][8]

### 1.2.2 Morfologia delle FFR

La morfologia delle FFR, come già anticipato nella sezione precedente, riflette direttamente e in modo dettagliato le caratteristiche acustiche dello stimolo che le genera. Di conseguenza, la forma del segnale varia considerevolmente in funzione del tipo di stimolo presentato.

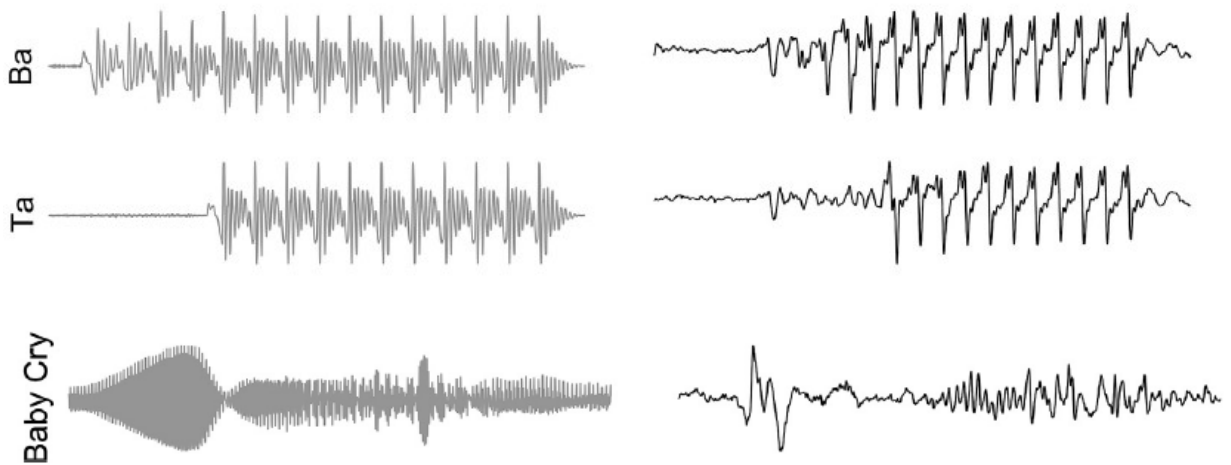


Figura 3: Confronto tra stimoli acustici (sinistra) e relative FFR registrate (destra) modificata da [9]

La Figura 3 mette a confronto tre tipi di stimoli e relative FFR ottenute. In particolare, si può osservare come le FFR ottenute da stimoli sillabici possono essere divise in due componenti distinte. La prima componente rappresenta una fase transiente, associata alle rapide variazioni acustiche tipiche degli onset consonantici, mentre la seconda componente descrive la fase sostenuta del segnale, che riflette la periodicità dello stimolo, come nel caso delle vocali.

Nel caso degli stimoli riportati in figura, le porzioni transitorie associate alle consonanti /b/ e /t/ mostrano marcate differenze, in linea con le diverse caratteristiche acustiche dei fonemi. Al contrario, durante la porzione vocalica /a/, le risposte risultano maggiormente sovrapponibili, evidenziando una codifica più stabile e periodica del segnale.

Per il pianto del bambino (Baby Cry), invece, la morfologia della FFR appare più variabile e meno regolare, riflettendo la natura dinamica e irregolare dello stimolo acustico.

Oltre alle caratteristiche dello stimolo, la morfologia della FFR può variare anche in funzione della polarità della stimolazione utilizzata, evidenziando ulteriormente la complessità dei meccanismi di codifica neurale sottostanti.[9]

In particolare, a livello cocleare, il processo di rettificazione a semionda determina una risposta differenziale alle diverse fasi dello stimolo acustico, influenzando il timing della risposta neurale. Per questo motivo, le FFR vengono frequentemente registrate utilizzando stimoli a polarità alternata (rarefazione e condensazione), consentendo di analizzare separatamente diverse componenti del segnale. In Figura 4, viene riportato un esempio delle diverse componenti estraibili considerando le diverse polarità.

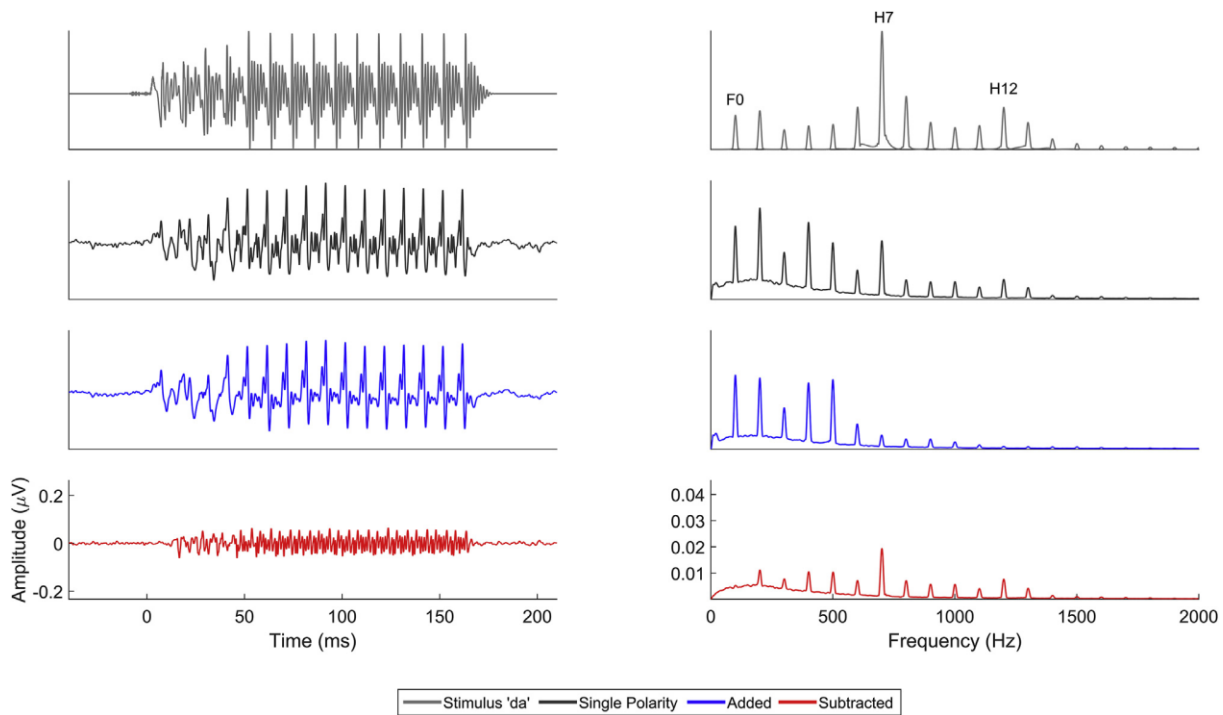


Figura 4: Confronto tra lo stimolo /da/ (grigio), la FFR in singola polarità (nero), la FFR ottenuta dalla somma delle polarità (blu) e la FFR ottenuta dalla sottrazione delle polarità (rosso), nel dominio del tempo (sinistra) e nel dominio delle frequenze (destra).[9]

La Figura 4 mostra come la somma e la sottrazione delle risposte ottenute dalle due polarità permette infatti di enfatizzare rispettivamente l'involuppo temporale (temporal envelope, ENV), rappresentato dal segnale blu, e la struttura fine del segnale (temporal fine structure, TFS), rappresentato dal segnale rosso. Il loro studio permette di ottenere informazioni complementari sui meccanismi di codifica uditiva.

### 1.2.3 Proprietà funzionali

Sebbene la morfologia delle FFR risulti unica, da ciascuna di esse è possibile estrarre le medesime componenti standard. Ognuna di queste componenti contribuisce a caratterizzare l'elaborazione cerebrale dello stimolo presentato. Le principali sono:

- **Il timing dei picchi:** i picchi in un segnale di stimolazione si verificano in momenti specifici. Tali picchi compaiono anche nella risposta a meno di un ritardo fisiologico, dovuto dal tipo necessario al suono per arrivare dall'orecchio al mesencefalo.[9] La loro rilevazione risulta quindi fondamentale per definire la precisione temporale del sistema uditivo e la corretta trasmissione dei segnali lungo le vie neuronali.
- **Frequenza fondamentale (F0):** rappresenta la più bassa frequenza di una forma d'onda periodica, corrispondente alla periodicità del suono. Può essere rilevata sia nella singola

polarità sia nella somma delle polarità ed è fondamentale perché consente al cervello di identificare la tonalità del suono e di distinguere tra diverse voci o strumenti musicali.

- **Armoniche:** multipli interi della  $F_0$ , rilevabili nella risposta fino a circa 1,2–1,3 kHz. Le armoniche contribuiscono alla qualità timbrica del suono e permettono al sistema uditivo di percepire la ricchezza spettrale di stimoli complessi come la voce umana o la musica.
- **Armoniche formanti:** armoniche di ampiezza significativamente maggiore e indipendenti dalla  $F_0$ , con un ruolo cruciale nella percezione fonetica. Queste componenti sono importanti perché forniscono informazioni essenziali per il riconoscimento dei fonemi e la discriminazione dei suoni del linguaggio.

In sintesi, l'analisi di queste componenti permette di comprendere come il sistema uditivo codifichi sia gli aspetti temporali sia quelli spettrali dello stimolo, offrendo informazioni preziose sul funzionamento cerebrale nella percezione dei suoni complessi.[9]

### **1.3 Protocollo standard di acquisizione di FFR**

Il protocollo di acquisizione delle FFR segue procedure standardizzate analoghe a quelle utilizzate per gli ABR evocati da click, con specifiche modifiche legate alla natura dello stimolo.

#### **1.3.1 Stimolazione**

Lo stimolo acustico è presentato a intensità sopraliminari calibrate prima di ogni sessione e generalmente comprese tra 60 e 85 dB SPL. La modalità di presentazione può essere monoaurale o binaurale, in funzione delle caratteristiche della popolazione oggetto di studio.

La durata di ciascun stimolo è compresa tra 40 e 100 ms, nei protocolli che richiedono la minimizzazione del tempo di registrazione, e tra 100 e 500 ms nei casi in cui si intenda preservare la naturalezza dello stimolo acustico.[10]

Per migliorare il rapporto segnale/rumore, gli stimoli vengono ripetuti un elevato numero di volte, tipicamente tra 2000 e 6000 sweeps, e le risposte vengono successivamente mediate mediante tecniche di averaging.

L'intervallo interstimolo (ISI) viene definito in modo da consentire il completo ritorno al baseline della risposta neurale, evitando fenomeni di sovrapposizione tra risposte successive

riducendo la variabilità temporale (jitter). Inoltre, possono essere utilizzate diverse polarità dello stimolo (condensazione e rarefazione), successivamente combinate per ridurre artefatti e isolare specifiche componenti della risposta [9][10].

Tra gli stimoli più utilizzati vi è il fonema /da/, la cui scelta è stata ampiamente motivata in letteratura. Questo fonema, molto diffuso nelle lingue europee, consente di ottenere risposte neurali chiare e altamente replicabili. Inoltre, presenta sia componenti transienti sia sostenute, risultando particolarmente adatto allo studio della codifica subcorticale del parlato. [10][11][12]

### 1.3.2 Registrazione

La registrazione viene eseguita mediante elettrodi di superficie, utilizzando una configurazione standard a tre canali: l'elettrodo attivo è collocato al vertice (Cz), l'elettrodo di riferimento sul lobo auricolare ipsilaterale (A1) e l'elettrodo di massa (ground) a livello frontale o controlaterale (Fpz). Questa configurazione consente di privilegiare la rilevazione dell'attività periferica.

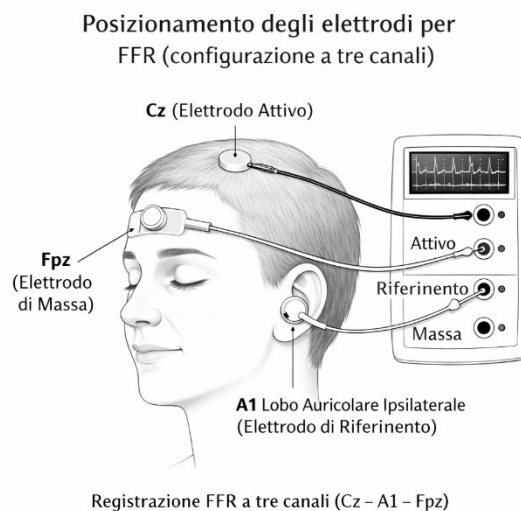


Figura 5: Posizionamento degli elettrodi per FFR (configurazione a tre canali) [33]

Per ridurre possibili interferenze legate all'esecuzione di compiti attivi e garantire condizioni di registrazione passive, ai soggetti viene richiesto di rimanere seduti in posizione rilassata durante l'esame, mentre guardano un film o leggono un libro.[10]

### 1.3.3 Pre-processing

I segnali acquisiti vengono filtrati tipicamente in una banda compresa tra 100 e 3000 Hz, in modo da isolare l'attività subcorticale e ridurre il contributo dei potenziali corticali e del rumore. Successivamente, il segnale viene campionato ad alta frequenza (fino a diverse migliaia di Hz) per garantire un'elevata risoluzione temporale, necessaria per l'analisi delle rapide

componenti della risposta.[10] Infine, le risposte vengono mediate e analizzate sia nel dominio temporale sia in quello spettrale.

Particolare attenzione viene dedicata alla riduzione degli artefatti, tra cui rumore elettrico, attività muscolare e microfonico cocleare, mediante tecniche di filtraggio, esclusione dei trial contaminati e utilizzo di ambienti schermati.[10]

## **1.4 Applicazione clinica e diagnostica delle FFR**

Le FFR rappresentano uno strumento di grande rilevanza in ambito clinico e diagnostico, grazie alla loro elevata versatilità e al crescente impiego nella valutazione dei processi di codifica neurale del suono.

Il loro utilizzo risulta differente rispetto ai click-ABR. Le FFR, infatti, consentono di indagare la precisione e la fedeltà dell'elaborazione uditiva di segnali complessi, mentre i click-ABR vengono principalmente impiegati per identificare condizioni quali perdita dell'udito o neuropatie uditive. Tuttavia, i click-ABR risultano spesso insufficienti nel rilevare deficit nell'elaborazione uditiva che influenzano la vita quotidiana, in quanto, in tali circostanze, i segnali generati dai click-ABR possono risultare normali [9][13][14][15].

Le FFR, integrando queste valutazioni, forniscono quindi informazioni più sensibili e dettagliate sul funzionamento del sistema uditivo.

Dal punto di vista clinico, le FFR consentono di comprendere in che modo il sistema uditivo gestisce e integra le informazioni provenienti da segnali acustici complessi con gli altri sistemi sensoriali [16][17]. Inoltre, rappresentano indicatori affidabili della plasticità neurale, mostrando come le esperienze di vita, la capacità di parlare lingue tonali o essere bilingue e l'allenamento musicale possano modulare la precisione della codifica neurale [17][18][19].

Le FFR si rivelano particolarmente utili anche in ambito diagnostico. Ad esempio, Skoe et al. [20] hanno dimostrato che possono essere utilizzate come indice di deprivazione linguistica. Infatti, lo studio da loro condotto ha evidenziato come adolescenti provenienti da contesti socioeconomici svantaggiati mostrano un'attività neurale più "rumorosa" e risposte al parlato più irregolari, indicando un sistema uditivo meno efficiente rispetto ai coetanei

Oltre alla deprivazione linguistica, le FFR sono di notevole rilevanza per identificare deficit uditivi associati a diverse condizioni cliniche, tra cui autismo, traumi cranici e iperbilirubinemia nei neonati. In particolare, negli studi su soggetti con autismo è stata osservata una difficoltà nella percezione dell'intonazione, suggerendo un potenziale meccanismo neurale alla base delle problematiche nella percezione della prosodia e nella comprensione del linguaggio [21].

In sintesi, le FFR si configurano come un indicatore sensibile e versatile dello stato funzionale del sistema uditivo e dei processi di plasticità neurale, offrendo una prospettiva privilegiata per l'analisi degli effetti delle esperienze ambientali, linguistiche e musicali sulla codifica neurale del suono, con rilevanti implicazioni in ambito sia di ricerca sia clinico.

## **1.5 Estrazione delle FFR**

Nel prossimo capitolo verranno illustrati tre diversi metodi utilizzabili per estrarre le FFR dai segnali acquisiti. Saranno inoltre descritte le principali analisi quantitative applicate ai dati, finalizzate a ottenere informazioni dettagliate sul funzionamento e sull'efficienza del sistema uditivo. Il confronto tra i diversi approcci metodologici permetterà di evidenziare i punti di forza e le limitazioni di ciascun metodo, nonché il tipo di informazioni che possono essere ricavate a livello neurale. Verranno quindi evidenziati problemi aperti e margini di miglioramento e quindi il raggiungimento degli obiettivi prefissati.

# Capitolo 2

## Estrazione ed analisi delle FFR

### 2.1 Metodi di letteratura per l'estrazione delle FFR

Visto che l'EEG domina la FFR si ha un basso SNR che, come è tipico delle registrazioni single-trial, rende necessario adottare tecniche di post-processing. Tra le principali metodologie, che saranno descritte nei paragrafi successivi, si annoverano la media convenzionale (CA), il filtro di Wiener a posteriori (APWF) e il SubBand Filter (SB).

#### 2.1.1 Media convenzionale (CA)

La media convenzionale (Conventional Averaging, CA) rappresenta il metodo più diffuso e consolidato per l'estrazione dei potenziali evocati. Questo approccio si basa sull'assunzione che il segnale neurale sia temporalmente allineato rispetto allo stimolo, mentre il rumore viene considerato casuale e non correlato tra le diverse registrazioni.

Sotto queste ipotesi, un segnale affidabile può essere ottenuto tramite la media aritmetica di un notevole numero di registrazioni acquisite. Formalmente, il procedimento può essere espresso con la seguente formula:

$$\bar{x}(t) = \frac{1}{N} \sum_{i=1}^N x_i \quad (1)$$

Dove  $\bar{x}(t)$  rappresenta il segnale medio,  $x_i$  il segnale rilevato all' $i$ -esima registrazione e  $N$  sono il numero di registrazioni totali.

#### 2.1.2 Filtro alla Wiener a Posteriori (APWF)

Il filtro alla Wiener a Posteriori rappresenta un metodo di filtraggio avanzato impiegato per il miglioramento del SNR.

Esso risulta particolarmente efficace nei casi in cui segnale d'interesse e rumore si sovrappongono in modo significativo. In tale condizione i metodi convenzionali possono risultare insufficienti nel garantire un'adeguata ricostruzione del potenziale evocato.

Il principio su cui si basa il metodo consiste nel preservare le componenti in frequenza caratterizzate da un elevato SNR e attenuare quelle in cui tale condizione non è soddisfatta.

Assumendo che il segnale osservato sia dato dalla somma di segnale utile e rumore, scorrelati e stazionari nel tempo. Ogni sweep può essere espressa come segue:

$$x_k(t) = s(t) + n_k(t) \quad (2)$$

Dove  $x_k(t)$  equivale al segnale registrato alla k-esima sweep,  $s(t)$  al potenziale evocato,  $n_k(t)$  al rumore contenuto nella k-esima sweep.

Successivamente, si procede al calcolo di due spettri di potenza distinti:

- Lo spettro medio delle singole sweep  $\bar{P}_x(\omega)$ , ottenuto mediando gli spettri di potenza di ciascuna acquisizione:

$$\bar{P}_x(\omega) = \frac{1}{N} \sum_{k=1}^N P_k(\omega) \quad (3)$$

- Lo spettro di potenza del segnale medio temporale  $P_{\bar{x}}(\omega)$ , che può essere espresso come segue:

$$P_{\bar{x}}(\omega) = P_s(\omega) + \frac{1}{N} P_n(\omega) \quad (4)$$

In cui  $P_s(\omega)$  è pari allo spettro del segnale e  $P_n(\omega)$  è lo spettro del rumore.

Combinando questi due spettri, si ottiene la funzione di trasferimento del filtro di Wiener a posteriori  $H(\omega)$ :

$$H(\omega) = \frac{P_s(\omega)}{P_s(\omega) + \frac{1}{N} P_n(\omega)} = \frac{N}{N-1} \left[ 1 - \frac{1}{k} \frac{\bar{P}_x(\omega)}{P_{\bar{x}}(\omega)} \right] \quad (5)$$

Questa funzione consente di selezionare e pesare le componenti spettrali più informative, attenuando quelle dominate dal rumore.

Infine, il segnale filtrato viene ottenuto, nel dominio della frequenza, attraverso il prodotto tra lo spettro di potenza del segnale medio  $\bar{X}(\omega)$  e la funzione di trasferimento  $H(\omega)$ :

$$\hat{S}(\omega) = \bar{X}(\omega) * H(\omega) \quad (6)$$

Il segnale nel dominio del tempo  $\hat{s}(t)$  si ricava quindi applicando l'anti-trasformata di Fourier.[22]

### 2.1.3 SubBand Filter (SB)

Il SubBand Filter (SB) rappresenta una tecnica avanzata di media ponderata progettata per adattarsi alle complesse caratteristiche spettrali delle FFR. L'obiettivo principale è estrarre le componenti armoniche che costituiscono la risposta uditiva. A tal fine, il segnale viene decomposto in diverse sottobande di frequenza mediante filtri dedicati e successivamente ricombinato, attraverso una procedura di pesatura, per ottenere il segnale ricostruito.

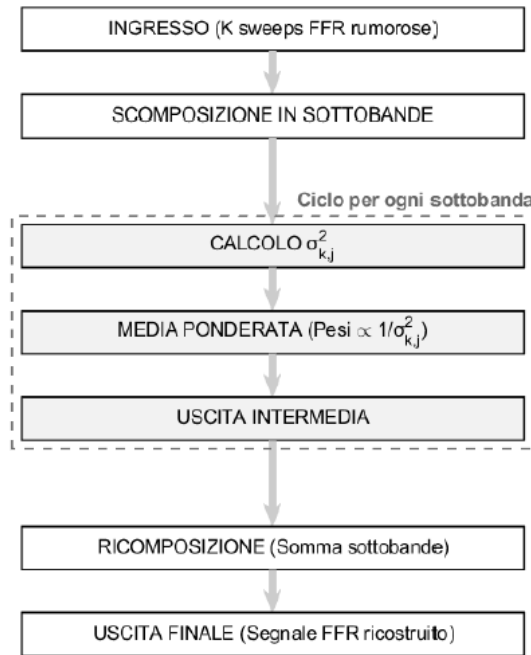


Figura 6: Diagramma di flusso del SubBand filter [25]

Come illustrato nel diagramma in Figura 5, ogni sweep viene suddivisa in diverse sottobande di frequenza. Per ciascuna sottobanda, e per ogni sweep filtrata al suo interno, viene calcolata la varianza del segnale. Successivamente, per ogni sottobanda si esegue una media ponderata tra le sweep filtrate, utilizzando pesi derivati dalla varianza, ottenendo così un unico segnale rappresentativo per ciascuna sottobanda. [23]

Il segnale per ogni sottobanda  $j$  è rappresentabile mediante l'Equazione 7:

$$\bar{y}_j(t) = \frac{\sum_{k=1}^N \left( \frac{1}{\sigma_k^{j/2}} y_k^j(t) \right)}{\sum_{k=1}^N \left( \frac{1}{\sigma_k^{j/2}} \right)} \quad (7)$$

Dove  $\sigma_k^{j^2}$  equivale la varianza della sweep  $k$  nella sottobanda  $j$  e  $y_k^j(t)$  rappresenta il segnale il segnale della sweep  $k$  nella sottobanda  $j$ .

Infine, i segnali ponderati ottenuti per ciascuna sottobanda vengono combinati mediante somma, al fine di ricostruire il segnale finale:

$$\bar{y}(t) = \sum_{j=1}^J \bar{y}_j(t) \quad (8)$$

In questo modo, il SB consente di valorizzare le componenti armoniche più stabili e informative, migliorando la qualità complessiva della stima della FFR, in linea con approcci avanzati di miglioramento del rapporto segnale-rumore proposti in letteratura.

#### **2.1.4 Metodi a confronto**

Confrontando tra loro i metodi precedentemente descritti, emergono differenze significative in diversi contesti applicativi.

La CA, sebbene rappresenti ancora oggi il metodo più utilizzato, presenta numerose limitazioni, soprattutto nell'analisi di popolazioni fragili. In particolare, essa richiede un elevato numero di registrazioni per ottenere una stima affidabile, condizione spesso impraticabile nel caso di neonati o pazienti clinicamente compromessi.[10]

Inoltre, l'incapacità di distinguere efficacemente tra segnale e rumore quando questi si sovrappongono limita ulteriormente le prestazioni del metodo, rendendolo meno efficace in condizioni di basso rapporto segnale-rumore.[10]

Questa limitazione risulta in parte superata dall'utilizzo di APWF, il quale consente di gestire in modo più efficace la sovrapposizione tra segnale e rumore, nonché eventuali disturbi dinamici non adeguatamente trattabili mediante media coerente. Tuttavia, l'ipotesi di stazionarietà e indipendenza tra segnale d'interesse e rumore può risultare restrittiva, rendendo talvolta necessario il ricorso a versioni più avanzate del metodo.[24]

Infine, il SB presenta caratteristiche in parte analoghe all'APWF. In particolare, consente un controllo accurato della qualità del segnale su diverse regioni dello spettro e un'adeguata estrazione delle componenti ad alta frequenza. Nonostante ciò, il metodo risulta computazionalmente più oneroso e la qualità della stima della varianza influisce in modo significativo sull'efficacia della media ponderata.[25]

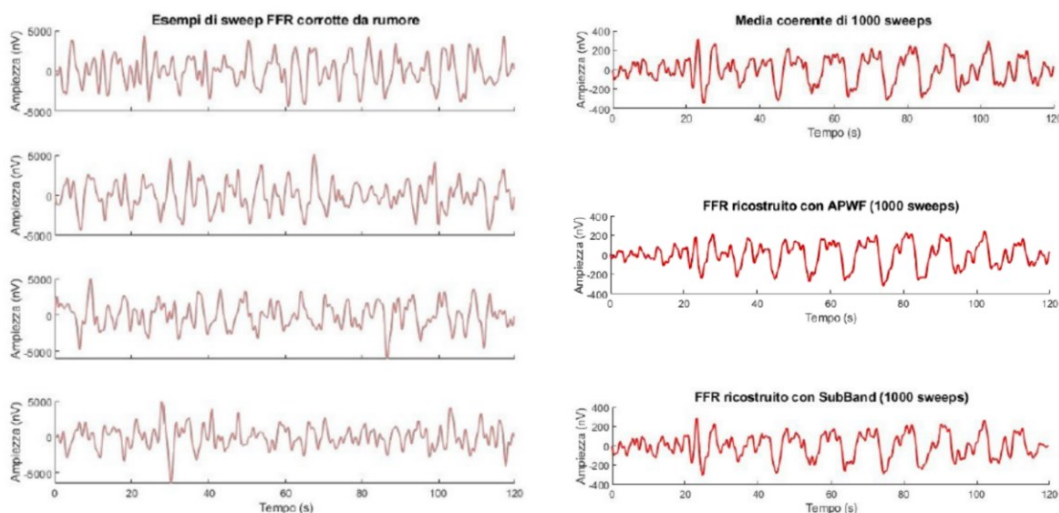


Figura 7: confronto segnali single-trial (sinistra) e FFR ricostruite con CA, APWF e SB (destra).  
Immagine modificata da [24]

In conclusione, mentre la media coerente rappresenta una soluzione semplice e consolidata, l'APWF e il SB offrono prestazioni superiori in condizioni di basso SNR, al costo però di una maggiore complessità computazionale e di ipotesi modellistiche più restrittive. La scelta del metodo più appropriato dipende pertanto dal contesto applicativo e dal compromesso tra accuratezza della stima e fattibilità operativa.

## 2.2 Indicatori proposti in letteratura per la quantizzazione delle FFR

Le FFR sono ricche di informazioni utili alla caratterizzazione di numerose condizioni fisiologiche e patologiche. Per poter estrarre in modo efficace tali informazioni, è necessario definire e calcolare specifici indicatori quantitativi.

### 2.2.1 Response consistency (RC)

La Response Consistency (RC) rappresenta una misura della correlazione tra i segnali acquisiti all'interno della stessa sessione di registrazione. Essa fornisce un indice della stabilità della FFR da prova a prova [9].

Il processo di calcolo prevede la suddivisione dei segnali registrati in due sottoinsiemi. La creazione di tali sottoinsiemi può avvenire secondo diverse modalità:

- *Sweep pari vs sweep dispari*: metodo utilizzato per valutare se un'eventuale ridotta consistenza sia dovuta a una variabilità costante della codifica neurale o a fenomeni di fatica neurale [14];
- *Prima metà vs seconda metà delle sweep*: approccio simile al precedente, utile per evidenziare eventuali variazioni temporali durante la registrazione;
- *Selezione casuale senza ripetizione (bootstrap)*: suddivisione casuale delle sweeps nei due sottoinsiemi, ripetuta tipicamente tra 100 e 300 iterazioni, al fine di ottenere una stima più robusta.

Successivamente, per ciascun sottoinsieme viene calcolata la media punto per punto, ottenendo due segnali FFR medi distinti. Tra questi viene quindi calcolata la cross-correlazione.

Dalla funzione di cross-correlazione si individua il valore massimo, indicato come  $r_{\max}$ , che viene direttamente trasformato in z-score nel caso in cui non si utilizzi la tecnica di bootstrap. Nel caso contrario, si procede dapprima al calcolo del valore massimo per ciascuna iterazione del bootstrap; successivamente, tali valori vengono mediati e il risultato ottenuto viene infine trasformato in z-score.

Valori di  $r_{\max}$  prossimi a 1 indicano un'elevata coerenza tra le FFR medie ottenute e, di conseguenza, una maggiore stabilità della risposta.

Dal punto di vista interpretativo, valori elevati della RC sono indicativi di una codifica neurale stabile e affidabile dello stimolo acustico. Al contrario, valori ridotti possono riflettere una maggiore variabilità della risposta neurale, potenzialmente associata a condizioni di deficit uditivi, difficoltà di elaborazione temporale o ridotta sincronizzazione neurale.[14]

### **2.2.2 Phase consistency (PC)**

Un ulteriore indicatore calcolabile direttamente sui segnali acquisiti, prima di ottenere la FFR complessiva, è la Phase Consistency (PC).

La PC rappresenta una tecnica di analisi del timing della risposta neurale, basata sulla coerenza di fase tra le diverse sweep registrate.

A differenza dei metodi tradizionali, essa consente di valutare la stabilità temporale della FFR utilizzando un numero inferiore di sweep rispetto a quello generalmente necessario per identificare con precisione i picchi della risposta neurale. Inoltre, essendo meno suscettibile al rumore, fornisce una misura più robusta della sincronizzazione neurale anche in condizioni di basso SNR [9][26].

L'analisi della coerenza di fase si basa sull'assunzione che le frequenze dello stimolo presenti nella ENV e/o nella TFS mostrino una fase coerente da prova a prova, mentre le componenti non appartenenti allo stimolo non presentino tale coerenza.

Sulla base di queste ipotesi, si procede al calcolo della Fast Fourier Transform (FFT) per ciascuna sweep registrata, considerando specifiche finestre temporali (ad esempio le componenti transitoria e sostenuta di una sillaba) [27].

Dalla FFT si ottiene, per ogni prova, un vettore complesso che descrive ampiezza e fase di ciascuna componente in frequenza. Tali vettori vengono quindi normalizzati a modulo unitario, eliminando l'informazione sull'ampiezza e preservando esclusivamente l'informazione di fase. Infine, i vettori così ottenuti vengono mediati tra loro, permettendo di stimare la coerenza di fase per ciascuna frequenza secondo la seguente espressione:

$$R(f) = \left| \frac{1}{N} \sum_{k=1}^N e^{i\theta_k(f)} \right| \quad (9)$$

Dove  $N$  rappresenta il numero di trial,  $e^{i\theta_k(f)}$  è il vettore unitario del  $k$ -esimo trial e  $R(f)$  il phase-locking della frequenza  $f$ , detto anche Inter-trial Phase Locking (ITPL).

Valori di  $R(f)$  prossimi a 1 indicano un'elevata sincronizzazione di fase tra le prove, mentre valori prossimi a 0 riflettono una distribuzione casuale delle fasi.

Questo tipo di indicatore può essere ricavato anche attraverso l'utilizzo di tecniche a finestra scorrevole o, come illustrato in Figura 7, utilizzando diverse polarità.

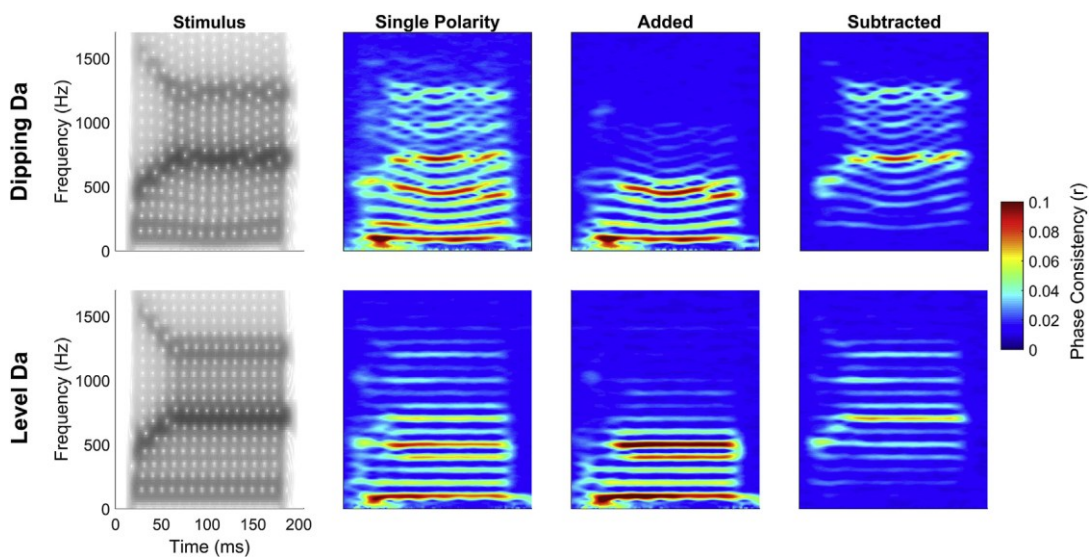


Figura 8: PC di un "dipping da" e di un "level-F0 da" utilizzando risposte a polarità singola, polarità sommate e polarità sottratte. [9]

L'esempio riportato nella figura precedente illustra come le diverse modalità di analisi delle polarità influenzino la coerenza di fase delle FFR. In particolare, per le risposte a polarità singola, la F0, la seconda armonica e le frequenze delle formanti mostrano la coerenza di fase più elevata, indicando una sincronizzazione ottimale delle componenti principali dello stimolo. La somma delle polarità enfatizza la coerenza delle frequenze più basse, tipicamente associate all' ENV, mentre la sottrazione delle polarità mette in evidenza la coerenza delle frequenze più alte, legate alla TFS. Questo esempio mostra chiaramente come la manipolazione delle polarità possa evidenziare selettivamente componenti diverse della risposta neurale.

Per ottenere una misura globale della coerenza di fase si utilizza il Global Inter-Trial Phase Locking (GITPL).

Tale indice viene calcolato considerando la F0 dello stimolo e le successive K armoniche. Per ciascuna di queste frequenze si analizza un intervallo di  $\pm 40$  Hz e si individua il valore massimo di coerenza di fase. I valori massimi ottenuti per tutte le frequenze considerate vengono mediati, fornendo così un indice complessivo della sincronizzazione di fase della risposta neurale rispetto allo stimolo. In termini matematici, il GITPL può essere espresso come segue:

$$GITPL = \frac{1}{K} \sum_{l=1}^K R(f_l) \quad (10)$$

Dal punto di vista diagnostico, valori elevati di GITPL indicano una forte sincronizzazione neurale alle componenti armoniche dello stimolo, riflettendo un'efficace codifica temporale da parte del tronco encefalico e delle vie uditive centrali. Al contrario, valori ridotti di GITPL riflettono una maggiore variabilità della fase tra prove, suggerendo che il sistema uditivo non riesce a seguire in modo coerente lo stimolo acustico. Tale ridotta coerenza può essere associata a difficoltà nella comprensione del parlato, ridotta plasticità neurale e maggiore sensibilità a rumori ambientali o fisiologici.[25]

### 2.2.3 Cross-correlazione

La RC e la PC rappresentano misure della fedeltà della risposta all'interno della stessa sessione di registrazione. Per valutare invece la fedeltà tra sessioni diverse, oppure tra la risposta neurale e lo stimolo che l'ha evocata, è possibile ricorrere all'analisi di cross-correlazione.

In questo contesto, è possibile distinguere due tipologie principali di misura: la cross-correlazione stimolo-risposta e la cross-correlazione risposta-risposta.

### 2.2.3.1 Cross-correlazione stimolo-risposta

La cross-correlazione tra stimolo e risposta consente di valutare la somiglianza morfologica tra la risposta neurale e lo stimolo che l'ha generata [9].

Il calcolo di tale metrica prevede, come primo passaggio, il filtraggio dello stimolo. Questa operazione è necessaria per ottenere una forma d'onda dello stimolo maggiormente confrontabile con la FFR registrata.

Successivamente, si procede al calcolo della cross-correlazione tra lo stimolo filtrato e la risposta evocata, all'interno di una specifica regione di interesse, secondo la seguente formulazione:

$$R_{xy}(\tau) = \sum_{m=0}^{\infty} x[m]y[m + \tau] \quad (11)$$

dove  $R_{xy}(\tau)$  rappresenta la somiglianza tra i due segnali in funzione dello spostamento temporale  $\tau$ , detto anche *lag* o ritardo.

È importante sottolineare che la traslazione della risposta evocata viene considerata solo per valori positivi, poiché la risposta neurale non può precedere lo stimolo. Una volta calcolata la funzione di cross-correlazione per tutte le traslazioni previste, si procede all'individuazione del valore massimo di correlazione e del corrispondente lag. Tali valori forniscono rispettivamente una misura della somiglianza tra i segnali e del ritardo di trasmissione neurale.

Dal punto di vista interpretativo, valori elevati di correlazione indicano una maggiore fedeltà della risposta neurale allo stimolo, suggerendo una codifica temporale accurata. Al contrario, valori ridotti possono riflettere una minore precisione nella rappresentazione dello stimolo, potenzialmente associata a difficoltà di elaborazione uditiva.

Il valore del lag fornisce invece informazioni sulla latenza della risposta neurale: latenze superiori al ritardo fisiologico, generalmente compreso tra 7 e 10 ms, possono indicare un rallentamento nella trasmissione o nell'elaborazione del segnale lungo le vie uditive.[9]

### 2.2.3.2 Cross-correlazione risposta-risposta

La cross-correlazione risposta-risposta consente di confrontare due risposte neurali, generalmente generate dallo stesso stimolo ma in condizioni diverse, ad esempio in presenza o assenza di rumore di fondo.

Il calcolo di tale metrica è analogo a quello della cross-correlazione stimolo-risposta, con la differenza che, in questo caso, le traslazioni temporali possono assumere anche valori negativi.

Ciò è possibile poiché entrambe le risposte rappresentano segnali neurali e non vi è un vincolo causale diretto come nel caso stimolo-risposta.

Nel caso di lag negativo, la risposta traslata risulta in anticipo rispetto a quella di riferimento, mentre per valori positivi si osserva un ritardo.

L'analisi del valore massimo di correlazione e del corrispondente lag consente quindi di valutare sia il grado di somiglianza tra le due risposte sia eventuali differenze temporali tra le condizioni considerate.[9]

## 2.2.4 Pitch Tracking

Il pitch tracking rappresenta una misura della capacità del sistema uditivo di seguire le variazioni temporali dell'altezza tonale (pitch). Il suo scopo principale è quello di valutare la precisione con cui il cervello decodifica la F0 dello stimolo acustico [28].

Il metodo più diffuso per ottenere tale misura consiste nel calcolo dell'autocorrelogramma a finestra scorrevole.

La procedura di calcolo è analoga a quella della cross-correlazione, con la differenza che in questo caso il segnale viene confrontato con una versione ritardata di sé stesso. In questo modo è possibile stimare, istante per istante, la periodicità del segnale e quindi ricavare l'andamento temporale della F0.

Il procedimento iterativo utilizzato può essere riassunto come segue.

Per ogni tempo  $t$ :

1. Si considerano due finestre temporali, solitamente di 40 ms con una sovrapposizione di 39 ms
2. Si calcola l'autocorrelazione massima ( $r_{max}$ ) e il ritardo corrispondente ( $lag_{max}$ ) per tale sovrapposizione, utilizzando la stessa formula impiegata per la cross-correlazione stimolo-risposta.
3. Al tempo  $t$  considerato si assegna il valore  $r_{max}$  e il reciproco di  $lag_{max}$  che corrisponde alla F0 in quel punto temporale.
4. Le finestre vengono traslate lungo il tempo e il procedimento viene ripetuto, ottenendo così un profilo temporale continuo della F0.

La Figura 8 mostra un'applicazione pratica dell'autocorrelazione per l'estrazione della F0 della sillaba /ya/ crescente.

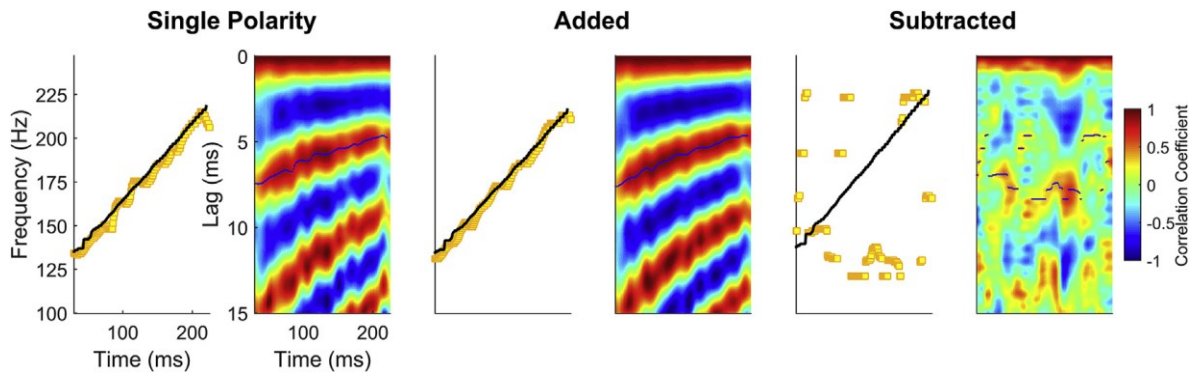


Figura 9: autocorrelogrammi formati per estrarre F0 della sillaba /ya/ crescente. A sinistra, il pitch estratto nel tempo (quadrati gialli) confrontato con il pitch dello stimolo (punti neri). A destra, l'autocorrelogramma della FFR, in cui il lag per ottenere la correlazione massima (rosso scuro) corrisponde alla periodicità della frequenza fondamentale. [9]

Nella figura sovrastante, si può notare come la periodicità è chiaramente evidente nelle risposte a polarità singola (sinistra) o nelle risposte sommate (centro), mentre non è rilevabile nelle risposte sottratte (destra). Questo esempio evidenzia ulteriormente come la scelta della polarità influenzi la capacità di rilevare la periodicità della F0 nelle FFR.[9]

In ambito clinico e diagnostico, il pitch tracking riveste un ruolo di particolare importanza, in quanto consente di valutare la precisione con cui il cervello rappresenta stimoli vocali e musicali e di analizzare come la plasticità cerebrale venga modulata da fattori quali età, apprendimento linguistico e addestramento musicale. [29][30][31]. Inoltre, il pitch tracking costituisce un potente indicatore della resilienza neurale in presenza di rumore di fondo, fornendo informazioni utili sulla capacità del sistema uditivo di mantenere una codifica temporale stabile anche in condizioni sfavorevoli [25].

#### 2.2.4.1 Frequency Error (FE)

Il Frequency Error (FE) è un parametro utilizzato per quantificare la capacità di un soggetto di seguire la frequenza fondamentale dello stimolo.

Per il calcolo di tale indice, si determinano innanzitutto gli spettrogrammi dello stimolo e della risposta, al fine di evidenziare la variazione delle componenti in frequenza nel tempo. Successivamente, l'intervallo temporale viene suddiviso in  $N$  segmenti e, per ciascun intervallo, viene individuata la frequenza corrispondente al massimo di ampiezza nel range della F0. Una volta ottenute le stime della F0 per stimolo e risposta, si calcola, per

ogni intervallo, la differenza tra le due frequenze, ottenendo così un errore di frequenza locale. Formalmente,  $\forall t = 1, \dots, N$  :

$$FE(t) = |F0_{stim}(t) - F0_{risp}(t)| \quad (12)$$

Il valore complessivo del FE si ottiene quindi come media aritmetica degli errori calcolati nei singoli intervalli:

$$FE = \frac{1}{N} \sum_{i=1}^N FE(t) \quad (13)$$

In letteratura è stato osservato che valori ridotti di FE sono associati a una maggiore capacità del sistema uditivo di seguire la F0 dello stimolo. Al contrario, valori elevati indicano una minore precisione nella codifica della F0 e sono stati associati a possibili difficoltà nel neurosviluppo.[21]

#### 2.2.4.2 Slope Error (SE)

Lo Slope Error (SE) è un ulteriore indice ricavabile dal confronto degli andamenti temporali della frequenza fondamentale dello stimolo e della risposta.

Il suo calcolo prevede inizialmente la stima delle traiettorie della F0 di stimolo e risposta nel tempo, attraverso l'utilizzo degli spettrogrammi, analogamente a quanto effettuato per il calcolo del FE. Tuttavia, in questo caso non viene considerata la differenza istante per istante; piuttosto, una volta ottenute le due curve, se ne calcola la pendenza mediante regressione lineare (o, alternativamente, tramite derivata). La differenza assoluta tra le pendenze delle due curve fornisce l'indice SE:

$$SE = |slope_{stim} - slope_{risp}| \quad (14)$$

Questo parametro esprime la capacità del sistema neurale di seguire le variazioni temporali del tono dello stimolo. In particolare, valori ridotti di SE indicano una buona capacità di tracciare la dinamica della F0. Valori elevati riflettono difficoltà nel seguire le variazioni temporali del tono associabili a invecchiamento, disturbi dello sviluppo o assenza di esperienza musicale.[21]

### 2.2.4.3 Pitch Strength (PS)

A completare gli indicatori derivabili dal pitch tracking si trova il Pitch Strength (PS).

Il PS rappresenta una misura della robustezza con cui i neuroni sottocorticali elaborano la frequenza fondamentale di uno stimolo acustico. Tale robustezza si riflette nella coerenza tra la periodicità della FFR e quella dello stimolo [26].

Il parametro viene ottenuto calcolando l'autocorrelazione del segnale su finestre temporali scorrevoli. Successivamente, per ciascun istante  $t$ , il valore di PS viene determinato normalizzando l'autocorrelazione rispetto al valore corrispondente a ritardo nullo:

$$PS(t) = \frac{R_{ww}(L_{F0}) - R_{ww}(0)}{1 - R_{ww}(0)} \quad (15)$$

Dove  $R_{ww}(L_{F0})$  corrisponde al valore di autocorrelazione in  $F0$  e  $R_{ww}(0)$  è il valore di autocorrelazione al lag 0.

Dal punto di vista interpretativo, valori elevati di PS indicano una forte rappresentazione della periodicità della  $F0$ , suggerendo una codifica neurale robusta e stabile dello stimolo. Al contrario, valori ridotti riflettono una minore sincronizzazione neurale e una rappresentazione meno precisa della frequenza fondamentale, potenzialmente associata ad invecchiamento uditivo o disturbi dello sviluppo.[10][32]

## 2.3 Riassunto su stato dell'arte e problemi aperti

La valutazione della percezione uditiva di un soggetto e delle modalità con cui essa si manifesta rappresentano due degli obiettivi principali nell'ambito della diagnosi audiologica differenziale. Tali finalità sono adeguatamente raggiungibili mediante l'analisi delle risposte FFR ottenute a seguito di una stimolazione sonora. Tuttavia, il limitato SNR rende necessario effettuare numerose stimolazioni ed acquisizioni, riducendo così l'applicabilità della metodica in popolazioni difficili da valutare.

Negli ultimi anni, diversi studi hanno affrontato questa problematica attraverso l'introduzione di tecniche di post-processing finalizzate ad ottenere FFR di qualità adeguata anche con un numero significativamente inferiore di registrazioni. In particolare, sono stati applicati filtri avanzati, come APWF e SB, per incrementare efficacemente SNR ed estrarre le componenti rilevanti del segnale, sfruttando le proprietà armoniche della risposta uditiva. Questi approcci permettono di raggiungere risultati soddisfacenti anche in presenza di una quantità di dati ridotta e in condizioni di elevato rumore [22][23] [24][25]. Ulteriori studi hanno

esplorato l'applicazione di tecniche di machine learning per il rilevamento e la classificazione della presenza o assenza di FFR. Tra questi, il metodo Source Separation No Negative Matrix Factorization (SSNMF) si è distinto per la capacità di fornire segnali adeguati all'analisi già dopo 250 registrazioni. [34]

Successivamente al post-processing, gli obiettivi sopra menzionati vengono perseguiti attraverso il calcolo di specifici indicatori diagnostici, come RC, PC, cross-correlazione, FE, SE e PS, che consentono di valutare caratteristiche come fedeltà, coerenza della risposta ed elaborazione del tono. Pur essendo notevolmente informativi, attualmente tali indicatori risultano di difficile interpretazione per operatori non specialisti, come medici generici, farmacisti o tecnici audioprotesisti.

## **2.4 Scopo della tesi e presentazione**

Il presente lavoro si propone di utilizzare alcuni degli indicatori definiti nella Sezione 2.2 per addestrare modelli di machine learning finalizzati alla classificazione della presenza/assenza di FFR e alla valutazione otoneurologica dei soggetti esaminati, impiegando un numero sensibilmente ridotto di registrazioni.

Gli obiettivi principali riguardano lo sviluppo di modelli affidabili, in grado di supportare attività di screening oggettivo in contesti di assistenza sanitaria di primo livello, quali farmacie, ambulatori di medicina generale e negozi di audioprotesi. Inoltre, il lavoro prevede la definizione di indici di capacità uditiva, utili per una valutazione multidimensionale delle prestazioni e per il monitoraggio della loro evoluzione nell'ambito dei percorsi riabilitativi. I modelli sviluppati verranno descritti nel dettaglio nel Capitolo 3.

Nel Capitolo 4, invece, verranno descritti i dataset impiegati e illustrato il processo metodologico adottato per il raggiungimento degli obiettivi prefissati. In particolare, verranno definite le metriche per la valutazione delle performance di classificazione di tre diversi modelli di machine learning: Logistic Regression (LR), Decision Tree (DT) e Random Forest (RF). Inoltre, saranno descritti gli approcci utilizzati per costruire e valutare la robustezza degli indici di capacità uditiva sviluppati.

I risultati ottenuti saranno presentati e discussi nei Capitoli 5 e 6: il Capitolo 5 sarà dedicato all'analisi delle prestazioni dei modelli di classificazione, mentre il Capitolo 6 riguarderà la valutazione degli indici sviluppati.

Tutte le analisi verranno condotte sugli indicatori FFR, ottenuti mediante post-elaborazione del segnale tramite CA, APWF e SB, considerando un numero crescente di sweep. Lo studio si

basa su due distinti dataset: il primo include soggetti normoudenti e soggetti patologici, mentre il secondo comprende individui portatori di impianto cocleare.

# Capitolo 3

## Metodi di machine learning per screening e valutazione otoneurologica

Nel presente capitolo vengono presentati alcuni dei possibili metodi impiegabili per la classificazione della presenza o assenza di FFR e per la definizione di uno o più indici rappresentativi della capacità uditiva. Le metodologie proposte si basano sulle proprietà estratte dal segnale FFR ottenuto a seguito delle procedure di post-processing convenzionali e non convenzionali descritte nel Capitolo 2.

### 3.1 Regressione Logistica (LR)

La Regressione Logistica (Logistic Regression, LR) rappresenta una tecnica di apprendimento supervisionato applicabile sia a problemi di classificazione binaria (presente/assente), sia a problemi di classificazione multiclasse (presente/assente/patologica).

Il modello associa ciascuna variabile indipendente, rappresentata dagli indicatori descritti nella Sezione 2.2, a un coefficiente indicativo della sua influenza sulla probabilità che la FFR venga classificata come presente o assente. Questo approccio permette di interpretare chiaramente il contributo delle singole caratteristiche, rendendo la LR un modello semplice da implementare, interpretabile e particolarmente adatto a contesti di screening in cui è necessario identificare i fattori determinanti della classificazione.

#### 3.1.1 Modello

Nel contesto della classificazione binaria, il modello determina la probabilità che la classe di appartenenza di una data istanza  $X = [x_1, x_2, x_3, \dots, x_d]$  sia la classe 1, anche definita classe positiva.

Questa probabilità è determinata tramite la funzione logistica, detta anche funzione sigmoide, definita come segue:

$$P(Y = 1|X) = \frac{1}{1 + e^{-\beta^T X}} \quad (16)$$

dove  $\beta \in R^d$  rappresenta il vettore dei coefficienti associati alle caratteristiche dell'istanza  $X$ .

Il termine “*sigmoide*” si riferisce alla caratteristica “forma ad S” del grafico della funzione logistica.



Figura 10: rappresentazione funzione logistica. [35]

Tale andamento, come mostrato in Figura 10, consente di trasformare il valore reale  $\beta^T X$  in una probabilità compresa tra 0 e 1.

Una volta definita la funzione di stima delle probabilità di appartenenza alla classe, è necessario stimare i parametri del modello, ovvero i coefficienti beta  $\beta$ . Tale operazione avviene nella fase di addestramento del modello mediante la risoluzione del problema di minimizzazione del rischio empirico (Empirical Risk Minimization, ERM).[36]

Nella LR, la funzione di perdita utilizzata è la logistic loss function:

$$l(\beta, (x_i, y_i)) = \log(1 + e^{-y_i(\beta^T x_i)}) \quad (17)$$

Pertanto, considerando un dataset di addestramento costituito da  $m$  coppie  $(x_i, y_i)$ , il relativo problema di ottimizzazione può essere formalizzato come segue:

$$\operatorname{argmin}_{\beta \in \mathbb{R}^d} \frac{1}{m} \sum_{i=1}^m \log(1 + e^{-y_i(\beta^T x_i)}) \quad (18)$$

Il problema in questione può essere affrontato efficacemente tramite metodi standard, in quanto la funzione di perdita logistica risulta convessa rispetto ai parametri  $\beta$ . Inoltre, tale formulazione risulta equivalente alla stima dei parametri tramite Maximum Likelihood Estimation (MLE), un approccio statistico che consiste nel determinare i parametri che massimizzano la probabilità di osservare i dati del dataset.[36]

Al termine dell'addestramento, la classe predetta viene determinata confrontando tale probabilità con una soglia  $\alpha$ , solitamente pari a 0.5. Se la probabilità stimata supera tale soglia l'istanza viene assegnata alla classe 1, altrimenti alla classe 0.

Formalmente, la regola di decisione del classificatore può essere espressa con il sistema seguente:

$$\hat{y} = \begin{cases} 1 & \text{se } P(Y = 1|X) \geq \alpha \\ 0 & \text{altrimenti} \end{cases} \quad (19)$$

### 3.1.2 Vantaggi

La LR presenta numerosi vantaggi dal punto di vista dell'implementazione, dell'addestramento e dell'interpretazione dei risultati. In particolare, il modello non richiede costi computazionali elevati, risultando quindi facilmente implementabile anche su dispositivi con risorse computazionali limitate.

Inoltre, sostituendo la funzione logistica con la funzione softmax, il modello può essere esteso in modo naturale a problemi di classificazione multiclasse.

Infine, attraverso il calcolo del log odds-ratio, è possibile interpretare l'effetto delle singole caratteristiche sull'aumento o sulla diminuzione della probabilità che un'istanza appartenga alla classe  $Y = 1$ . Questo rende la LR un modello facilmente interpretabile, caratteristica particolarmente importante in ambito clinico e biomedico.[36][37]

### 3.1.3 Limiti

Nonostante offra numerosi vantaggi, la LR è caratterizzata da alcune limitazioni rilevanti tra cui:

- L'assunzione di una relazione lineare tra le variabili di input e il *log odds-ratio*. Un presupposto che può risultare non soddisfatto in applicazioni reali, soprattutto quando emergono relazioni non lineari tra le caratteristiche. [37]
- L'incapacità di rappresentare strutture decisionali complesse. Infatti, rispetto ad approcci di classificazione più sofisticati, quali Decision Tree (DT) e Random Forest (RF), il metodo potrebbe non riuscire a cogliere pattern complessi presenti nei dati. [36][37]

- La sensibilità a caratteristiche fortemente correlate, condizione che potrebbe rendere instabili le stime dei coefficienti di regressione. [38]

### 3.2 Decision Tree (DT)

Un albero decisionale (Decision Tree, DT) è un modello di apprendimento supervisionato utilizzato per predire la classe di appartenenza di una data istanza percorrendo un cammino che parte dal nodo radice (root node) e termina in un nodo foglia (leaf node).

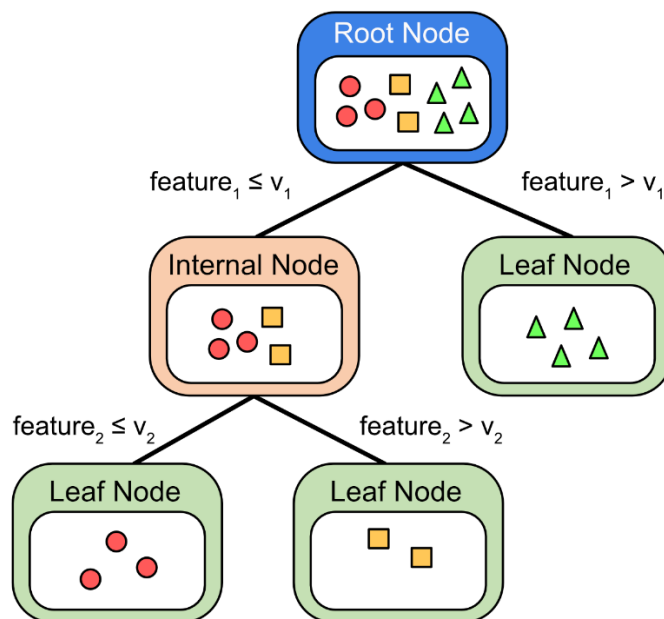


Figura 11: rappresentazione albero decisionale [39]

Come illustrato in Figura 11, ad ogni nodo interno dell'albero, il nodo successivo viene determinato sulla base di un criterio di suddivisione (split), definito da una delle caratteristiche delle istanze del dataset.

Una volta raggiunto un nodo foglia, la classe assegnata all'istanza corrisponde alla classe più frequente tra le istanze che hanno raggiunto quella foglia durante la fase di training. [36]

L'applicabilità del DT in ambito di screening risulta particolarmente rilevante grazie alla sua elevata interpretabilità. La struttura gerarchica del modello consente infatti di tracciare in modo esplicito il processo decisionale, evidenziando gli indicatori che determinano la classificazione della FFR come presente o assente. Ciò permette di individuare le variabili più influenti e di supportare l'identificazione dei principali fattori critici nei soggetti analizzati.

### 3.2.1 Costruzione dell'albero decisionale

La costruzione di un DT viene effettuata in modo ricorsivo.

Per ogni nodo si analizzano tutte le possibili modalità di partizione delle caratteristiche disponibili e si individua quella che consente di organizzare i dati in modo da rendere i nodi figli quanto più omogenei possibile rispetto alla classe di appartenenza.

La determinazione del miglior criterio di suddivisione avviene utilizzando una delle diverse misure di impurità disponibili. Le più utilizzate sono:

- *Classification error rate*: rappresenta il tasso d'errore di classificazione all'interno di un nodo e corrisponde alla quantità di istanze che non appartengono alla classe più comune nella regione considerata. Formalmente:

$$E = 1 - \max_k(p_{mk}) \quad (20)$$

Dove  $p_{mk}$  rappresenta la porzione di osservazioni del training set che nel m-esima regione appartengono alla classe k.

- *Indice di Gini*: misura di impurità del nodo utilizzato nell'algoritmo CART. Tale indice assume valori più bassi quando le osservazioni nel nodo appartengono prevalentemente alla stessa classe. Nel caso di  $K$  classi:

$$G = \sum_{i=1}^K p_{mk}(1 - p_{mk}) \quad (21)$$

- *Cross-entropia*: indice alternativo all'indice di Gini, spesso utilizzato nell'algoritmo IDE. In termini matematici:

$$D = - \sum_{i=1}^K p_{mk} \log(p_{mk}) \quad (22)$$

In generale, l'indice di Gini e la cross-entropia sono i criteri più utilizzati per individuare il miglior split, poiché risultano più sensibili alle variazioni di impurità dei nodi rispetto al classification error rate. Quest'ultimo, invece, viene spesso utilizzato nella fase di potatura dell'albero (pruning), il cui obiettivo è ridurre la complessità del modello e migliorare la capacità di generalizzazione, evitando fenomeni di overfitting.[36] [38]

### 3.2.2 Vantaggi

L'interpretabilità e la trasparenza rappresentano due dei principali punti di forza degli alberi decisionali. Questi modelli sono infatti facilmente comprensibili e visualizzabili, consentendo di motivare in modo chiaro ciascuna decisione assunta. [38] [40]

Inoltre, la loro semplice implementazione e la capacità di gestire dati eterogenei (numerici e categoriali) li rendono versatili, adatti a una vasta gamma di dataset e implementabili nella maggior parte dei dispositivi. [38]

Infine, rispetto a modelli lineari come la LR, gli alberi decisionali sono in grado di identificare relazioni complesse tra le variabili senza necessità di specificarle preventivamente, catturando automaticamente interazioni non lineari. [41]

### 3.2.3 Limiti

Nonostante la loro ampia diffusione, gli alberi decisionali presentano alcune importanti limitazioni. In particolare:

- L'utilizzo di alberi profondi può favorire il fenomeno dell'overfitting, compromettendo la capacità del modello di generalizzare su dati non visti. Tale problematica può essere attenuata mediante tecniche di potatura o restrizioni sulla profondità massima dell'albero. [38] [40]
- I singoli alberi mostrano una certa instabilità: lievi modifiche nei dati di input possono determinare significativi cambiamenti nella struttura dell'albero, rendendo il modello altamente sensibile ai dati di addestramento. [41]
- In contesti caratterizzati da dataset particolarmente complessi e di grandi dimensioni, un singolo albero decisionale tende a fornire prestazioni inferiori rispetto a modelli più avanzati, quali RF o Gradient Boosting, che sfruttano l'integrazione di molteplici alberi per incrementare l'accuratezza. [38]

## 3.3 Random Forest (RF)

Come suggerisce il nome, una Random Forest (RF) è un metodo di apprendimento supervisionato basato su un insieme di alberi decisionali indipendenti tra loro.

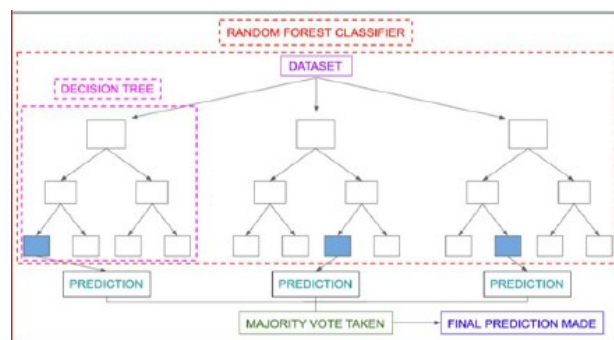


Figura 12: rappresentazione Random Forest. Immagine modificata da [41]

La RF, costruita come insieme di più DT, consente di ridurre la varianza del modello e migliorare la capacità predittiva. Pur essendo meno trasparente di un singolo albero, è possibile stimare l'importanza delle variabili e ottenere indicazioni sulle caratteristiche che maggiormente influenzano la classificazione della FFR come presente o assente. Questo la rende uno strumento adatto a contesti di screening dove è necessario bilanciare un'elevata accuratezza predittiva con un livello, seppur parziale, di interpretabilità delle decisioni.

### 3.3.1 Costruzione e predizione della RF

La costruzione e l'assegnazione delle classi da parte di una RF avvengono attraverso tre fasi principali.

La prima fase riguarda il campionamento dei dati, mediante una tecnica nota come Bagging. A partire dal dataset di addestramento originale vengono generati  $n$  sottoinsiemi di training mediante campionamento con reinserimento (bootstrapping). Questo processo è noto come Bootstrap Aggregating (bagging) e consente di creare dataset leggermente diversi tra loro, riducendo la varianza del modello.

Successivamente alla definizione degli  $n$  dataset di training si procede alla costruzione degli alberi decisionali. Per ciascun sottoinsieme di dati viene costruito un albero decisionale. Durante la crescita dell'albero, a ogni nodo non vengono considerate tutte le  $d$  caratteristiche disponibili, ma solo un sottoinsieme casuale di  $m$  caratteristiche, selezionate tra quelle originali. Questa procedura introduce ulteriore casualità nel modello e riduce la correlazione tra gli alberi. Una volta costruiti tutti gli alberi costituenti la RF, la classificazione di una nuova istanza viene ottenuta tramite un meccanismo di votazione (majority voting). In particolare, ogni albero fornisce una predizione e la classe finale assegnata all'istanza corrisponde alla classe più frequentemente predetta dagli alberi della foresta.

Grazie alla combinazione di bagging e selezione casuale delle feature, le RF riescono a ridurre la varianza senza intaccare il bias, migliorando la capacità di generalizzazione del

modello. Inoltre, questo metodo risulta generalmente robusto al rumore nei dati e capace di gestire dataset con un elevato numero di variabili. [38] [40]

### **3.3.2 Vantaggi**

Come già evidenziato, uno dei principali vantaggi delle RF risiede nella loro capacità di mitigare il problema dell'overfitting, tipico dei singoli alberi decisionali. Questo risultato è ottenuto grazie alla combinazione di molteplici alberi, costruiti su differenti sottoinsiemi del dataset e mediante una selezione casuale delle variabili, rendendo il modello complessivamente più stabile e con una migliore capacità di generalizzazione su dati non osservati [38][40].

Un ulteriore punto di forza è rappresentato dalla robustezza nei confronti del rumore presente nei dati. Poiché la previsione finale deriva dall'aggregazione delle predizioni di numerosi alberi, eventuali errori commessi dai singoli modelli tendono a compensarsi reciprocamente [38][40].

Le RF sono inoltre in grado di gestire dataset caratterizzati da un elevato numero di variabili senza richiedere particolari trasformazioni preliminari. A ciò si aggiunge la possibilità di stimare l'importanza delle variabili (feature importance), consentendo di individuare le caratteristiche che contribuiscono maggiormente al processo decisionale del modello [38].

Infine, questo metodo garantisce generalmente buone prestazioni predittive, risultando spesso più accurato rispetto a un singolo albero decisionale [36][38][41].

### **3.3.3 Limiti**

Uno dei principali limiti delle RF, rispetto ai singoli alberi decisionali, è rappresentato dalla minore interpretabilità del modello. Sebbene sia possibile analizzare l'importanza delle variabili, la presenza di un elevato numero di alberi rende complesso ricostruire in modo intuitivo il processo decisionale complessivo [38].

Un ulteriore svantaggio riguarda il costo computazionale. La costruzione e l'utilizzo di molteplici alberi comportano infatti un maggiore impiego di tempo e risorse, soprattutto nel caso di dataset di grandi dimensioni [36].

Infine, sebbene le RF siano generalmente robuste, possono mostrare prestazioni meno ottimali in presenza di dataset fortemente sbilanciati o caratterizzati da variabili altamente correlate. In tali situazioni, può risultare necessario ricorrere a tecniche aggiuntive di pre-processing o di bilanciamento dei dati per migliorare le performance del modello. [36][40]

### 3.4 Principal Component Analysis (PCA)

L'Analisi delle Componenti Principali (Principal Component Analysis, PCA) rappresenta una tecnica di apprendimento non supervisionato ampiamente in uso nell'ambito del machine learning e dell'analisi statistica dei dati.

Il suo scopo principale consiste nella riduzione della dimensionalità di un dataset preservando la maggior quantità possibile di informazione contenuta nei dati originali.

In particolare, la PCA consente di trasformare un insieme di variabili potenzialmente correlate in un nuovo insieme di variabili 'artificiali' tra loro ortogonali e non correlate, denominate componenti principali. Tali componenti sono ottenute come combinazioni lineari delle variabili originali e sono ordinate in modo tale che la prima componente spieghi la massima varianza possibile dei dati, la seconda la massima varianza residua sotto vincolo di ortogonalità rispetto alla prima, e così via. [36][37][40]

Nel caso in esame, utilizzando come input gli indicatori descritti nella Sezione 2.2, la PCA permette di sviluppare indici sintetici. Questi indici risultano particolarmente utili per definire la capacità uditiva complessiva del soggetto e per identificare eventuali problematiche nell'elaborazione cerebrale del suono, come l'incapacità di seguire la F0 dello stimolo o la difficoltà nell'elaborazione della fase transitoria dello stimolo presentato.

#### 3.4.1 Formulazione matematica

Supponendo di disporre di un dataset composto da  $N$  osservazioni, ciascuna descritta da  $d$  variabili numeriche. Ogni osservazione può essere rappresentata tramite un vettore:

$$x_i = [x_{i1}, x_{i2}, \dots, x_{id}]^T \quad (23)$$

L'intero dataset può quindi essere organizzato all'interno di una matrice dei dati  $X \in \mathbb{R}^{N \times d}$ , dove le righe rappresentano i campioni e le colonne rappresentano le caratteristiche.

Il primo passo della tecnica PCA consiste nel centrare i dati, ovvero nel sottrarre la media di ciascuna variabile. La media campionaria del dataset è definita come:

$$\bar{x} = \frac{1}{N} \sum_{i=1}^N x_i \quad (24)$$

I dati centrati si ottengono quindi tramite:

$$m_i = x_i - \bar{x} \quad (25)$$

Raccogliendo tutti i vettori centrati si ottiene la matrice dei dati centrati:

$$M = X - \bar{x} \quad (26)$$

In questo modo ogni variabile presenta media nulla, condizione necessaria per il corretto calcolo della matrice di covarianza. Quest'ultima riveste un ruolo fondamentale nell'analisi delle relazioni di dipendenza lineare tra le diverse caratteristiche del dataset. Il suo calcolo avviene secondo quanto riportato nell'Equazione 27:

$$\Sigma = \frac{1}{N} \sum_{i=1}^N (x_i - \bar{x})(x_i - \bar{x})^T \quad (27)$$

La matrice  $\Sigma$  è simmetrica e definita positiva, proprietà che consentono di effettuare una decomposizione agli autovalori, mediante il processo di Singular Value Decomposition (SVD). [37]

Il processo SVD prevede la diagonalizzazione della matrice di covarianza nel seguente modo:

$$U^T \Sigma U = \Lambda \quad (28)$$

dove:

- $U$  è una matrice ortogonale contenente gli autovettori della matrice di covarianza
- $\Lambda$  è una matrice diagonale contenente gli autovalori  $\lambda_1 \geq \lambda_2 \geq \dots \geq \lambda_d$

Gli autovettori identificano le direzioni principali di variazione dei dati, mentre gli autovalori quantificano la varianza spiegata lungo tali direzioni. [40]

I dati centrati possono essere proiettati nel nuovo sistema di coordinate definito dagli autovettori tramite la trasformazione:

$$r = U^T m \quad (29)$$

Il vettore  $r$  rappresenta la nuova rappresentazione dei dati nello spazio delle componenti principali. In questo nuovo spazio la matrice di covarianza assume la forma:

$$\text{cov}(r) = \Lambda \quad (30)$$

ovvero una matrice diagonale, il che implica che le nuove variabili risultano non correlate tra loro.

Uno dei principali vantaggi della PCA consiste nella possibilità di ridurre la dimensionalità dei dati selezionando soltanto le prime  $k$  componenti principali, cioè quelle associate agli autovalori più elevati.

Indicando con  $U_k$  la matrice contenente i primi  $k$  autovettori, la trasformazione diventa:

$$r_k = U_k^T m \quad (31)$$

In questo modo i dati vengono rappresentati in uno spazio di dimensione ridotta mantenendo la maggior parte della varianza del dataset originale.

### 3.4.2 Vantaggi

Oltre alla capacità di ridurre la dimensionalità dei dati rendendoli più facilmente interpretabili, la PCA presenta ulteriori vantaggi.

In primo luogo, si tratta di una tecnica computazionalmente efficiente, soprattutto se confrontata con metodi di riduzione della dimensionalità di tipo non lineare. La sua formulazione matematica, basata sulla decomposizione della matrice di covarianza, permette infatti di applicarla in modo efficiente anche a dataset di grandi dimensioni. [37]

Un ulteriore vantaggio consiste nella possibilità di analizzare il contributo delle variabili originali alle componenti principali. Osservando i coefficienti degli autovettori, infatti, è possibile comprendere quali caratteristiche influenzano maggiormente ciascuna componente, facilitando l'individuazione di pattern o strutture latenti nei dati. [36] [37][40]

Infine, la PCA permette spesso di ridurre il rumore presente nei dati, poiché le componenti principali associate agli autovalori più piccoli tendono a catturare variazioni meno significative o rumore statistico. [37]

### 3.4.3 Limiti

Uno dei principali limiti riguarda l'assunzione di linearità: la PCA identifica infatti solo relazioni lineari tra le variabili. Di conseguenza, nel caso in cui la struttura dei dati sia

caratterizzata da relazioni non lineari, la PCA potrebbe non riuscire a rappresentare adeguatamente la struttura sottostante del dataset. [37]

Un'ulteriore limitazione è legata alla sensibilità alla scala delle variabili. Se le caratteristiche del dataset sono misurate su scale molto diverse, le variabili con varianza maggiore possono dominare l'analisi. Per questo motivo, nella pratica è spesso necessario applicare una fase preliminare di standardizzazione dei dati.[37]

Infine, le componenti principali ottenute sono combinazioni lineari delle variabili originali, il che può rendere talvolta difficile interpretarne il significato in termini delle caratteristiche iniziali del dataset. [37][40]

# Capitolo 4

## Materiali e Metodi

Nelle sezioni successive verranno descritte le diverse fasi di acquisizione e preelaborazione delle FFR, che costituiscono la base per la successiva post-elaborazione, estrazione degli specifici indicatori e valutazione dell'efficacia dei modelli di machine learning presentati nel Capitolo 3. Le FFR sono state acquisite da tre differenti gruppi di soggetti: il primo gruppo è costituito da individui normoudenti, il secondo comprende un paziente con lesione ischemica, mentre il terzo è composto esclusivamente da soggetti che presentano impianto cocleare.

### 4.1 Dataset reale 1: soggetti normoudenti e patologici

Il primo dataset è costituito da registrazioni FFR provenienti da soggetti normoudenti e da un individuo affetto da lesione ischemica. Per garantire la ripetibilità e l'affidabilità degli indicatori estratti, è stato adottato un protocollo sperimentale progettato per minimizzare le interferenze e ottimizzare l'SNR.

La procedura ha previsto l'utilizzo di strumentazione specializzata, un accurato posizionamento degli elettrodi, condizioni controllate di stimolazione e la definizione di parametri di acquisizione conformi agli standard riportati nella letteratura scientifica.

#### 4.1.1 Condizioni sperimentali

Il dataset analizzato, fornito da Hedera Biomedics S.R.L, include 49 soggetti, di cui 48 sani e uno affetto da lesione ischemica, tutti sottoposti precedentemente a test ABR. I partecipanti hanno un'età compresa tra 18 e 70 anni ( $45,28 \pm 21,12$ ) e presentano una buona diversità di genere (23 femmine e 25 maschi).

Durante l'acquisizione, i soggetti sono stati mantenuti svegli in posizione prona con occhi chiusi, condizione standard per l'acquisizione delle risposte FFR non influenzate dall'attenzione cosciente. Le registrazioni sono state effettuate in un ambiente silenzioso e schermato, al fine di minimizzare qualsiasi interferenza ambientale. Per ogni partecipante sono stati applicati i medesimi parametri sperimentali e la stessa procedura di preelaborazione, come dettagliato in seguito.

La stimolazione sonora presentata consisteva in fonemi sintetici sviluppati dal gruppo della Northwestern University. La velocità di stimolazione è stata variata a seconda della durata dello stimolo per bilanciare alto SNR con il rischio di adattamento o di risposte sovrapposte. [25]

La sillaba /da/ è stato lo stimolo maggiormente analizzato tra quelli presentati. Esso è stato somministrato tre volte al secondo ottenendo 1000 sweeps per paziente. La finestra temporale di acquisizione era di 120 ms e ciascuna sweep includeva 500 campioni, ottenendo una frequenza di campionamento di 4167 Hz. Tutte le registrazioni si sono svolte in un ambiente schermato per limitare interferenze elettriche e artefatti muscolari o oculari.

Sebbene gli artefatti da movimento siano meno frequenti nell'FFR rispetto ai potenziali corticali, sono state considerate e mitigate tutte le possibili fonti di disturbo fisiologico per ridurre al minimo le distorsioni della risposta.

#### **4.1.2 Stimolazione**

Gli stimoli acustici sono stati erogati a tutti i soggetti tramite auricolari insert Etymotic ER-3C, conformi allo standard ISO 0389-1-9 e calibrati con strumenti professionali. La scelta di questi auricolari ha garantito una protezione efficace dalle interferenze elettromagnetiche e una chiusura ottimale del canale uditivo, preservando così l'integrità del segnale.

La sillaba analizzata in questo studio corrisponde al fonema /da/, caratterizzato da una F0 di 100 Hz e da una durata complessiva di 120 ms, così suddivisa: i primi 10 ms costituiscono il burst consonantico, seguiti da 50 ms di transizione, mentre gli ultimi 60 ms rappresentano la fase stazionaria della vocale.

Per evidenziare le componenti della risposta, sia in termini di ENV che di TFS, e per ridurre i fenomeni di artefatto cocleare e microfonico, gli stimoli sono stati presentati in polarità alternata (condensazione e rarefazione).

#### **4.1.3 Strumentazione per l'acquisizione**

Le FFR sono state acquisite tramite l'applicazione di elettrodi sullo scalpo dei soggetti, collegati al dispositivo Socrates (Figura 13). Tale dispositivo è stato sviluppato da Hedera Biomedics S.R.L., un'azienda italiana specializzata nella progettazione e produzione di dispositivi biomedicali e software clinici.



*Figura 13: Dispositivo base Socrates utilizzato per registrare i segnali FFR. A sinistra visione frontale, a destra visione posteriore.[24]*

Socrates, grazie all'impiego di due canali di acquisizione indipendenti, è in grado di rilevare in modo affidabile i potenziali evocati uditivi a bassa intensità generati da stimolazione acustica.

Come supporto al dispositivo, l'azienda ha sviluppato l'interfaccia operativa, chiamata Helix, utile per ottimizzare l'efficienza dei processi di acquisizione e minimizzare il rischio di errore umano. Un esempio di acquisizione viene mostrato in Figura 14.

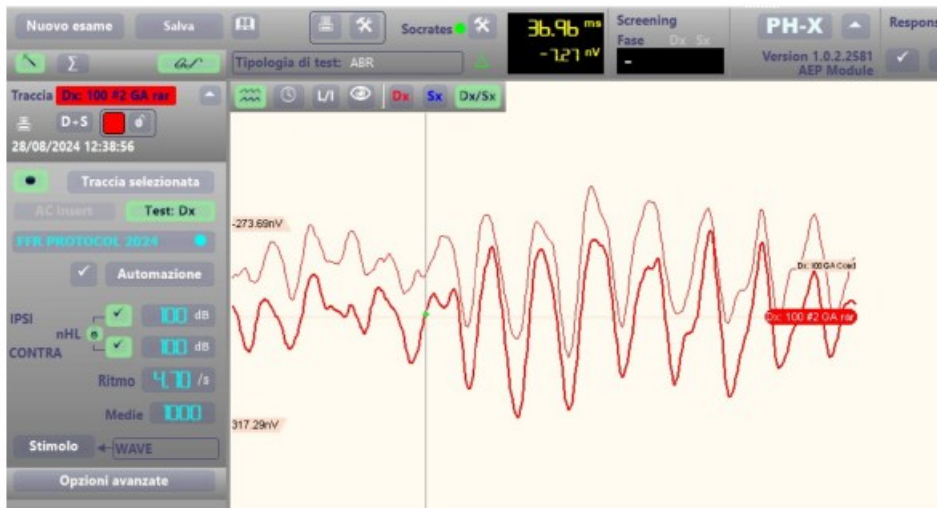


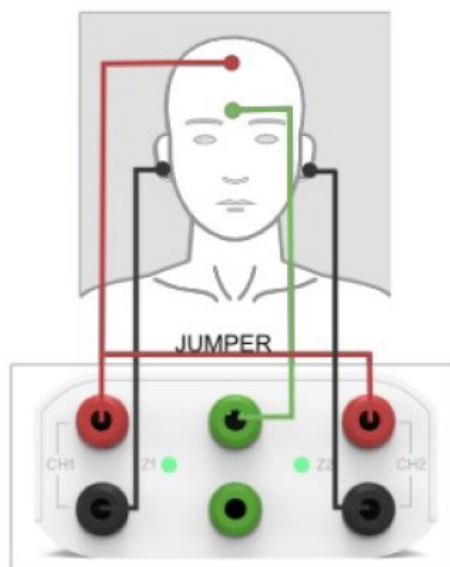
Figura 14: esempio di FFR registrata e visualizzata attraverso interfaccia Helix[28]

Attraverso l'interfaccia *Helix* è possibile:

- *Gestire i dati*: il software consente di ottenere dati già predisposti per l'analisi in vari formati, applicando automaticamente filtraggio, riduzione del rumore e preelaborazione;
- *Configurare i parametri di acquisizione in modo efficiente*: la gestione del secondo livello di amplificazione, successiva a una pre-amplificazione fissa che assicura la stabilità del segnale, garantisce elevata flessibilità d'utilizzo;
- *Visualizzare il segnale in tempo reale*: la rappresentazione istantanea del segnale medio acquisito permette di monitorare costantemente la qualità delle registrazioni.

Complessivamente, tali caratteristiche rendono *Socrates* un dispositivo estremamente versatile e in grado di limitare efficacemente le interferenze di origine muscolare o ambientale.

Il montaggio utilizzato per gli elettrodi è il montaggio verticale a canale singolo, che prevede l'impiego di tre elettrodi Ag-AgCl monouso e pre-gelati, posizionati come illustrato in Figura 15.



*Figura 15: Montaggio verticale degli elettrodi per acquisizione FFR: CH indica il canale, Z il led di impedenza. Elettrodo positivo in rosso, negativo in nero, comune in verde.*

L'elettrodo attivo viene collocato sulla sommità del cranio, in corrispondenza della posizione Cz. L'elettrodo di riferimento viene posizionato in A2, corrispondente al lobo auricolare destro o sinistro, in base al lato della stimolazione acustica. Infine, l'elettrodo di massa viene collocato sulla fronte del soggetto, in corrispondenza della posizione Fpz.

Nel rispetto delle norme vigenti per la sicurezza del paziente, la riduzione degli artefatti e il mantenimento di una buona qualità del segnale, i limiti superiori delle impedenze sono stati fissati a 5 k $\Omega$  per ciascun elettrodo e a 2 k $\Omega$  per le impedenze interelettrodiche. Tali valori sono stati verificati e controllati prima dell'inizio dell'acquisizione dei segnali FFR.

#### **4.1.4 Pre-processing**

Successivamente all'acquisizione, ogni sweep è stata sottoposta alla seguente procedura. In primo luogo, i segnali sono stati sottoposti ad amplificazione e filtraggio analogico mediante un filtro passa-banda 0–4000 Hz. Il filtro Notch a 50 Hz è stato disattivato per evitare distorsioni sulle armoniche con frequenze prossime a quella di rete.

Successivamente, i dati grezzi sono stati importati in Matlab (R2024B) tramite la funzione 'readfile.m', sviluppata da Hedera Biomedics S.R.L. Questa funzione converte i segnali in nanovolt (nV) e li organizza in una matrice, dove le righe rappresentano il numero di sweep e le colonne corrispondono al numero di campioni acquisiti.

Infine, i dati sono stati sottoposti a un filtraggio digitale aggiuntivo mediante un filtro Butterworth del quarto ordine nella banda 80–2000 Hz, al fine di isolare le componenti spettrali principali dell'FFR, come la F0 e le armoniche, escludendo le componenti derivanti da attività corticali e i rumori ad alta frequenza che possono degradare il segnale.[25]

## 4.2 Dataset reale 2: soggetti portatori di impianto cocleare (IC)

### 4.2.1 Condizioni sperimentali utilizzate

I pazienti inclusi nel dataset in analisi, fornito da Hedera Biomedics S.R.L, sono pazienti con IC seguiti dall'ospedale di Palermo. Il dataset acquisito comprende 14 soggetti, di cui uno utilizza anche un apparecchio acustico.

Le condizioni sperimentali e i dispositivi impiegati per la registrazione delle FFR sono analoghi a quelli del dataset precedente, ad eccezione di due differenze relative alla modalità di presentazione dello stimolo e al numero di registrazioni acquisite. Tali differenze sono descritte nelle sezioni successive.

Una breve descrizione dell'IC è riportata nell'Appendice A.

### 4.2.2 Stimolazione

A differenza dei soggetti sani, lo stimolo sonoro è stato presentato in campo libero mediante l'uso di due casse attive, calibrate e poste 80 cm dal soggetto con un'inclinazione dei coni a 45 gradi. Questo tipo di approccio, come dimostrato anche in Figura 17, può aver influenzato notevolmente l'intensità delle risposte, a causa dell'incapacità del microfono di riprodurre risposte simili a quelle generate dagli auricolari nei soggetti sani.

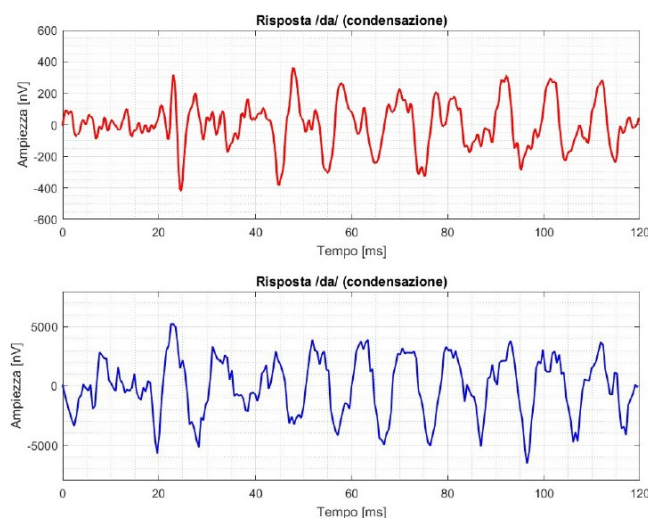


Figura 16: confronto FFR in soggetto sano (in alto) vs FFR in soggetto con IC (in basso) [25]

### **4.2.3 Registrazione**

La seconda differenza, invece, riguarda la quantità di registrazioni ottenute. Mentre nei soggetti sani venivano rilevate 1000 registrazioni per paziente e per stimolazione, nel presente dataset ne sono state registrate solo 500 per stimolazione. Inoltre, ogni registrazione presenta una durata 240 ms, al fine di acquisire anche 120 ms di prestimolo, utile per le analisi successive.

## **4.3 Addestramento ed analisi dei modelli**

Nella presente sezione verranno descritti i diversi passaggi utilizzati per l'estrazione dei parametri impiegati nell'addestramento dei modelli, le modalità di addestramento adottate e le metriche utilizzate per la valutazione delle prestazioni ottenute sui diversi dataset.

Le procedure descritte sono state applicate in maniera analoga sia al dataset composto da pazienti sani e patologici, sia al dataset costituito da soggetti portatori di IC.

### **4.3.1 Estrazione delle caratteristiche e assegnazione classe**

Le caratteristiche estratte corrispondono agli indicatori descritti nella Sezione 2.2. In particolare, gli indicatori sono stati calcolati considerando un numero crescente di sweep, portando alla formazione di otto diversi dataset per i soggetti normoudenti e patologici (50,100,175,250,350,500,750 e 1000 sweeps) e sei diversi dataset per i soggetti portatori di impianto cocleare (50,100,175,250,350 e 500 sweeps) per ciascun tipo di post-elaborazione utilizzata (CA, APWF e SB, vedi capitolo 2.1). In totale sono stati quindi ottenuti 24 dataset per i soggetti normoudenti e patologici e 18 dataset per i pazienti con IC.

Questa procedura ha consentito di valutare le prestazioni dei modelli al variare della quantità di registrazioni considerate, permettendo in tal modo di definire se i modelli sono affidabili anche utilizzando un numero inferiore di registrazioni rispetto agli standard convenzionali.

Successivamente, tra gli indicatori estratti sono stati selezionati come ingressi dei modelli solamente quelli rilevati in assenza di rumore. Tale selezione è avvenuta per evitare l'eventuale mancanza di dati nelle diverse condizioni di rumore (White Noise e Two Babble Talker), poiché non tutti i soggetti presentavano valutazioni per entrambe le condizioni considerate.

Gli input utilizzati nei diversi modelli sono quindi: RC della vocale, RC della consonante, PC della vocale, PC della consonante, FE, SE e PS (sezione 2.2.1, 2.2.2 e 2.2.4).

A queste si è aggiunta un'ulteriore caratteristica dicotomica nominata 'tipo di stimolazione' (Rarefazione/Condensazione). La sua aggiunta è stata eseguita poiché non per tutti i soggetti

era presente la doppia stimolazione. In totale, quindi, si sono state considerate otto caratteristiche per istanza in analisi.

I modelli di machine learning utilizzati sono modelli supervisionati di classificazione, il che implica che la classe reale di appartenenza delle istanze nel dataset di addestramento sia già nota. Poiché nei dataset analizzati questa informazione non era disponibile, è stato necessario procedere preliminarmente all'assegnazione della classe a ciascuna istanza.

A tal fine, sono state definite come assenti le FFR registrate nel prestimolo oppure, nel caso dei soggetti normoudenti e patologici, quando il test eseguito con click-ABR risultava piatto a qualsiasi tipo di soglia.

Definite le classi di appartenenza, sono state ottenute approssimativamente le seguenti percentuali per ogni dataset in analisi: per i dataset relativi ai soggetti normoudenti e patologici l'84% delle istanze è stato classificato come FFR presente e il 16% come FFR assente; mentre per i soggetti con IC il 58% delle istanze è stato classificato come FFR presente e il 42% come FFR assente.

Lo squilibrio tra i due dataset è dovuto al fatto che il prestimolo non era disponibile per tutti i soggetti normoudenti mentre per i soggetti con IC, come descritto nella sezione 4.2.3, il prestimolo viene sempre acquisito.

## **4.3.2 Addestramento e valutazione metodi di classificazione**

### **4.3.2.1 Addestramento**

Al fine di garantire la confrontabilità tra i diversi modelli di classificazione, è stata adottata la seguente procedura:

#### *1. Divisione del dataset*

Il dataset di partenza è stato diviso in training set, pari al 75% del dataset completo, e test set, pari al restante 25%. Tale divisione è avvenuta considerando di mantenere omogenee le classi all'interno dei sottoinsiemi con lo scopo di preservare un'adeguata rappresentazione delle classi sia nell'addestramento che nella validazione.

#### *2. Validazione incrociata*

Sul training set è stata applicata una 5-fold cross-validation al fine di selezionare gli iperparametri ottimi e valutare la robustezza dei modelli [36][38]

#### *3. Addestramento algoritmo sull'intero training set*

Determinati gli iperparametri ottimali, i modelli sono stati addestrati nuovamente sul set di training completo.

#### 4. Determinazione soglia decisionale ottimale

Le probabilità di appartenenza alla classe positiva stimate dai modelli sono state utilizzate per calcolare le principali metriche di classificazione e individuare la soglia decisionale ottimale, oltre la quale l'osservazione viene assegnata alla classe 1.

Lo sviluppo degli algoritmi utilizzati viene riportato in appendice B. In particolare, la LR binaria è stata implementata mediante Ridge Regression. Mentre il DT e la RF seguono algoritmo CART (Classification and Regression Trees).[36][38][40]

#### 4.3.2.2 Metriche di valutazione

Le performance dei modelli sono state valutate nel test set mediante l'utilizzo delle seguenti metriche: Accuratezza, Sensibilità, Specificità, ROC e AUC.

Le prime tre metriche possono essere ottenute facilmente costruendo la Confusion Matrix, una matrice contenente tutti i possibili output del modello in analisi.

	Classe Reale 1	Classe Reale 0
Classe Predetta 1	True Positive	False Positive
Classe Predetta 0	False Negative	True Negative

Tabella 1: esempio di Confusion Matrix

In Tabella 2, si riporta un esempio di Confusion Matrix, in cui:

- True Positive (TP) = numero di volte in cui la classe reale è 1 e la classe predetta è 1
- False Positive (FP) = numero di volte in cui la classe reale è 0 e la classe predetta è 1
- False Negative (FN) = numero di volte in cui la classe reale è 1 e la classe predetta è 0
- True Negative (TN) = numero di volte in cui la classe reale è 0 e la classe predetta è 0

Sfruttando tali quantità, si ricavano facilmente le seguenti metriche di valutazione del modello:

- *Accuratezza di predizione*: quantità di volte in cui il modello ha determinato correttamente la classe di appartenenza delle istanze. Formalmente:

$$Accuratezza = \frac{TP + TN}{TP + TN + FP + FN} \quad (32)$$

- *Sensibilità (True Positive Ratio, TPR)*: rapporto tra un numero di volte in cui il modello ha identificato la classe 1 e il numero totale di istanze appartenenti alla classe 1.

$$TPR = \frac{TP}{TP + FN} \quad (33)$$

- *Specificità*: rapporto tra un numero di volte in cui il modello ha identificato la classe 0 e il numero totale di istanze appartenenti alla classe 0.

$$\text{Specificità} = \frac{TN}{TN + FP} \quad (34)$$

L'utilizzo congiunto di queste metriche consente di ottenere una valutazione più completa delle prestazioni del modello, poiché la sola accuratezza potrebbe risultare fuorviante, soprattutto in presenza di dataset sbilanciati.

La Receiver Operating Characteristic Curve (ROC) è una curva che rappresenta la relazione tra il True Positive Rate (TPR), riportato sull'asse delle ordinate, e il False Positive Rate (FPR), riportato sull'asse delle ascisse, dove:

$$FPR = \frac{FP}{FP + TN} = 1 - \text{specificità} \quad (35)$$

Questa curva risulta particolarmente utile quando si vogliono confrontare diversi classificatori oppure lo stesso classificatore con differenti parametri di apprendimento.

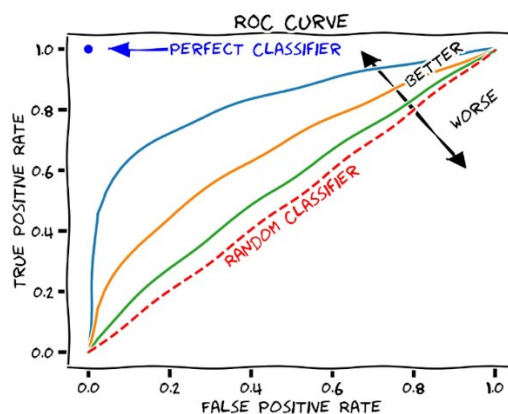


Figura 18: rappresentazione ROC [46]

Come illustrato in Figura 18, un classificatore che si colloca lungo la diagonale principale del grafico ROC presenta prestazioni equivalenti a quelle di un classificatore casuale (Random Classifier o Chance Classifier), ovvero assume che le classi positive e negative siano

equiprobabili. L'obiettivo di un classificatore è quindi quello di ottenere prestazioni superiori a tale riferimento, posizionandosi al di sopra della diagonale.

In generale, quanto più la curva ROC si avvicina all'angolo superiore sinistro del grafico, tanto migliori risultano le prestazioni del modello.

Un ulteriore metodo per confrontare i classificatori consiste nel calcolare l'area sottesa alla curva ROC, nota come Area Under the Curve (AUC). Questa metrica fornisce una misura sintetica delle prestazioni del classificatore. Ovvero valori di AUC più elevati indicano una migliore capacità del modello di distinguere tra le diverse classi.

### 4.3.3 Formazione e valutazione indici di capacità uditiva

#### 4.3.3.1 Costruzione indici

Gli indici di capacità uditiva sono stati sviluppati secondo la seguente procedura:

1. *Applicazione PCA*

La PCA è stata applicata agli indicatori selezionati al fine di ridurre la dimensionalità del dataset. È stato selezionato il numero minimo di componenti principali in grado di spiegare almeno il 90% della varianza totale dei dati.

2. *Costruzione degli indici multipli.*

I dati sono stati rappresentati nello spazio delle componenti principali selezionate. Le prime due componenti sono state utilizzate per costruire indici parziali, normalizzati nel range 0–100 rispetto ai valori minimo e massimo osservati, ottenendo valori percentuali per ogni soggetto.

3. *Costruzione dell'indice globale (GI)*

L'indice complessivo è stato ottenuto come combinazione lineare delle componenti principali più informative, pesate in base alla loro varianza spiegata. Formalmente:

$$\text{Global Index} = w_1PC_1 + w_2PC_2 + \dots + w_nPC_n \quad (36)$$

Dove  $w = [w_1, \dots, w_n]$  è il vettore dei pesi associati alle componenti principali.

Un approccio analogo è stato proposto in letteratura, in cui la PCA viene applicata direttamente ai segnali FFR per estrarre componenti rilevanti ai fini della discriminazione percettiva.[47]

### 4.3.3.2 Metriche di valutazione

La valutazione degli indici formati è stata effettuata in termini di affidabilità e robustezza utilizzando i seguenti metodi:

- *Bartlett test*, per verifica della presenza di correlazioni tra le variabili e quindi la possibilità di applicare la PCA.
- *5-fold cross validation*, utilizzata per valutare la robustezza degli indici. La stabilità per ogni indice è stata quantificata mediante correlazione tra l'indice ottenuto sull'intero dataset e quello ottenuto nei diversi fold di validazione.
- *Correlazione tra indici e indicatori FFR*, per analisi della relazione tra l'indice sintetico e le proprietà originali.
- *Test in soggetto con ictus pre e post Deep Brain Stimulation (DBS)*: valutazione dell'efficacia dell'indice nel rilevare i miglioramenti ottenuti dal soggetto durante il percorso riabilitativo.

# Capitolo 5

## Risultati sulla rilevazione della presenza/assenza delle FFR

Nel seguente capitolo verranno discussi i risultati ottenuti dai diversi modelli per i diversi metodi di post elaborazione. Per motivi di sintesi si riporteranno solamente i risultati ottenuti per numero di sweeps pari a 100, 250,500 e 1000 per il dataset dei soggetti sani e patologici; mentre per il dataset dei soggetti con IC si riporteranno i risultati ottenuti per 100, 250 ,350 e 500 sweeps. I risultati complessivi possono essere consultati all'interno dell'Appendice A.

### 5.1 Dataset normoudenti e patologici

#### *Rilevazione FFR con LR*

Di seguito viene riportata la tabella contenente i valori dell'area sottostante alla ROC per la LR applicata al test set.

AUC				
<b>SWEEP</b>	<b>100</b>	<b>250</b>	<b>500</b>	<b>1000</b>
<b>CA</b>	100%	94%	89%	100%
<b>APWF</b>	53%	89%	78%	100%
<b>SB</b>	89%	89%	100%	100%

*Tabella 2: AUC LR per normoudenti e patologici*

Come illustrato in tabella, gli indicatori derivati dalle FFR ottenute tramite CA e SB mostrano che, già con sole 100 registrazioni, è possibile ottenere modelli con valori di AUC nettamente superiori a quelli di uno Chance Classifier. Questo implica una buona capacità discriminativa tra FFR presenti e assenti anche con un numero limitato di osservazioni.

Al contrario, APWF sembrerebbe richiedere più registrazioni per ottenere prestazioni comparabili con gli altri due metodi. Le sue prestazioni, infatti, diventano comparabili con CA e SB solo dopo 250 registrazioni. Questa differenza potrebbe essere dovuta a una minore robustezza nella stima delle caratteristiche con pochi dati disponibili.

<b>ACCURATEZZA</b>				
<b>SWEEP</b>	<b>100</b>	<b>250</b>	<b>500</b>	<b>1000</b>
<b>CA</b>	91%	91%	82%	91%
<b>APWF</b>	64%	73%	82%	100%
<b>SB</b>	91%	73%	82%	100%

<b>SENSIBILITA'</b>				
<b>SWEEPS</b>	<b>100</b>	<b>250</b>	<b>500</b>	<b>1000</b>
<b>CA</b>	89%	100%	89%	100%
<b>APWF</b>	67%	78%	89%	100%
<b>SB</b>	89%	78%	78%	100%

<b>SPECIFICITA'</b>				
<b>SWEEPS</b>	<b>100</b>	<b>250</b>	<b>500</b>	<b>1000</b>
<b>CA</b>	100%	50%	50%	50%
<b>APWF</b>	50%	50%	50%	100%
<b>SB</b>	100%	50%	100%	100%

*Tabella 3,4,5: tabelle metriche di valutazione LR*

I risultati relativi ad accuratezza, sensibilità e specificità confermano quanto evidenziato dall'AUC. In generale, accuratezza e sensibilità assumono valori elevati per tutti i metodi, indicando una buona capacità del modello sia nel classificare correttamente i casi complessivi sia nell'identificare le FFR presenti.

La specificità, invece, assume esclusivamente valori pari al 50% o al 100%. Questo comportamento è riconducibile allo sbilanciamento delle classi dataset: il numero di soggetti appartenenti alla classe 0 (FFR assenti) risulta infatti molto ridotto. Di conseguenza, anche un singolo errore di classificazione su questi pochi casi comporta una variazione significativa della specificità, riducendola drasticamente.

#### *Rilevazione FFR con DT*

<b>AUC</b>				
<b>SWEEPS</b>	<b>100</b>	<b>250</b>	<b>500</b>	<b>1000</b>
<b>CA</b>	64%	72%	89%	69%
<b>APWF</b>	61%	75%	69%	94%
<b>SB</b>	61%	75%	69%	69%

*Tabella 6: AUC DT*

Nel caso del DT, si osserva come i valori di AUC risultino complessivamente inferiori rispetto a quelli ottenuti con la LR, pur mantenendosi su livelli discreti e mai inferiori al 60%.

<b>ACCURATEZZA</b>				
<b>SWEEPS</b>	<b>100</b>	<b>250</b>	<b>500</b>	<b>1000</b>
<b>CA</b>	64%	82%	82%	82%
<b>APWF</b>	64%	91%	82%	91%
<b>SB</b>	64%	91%	82%	82%

<b>SENSIBILITA'</b>				
<b>SWEEPS</b>	<b>100</b>	<b>250</b>	<b>500</b>	<b>1000</b>
<b>CA</b>	67%	89%	78%	89%
<b>APWF</b>	67%	100%	89%	89%
<b>SB</b>	67%	100%	89%	89%

<b>SPECIFICITA'</b>				
<b>SWEEPS</b>	<b>100</b>	<b>250</b>	<b>500</b>	<b>1000</b>
<b>CA</b>	50%	50%	100%	50%
<b>APWF</b>	50%	50%	50%	100%
<b>SB</b>	50%	50%	50%	50%

*Tabella 7,8,9: tabelle metriche di valutazione DT*

Analizzando le metriche di accuratezza, sensibilità e specificità, si conferma quanto suggerito dall'AUC. In particolare, quando l'AUC supera circa il 65%, si osserva un incremento significativo sia dell'accuratezza sia della sensibilità, indicando una migliore capacità del modello di identificare correttamente i casi positivi.

La specificità, invece, mostra un andamento analogo a quello riscontrato nel caso della LR, assumendo valori pari al 50% o al 100%. Anche in questo caso, tale comportamento è attribuibile allo sbilanciamento delle classi.

Nel complesso, i tre metodi di post-elaborazione (CA, APWF e SB) producono indicatori che conducono a prestazioni piuttosto simili tra loro nel contesto del DT. Questo comportamento differisce da quanto osservato con la LR, dove emergevano differenze più marcate tra i metodi. Il motivo di tale differenza tra i due modelli risiede nella natura del DT. Esso, infatti, essendo un modello non lineare basato su soglie, tende a ridurre l'impatto di differenze minime tra le caratteristiche dell'istanza portando a generare prestazioni simili anche in presenza di diversi metodi di post-elaborazione. Al contrario, la LR, essendo un modello lineare, risulta più

sensibile alla qualità e alla distribuzione degli indicatori, evidenziando maggiormente le differenze tra i metodi considerati.

#### *Rilevazione FFR Con RF*

La RF mostra prestazioni complessivamente buone applicata agli indicatori ottenuti tramite SB e APWF già a partire da 100 sweeps. Al contrario, per il metodo CA sembra essere necessario un numero maggiore di registrazioni (almeno 500) per ottenere prestazioni comparabili.

<b>AUC</b>				
<b>SWEEPS</b>	<b>100</b>	<b>250</b>	<b>500</b>	<b>1000</b>
<b>CA</b>	75%	56%	69%	92%
<b>APWF</b>	92%	89%	86%	100%
<b>SB</b>	69%	92%	89%	83%

*Tabella 10: AUC RF*

Analizzando i valori dell'AUC, si osserva come APWF presenti risultati particolarmente elevati, fino a raggiungere il 100% con 1000 sweeps. Tuttavia, tale andamento non si riflette in modo coerente sull'accuratezza, che risulta invece ridotta per 500 e 1000 registrazioni. Un comportamento analogo si osserva anche per la sensibilità, mentre la specificità continua a mostrare valori estremi (50% o 100%), in linea con quanto osservato in precedenza.

<b>ACCURATEZZA</b>				
<b>SWEEPS</b>	<b>100</b>	<b>250</b>	<b>500</b>	<b>1000</b>
<b>CA</b>	91%	91%	73%	82%
<b>APWF</b>	91%	91%	45%	64%
<b>SB</b>	73%	73%	91%	73%

<b>SENSIBILITA'</b>				
<b>SWEEPS</b>	<b>100</b>	<b>250</b>	<b>500</b>	<b>1000</b>
<b>CA</b>	100%	100%	78%	78%
<b>APWF</b>	100%	100%	33%	56%
<b>SB</b>	78%	67%	100%	67%

<b>SPECIFICITA'</b>				
<b>SWEEPS</b>	<b>100</b>	<b>250</b>	<b>500</b>	<b>1000</b>
<b>CA</b>	50%	50%	50%	100%
<b>APWF</b>	50%	50%	100%	100%
<b>SB</b>	50%	100%	50%	100%

*Tabella 11,12,13: tabelle metriche di valutazione RF*

Questa discrepanza tra AUC e le altre metriche può essere attribuita alle caratteristiche della RF in presenza di dataset di dimensioni ridotte. In tali condizioni, gli alberi decisionali che compongono la foresta possono risultare tra loro più correlati, riducendo l'efficacia del meccanismo di aggregazione e aumentando il rischio di overfitting. Questo può portare a una minore capacità di generalizzazione, con conseguente riduzione di accuratezza e sensibilità, nonostante valori di AUC elevati.

### Modelli A Confronto

Nel complesso, tutti i modelli analizzati risultano efficacemente utilizzabili per l'identificazione della presenza o assenza di FFR già a partire da circa 250 registrazioni.

A supporto di queste considerazioni, è possibile analizzare le curve ROC dei diversi modelli, che consentono un confronto diretto della loro capacità discriminativa.

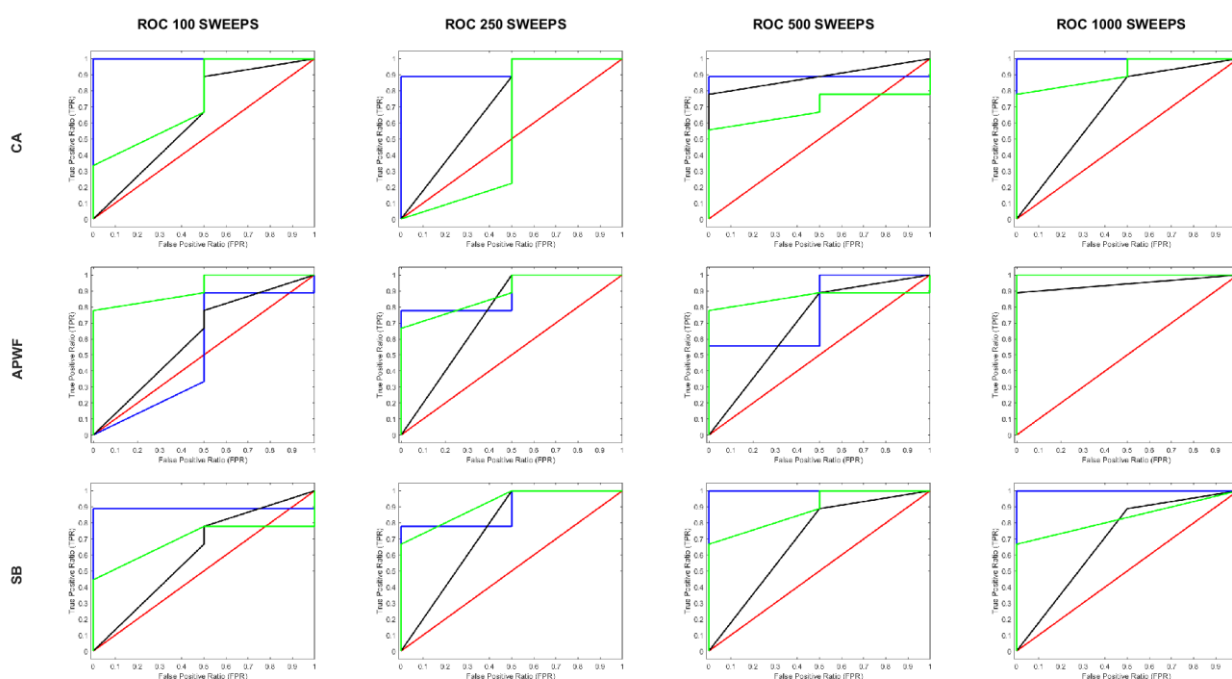


Figura 19: confronto tra ROC di Chance Classifier (rosso), LR (blu), DT (nero) e RF (verde) per numero crescente di registrazioni per i diversi metodi di post elaborazione

Come illustrato nelle figure precedenti, il confronto tra le curve ROC evidenzia come i modelli presentino capacità discriminative differenti in funzione sia del metodo di post-elaborazione sia del numero di sweeps considerati.

In particolare, le prime tre immagini in alto mostrano le prestazioni dei diversi modelli di machine learning applicati agli indicatori ottenuti tramite CA per numero crescente di

registrazioni. Dall'analisi delle ROC si osserva come la LR rappresenti il metodo più performante già a partire da 100 sweeps.

In linea con quanto evidenziato dall'AUC, la RF presenta prestazioni inferiori, risultando al di sotto dello Chance Classifier per una porzione significativa della curva a 250 registrazioni, per poi migliorare a partire da 500 sweeps. Il DT, invece, si mantiene costantemente al di sopra dello Chance Classifier, ma con prestazioni inferiori rispetto alla LR.

Per quanto riguarda gli indicatori ottenuti tramite APWF, le cui ROC sono rappresentate nella seconda riga, si osserva un comportamento differente. In questo caso, la LR presenta prestazioni inferiori allo Chance Classifier a 100 sweeps per una parte rilevante della curva ROC, migliorando progressivamente con l'aumentare delle registrazioni fino a raggiungere valori di AUC pari a 1 a 1000 sweeps. Tuttavia, tali miglioramenti non risultano sufficienti a superare le prestazioni della RF, che si mantiene al di sopra dello Chance Classifier già a partire da 100 sweeps. Questa differenza rispetto al caso CA può essere attribuita al fatto che il metodo APWF introduce una maggiore complessità negli indicatori, che sembra essere meglio catturata da un modello non lineare come la RF.

Infine, considerando gli indicatori ottenuti tramite SB, tutti i modelli si posizionano al di sopra dello Chance Classifier già a partire da 250 sweeps. In particolare, la LR mostra prestazioni elevate, superando RF e DT a partire da 500 sweeps. La RF evidenzia un miglioramento fino a 500 sweeps, seguito da una lieve riduzione delle prestazioni, mentre il DT mantiene un comportamento più stabile, con variazioni contenute al variare del numero di registrazioni.

Alla luce di queste osservazioni, l'estrazione di indicatori tramite CA, combinata con la LR, rappresenta un approccio particolarmente adatto a contesti di screening preliminare, ad esempio in ambito farmaceutico o clinico di primo livello. Ciò è dovuto alla semplicità del modello, alla rapidità di esecuzione e alle buone prestazioni ottenute anche con un numero limitato di registrazioni, consentendo una riduzione significativa dei tempi di acquisizione rispetto alle procedure tradizionali di analisi delle FFR.

Modelli più complessi, come la RF applicata a input ottenuti da APWF, risultano invece maggiormente indicati in ambito riabilitativo o di ricerca, dove è utile individuare pattern più complessi e relazioni non lineari nelle FFR. La loro maggiore capacità espressiva consente infatti di cogliere caratteristiche più articolate del segnale, che potrebbero risultare rilevanti per analisi più approfondite o personalizzate.

## 5.2 Dataset soggetti portatori IC

Successivamente all'addestramento e alla valutazione delle performance dei modelli nei soggetti normoudenti e patologici, gli stessi modelli sono stati addestrati e testati su soggetti con IC.

### *Rilevazione FFR Con LR*

Le tabelle successive evidenziano come le prestazioni del modello di LR nella rilevazione della presenza/assenza di FFR nei soggetti con IC dipendano dal metodo di post-elaborazione e dal numero di sweeps considerati.

<b>AUC</b>				
<b>SWEEPS</b>	<b>100</b>	<b>250</b>	<b>350</b>	<b>500</b>
<b>CA</b>	65%	74%	75%	95%
<b>APWF</b>	85%	78%	79%	97%
<b>SB</b>	85%	81%	73%	97%

*Tabella 14: AUC LR*

<b>ACCURATEZZA</b>				
<b>SWEEPS</b>	<b>100</b>	<b>250</b>	<b>350</b>	<b>500</b>
<b>CA</b>	64%	57%	64%	92%
<b>APWF</b>	78%	50%	71%	84%
<b>SB</b>	71%	71%	71%	92%

<b>SENSIBILITA'</b>				
<b>SWEEPS</b>	<b>100</b>	<b>250</b>	<b>350</b>	<b>500</b>
<b>CA</b>	62%	38%	63%	100%
<b>APWF</b>	63%	25%	75%	100%
<b>SB</b>	62%	63%	75%	100%

<b>SPECIFICITA'</b>				
<b>SWEEPS</b>	<b>100</b>	<b>250</b>	<b>350</b>	<b>500</b>
<b>CA</b>	67%	83%	67%	83%
<b>APWF</b>	100%	83%	67%	67%
<b>SB</b>	83%	83%	67%	83%

*Tabella 15,16,17: tabelle metriche di valutazione LR*

In particolare, la combinazione che risulta più stabile è quella basata sull'estrazione degli indicatori tramite SB. Infatti, LR applicata a tali indicatori mostra valori da discreti a buoni per

tutte le quantità di registrazioni considerate. Al contrario, i metodi CA e APWF presentano prestazioni più variabili, come evidenziato soprattutto dai valori di accuratezza e sensibilità, che risultano meno stabili al variare degli sweeps.

Confrontando questi risultati con quelli ottenuti nel dataset dei soggetti normoudenti, si osserva una generale riduzione delle performance. In particolare, l'AUC assume valori più contenuti a basse quantità di sweeps, pur mostrando un miglioramento significativo all'aumentare delle registrazioni, fino a raggiungere valori elevati (fino al 97%) a 500 sweeps per tutti i metodi di post-elaborazione.

Le metriche di accuratezza e sensibilità risultano generalmente inferiori rispetto al caso dei normoudenti, soprattutto per 100 e 250 sweeps, evidenziando una maggiore difficoltà del modello nel classificare correttamente i segnali in presenza di dati limitati.

Al contrario, la specificità si mantiene su valori relativamente stabili, compresi tra il 67% e il 100%, suggerendo una buona capacità del modello di identificare i casi negativi. Questo comportamento può essere attribuito a una distribuzione più equilibrata delle classi nel dataset, che rende la metrica meno sensibile a singoli errori di classificazione.

#### *Rilevazione FFR Con DT*

Dai valori dell'AUC riportati di seguito, sembra che il modello DT sia in grado di discriminare adeguatamente tra le due classi, con prestazioni che dipendono sia dal numero di sweeps sia dal metodo di post-elaborazione adottato.

<b>AUC</b>				
<b>SWEEPS</b>	<b>100</b>	<b>250</b>	<b>350</b>	<b>500</b>
<b>CA</b>	52%	63%	77%	83%
<b>APWF</b>	79%	50%	71%	91%
<b>SB</b>	83%	71%	77%	92%

*Tabella 19: tabelle metriche di valutazione DT*

In particolare, per i metodi CA e APWF, le capacità discriminative migliorano progressivamente all'aumentare del numero di registrazioni, raggiungendo valori soddisfacenti a partire da circa 350 sweeps.

Al contrario, utilizzando SB, il modello mostra prestazioni già buone a partire da 100 sweeps, suggerendo una maggiore efficacia di questo approccio nelle fasi iniziali di acquisizione.

<b>ACCURATEZZA</b>				
<b>SWEEPS</b>	<b>100</b>	<b>250</b>	<b>350</b>	<b>500</b>
<b>CA</b>	57%	64%	71%	84%
<b>APWF</b>	71%	50%	71%	92%
<b>SB</b>	86%	71%	71%	92%

<b>SENSIBILITA'</b>				
<b>SWEEPS</b>	<b>100</b>	<b>250</b>	<b>350</b>	<b>500</b>
<b>CA</b>	75%	75%	75%	100%
<b>APWF</b>	63%	62%	75%	100%
<b>SB</b>	88%	75%	75%	100%

<b>SPECIFICITA'</b>				
<b>SWEEPS</b>	<b>100</b>	<b>250</b>	<b>350</b>	<b>500</b>
<b>CA</b>	33%	50%	67%	67%
<b>APWF</b>	83%	33%	67%	83%
<b>SB</b>	83%	67%	67%	83%

*Tabella 19,20,21: tabelle metriche di valutazione DT*

Le metriche di accuratezza e sensibilità risultano coerenti con i valori di AUC, mostrando un incremento generale delle prestazioni all'aumentare delle sweeps. In particolare, la sensibilità assume valori elevati già a basse quantità di registrazioni, indicando una buona capacità del modello di identificare correttamente la presenza di FFR.

La specificità mostra invece una maggiore variabilità, soprattutto per un numero ridotto di sweeps, con valori inizialmente più bassi che tendono a migliorare con l'aumentare delle registrazioni.

Confrontando questi risultati con quelli ottenuti nel caso dei soggetti normoudenti, si osserva come le prestazioni risultino complessivamente migliori in diversi casi, in particolare per 100 e 500 sweeps, dove i valori delle metriche superano quelli precedentemente osservati. Questo comportamento potrebbe essere attribuito a una maggiore separabilità delle caratteristiche nei soggetti con IC oppure a una diversa distribuzione dei dati, che favorisce il processo decisionale basato su soglie tipico del DT.

Rilevazione FFR Con RF

AUC				
SWEEPS	100	250	350	500
CA	65%	73%	82%	93%
APWF	38%	75%	67%	81%
SB	50%	65%	74%	86%

Tabella 22: AUC RF

L'AUC mostra come il modello RF presenti prestazioni complessivamente buone a partire da 250, sebbene con una maggiore variabilità rispetto agli altri modelli, in funzione sia del numero di sweeps sia del metodo di post-elaborazione utilizzato.

In particolare, per il metodo CA si osserva un miglioramento progressivo delle capacità discriminative all'aumentare delle registrazioni, con valori già discreti a partire da 250 sweeps e prestazioni elevate a 500 sweeps. Analogamente, per SB il modello raggiunge risultati soddisfacenti a partire da circa 350 sweeps. Al contrario, per APWF le prestazioni risultano più instabili ma complessivamente sopra lo Chance Classifier già dopo 250 registrazioni.

ACCURATEZZA				
SWEEPS	100	250	350	500
CA	64%	64%	78%	92%
APWF	36%	71%	57%	77%
SB	43%	57%	64%	85%

SENSIBILITA'				
SWEEPS	100	250	350	500
CA	75%	75%	75%	100%
APWF	25%	87%	25%	100%
SB	75%	50%	75%	85%

SPECIFICITA'				
SWEEPS	100	250	350	500
CA	50%	50%	83%	83%
APWF	50%	50%	100%	50%
SB	0%	67%	50%	83%

Tabella 23,24,25: tabelle metriche di valutazione RF

Le metriche di accuratezza e sensibilità risultano coerenti con quanto osservato per l'AUC. In particolare, per CA si evidenzia un incremento progressivo delle prestazioni, mentre per APWF e SB si osserva una maggiore variabilità, soprattutto a basse quantità di sweeps.

In alcuni casi, la sensibilità assume valori molto bassi (ad esempio per APWF a 100 e 350 sweeps), indicando una difficoltà del modello nell'identificare correttamente i casi positivi in presenza di dati limitati.

La specificità mostra un andamento irregolare, con valori che variano significativamente tra i diversi metodi e numeri di registrazioni, arrivando anche a valori molto bassi (fino allo 0% nel caso di SB a 100 sweeps). Questo comportamento può essere attribuito sia alla scelta di una soglia non ottimale sia alla possibile presenza di alberi fortemente correlati tra loro, dovuta al numero ridotto di campioni disponibili per l'addestramento.

Nel complesso, l'applicazione della RF ai soggetti con impianto cocleare risulta meno affidabile rispetto al caso dei normoudenti, poiché le prestazioni appaiono generalmente inferiori e più instabili nella maggior parte delle configurazioni considerate.

#### *Modelli A Confronto*

Complessivamente, i modelli analizzati mostrano prestazioni discrete e leggermente inferiori nella rilevazione della presenza o assenza delle FFR nei soggetti con impianto cocleare rispetto alla loro applicazione in soggetti normoudenti o patologici, visibili anche nelle ROC a poste confronto per i diversi metodi di post elaborazione e rappresentate nella figura sottostante.

Questa riduzione delle performance può essere attribuita alle differenze intrinseche delle FFR nei soggetti con IC. L'impianto può introdurre artefatti elettrici significativi, che possono mascherare o alterare la risposta neurale, rendendo più difficile per il modello discriminare correttamente tra presenza e assenza di FFR. Inoltre, le caratteristiche fisiologiche del segnale nei soggetti con IC differiscono da quelle dei soggetti normoudenti, il che richiede ai modelli una maggiore capacità di adattamento per ottenere performance comparabili.

In generale, considerando nell'analisi anche le ROC illustrate in figura seguente, le combinazioni più performanti risultano essere SB e APWF applicate alla LR a partire da circa 350 sweeps, nonché CA applicato alla RF nello stesso intervallo di registrazioni. Queste configurazioni permettono di ottenere un buon compromesso tra capacità discriminativa e stabilità delle metriche.

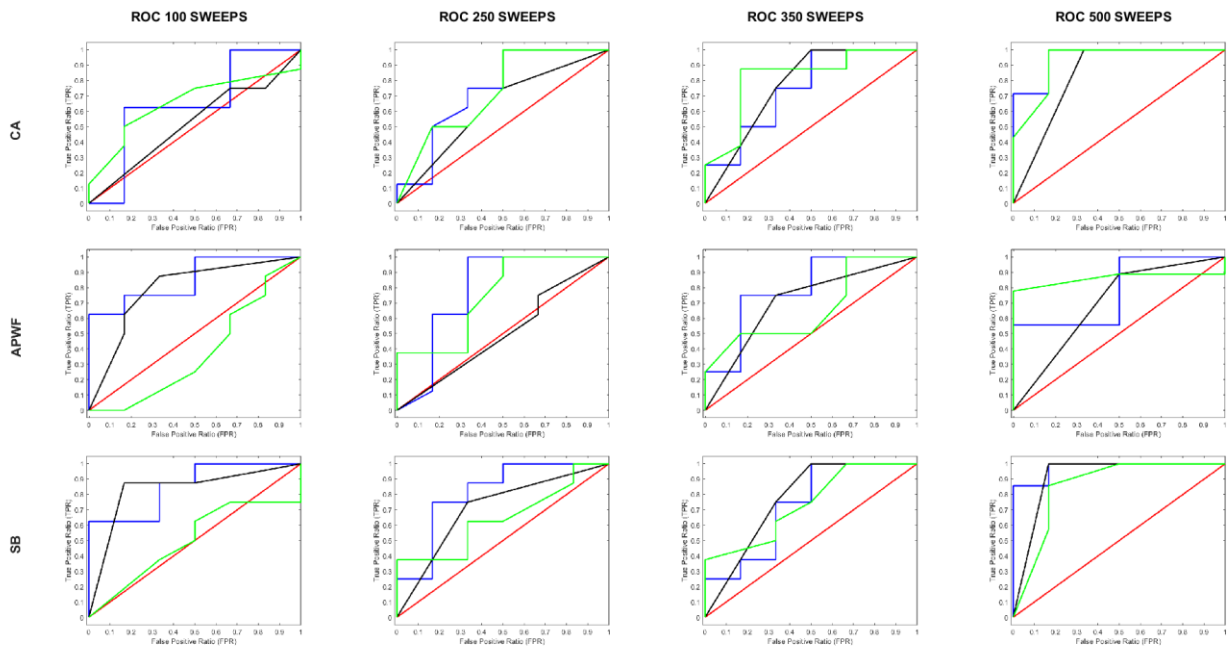


Figura 20: confronto tra ROC di Chance Classifier (rosso), LR (blu), DT (nero) e RF (verde) per numero crescente di registrazioni per i diversi metodi di post elaborazione

La scelta del modello più appropriato dipende fortemente dallo scopo applicativo. Nel contesto di questo studio, i modelli sono finalizzati al supporto del fitting protesico successivo all'inserimento dell'IC. In tale ambito, è fondamentale disporre di un metodo affidabile ma anche rapido, che consenta di ridurre i tempi di acquisizione mantenendo al contempo una buona accuratezza diagnostica.

In quest'ottica, la LR rappresenta una soluzione particolarmente vantaggiosa, grazie alla sua semplicità, interpretabilità e alle buone prestazioni ottenute anche con un numero relativamente contenuto di registrazioni. Modelli più complessi, come la RF, pur mostrando in questo contesto una maggiore variabilità e prestazioni talvolta inferiori, possono offrire un valore aggiunto nell'individuazione di pattern più articolati e relazioni non lineari, risultando potenzialmente utili in ambiti di analisi più approfondita o di ottimizzazione fine del fitting protesico. La sua variabilità potrebbe essere risolta ampliando il dataset.

# Capitolo 6

## Risultati sugli indici di capacità uditiva

La sola rilevazione della presenza o assenza delle FFR non è sufficiente a caratterizzare in modo completo la capacità di decodifica del segnale sonoro. Per questo motivo, si è scelto di affiancare ai metodi precedentemente descritti uno o più indici sintetici, in grado di fornire una valutazione più rappresentativa della capacità uditiva del soggetto.

### 6.1 Dataset normoudenti e patologici

Il primo test eseguito per accertare che la PCA fosse giustificata per il dataset in questione, è stato il Bartlett Test. I risultati ottenuti hanno evidenziato la presenza di correlazioni significative tra le variabili ( $p\text{-value} < 0.05$ ), giustificando quindi l'applicazione della PCA. In particolare, il  $p\text{-value}$  medio risulta pari a  $6.16 \cdot 10^{-8}$ , con deviazione standard pari a  $\pm 1.67 \cdot 10^{-7}$ .

A seguito di tale verifica, per numero crescente di sweep considerate e per ciascun metodo di post elaborazione, sono stati creati tre indici sintetici.

Il primo indice, denominato Global Index (GI), è stato calcolato come descritto al punto tre della sezione 4.3.3.1. Gli indici ICU1 e ICU2 sono stati invece ricavati rispettivamente dalla prima e dalla seconda componente principale della nuova rappresentazione dei dati (punto 2 sezione 4.3.3.1).

La scelta di introdurre ICU1 e ICU2 deriva dall'analisi dell'importanza delle variabili nelle prime due componenti principali. È stato infatti osservato che le proprietà con maggiore contributo si stabilizzano dopo circa 250 registrazioni e differiscono tra indici. Ovvero una proprietà che risulta importante per l'indice ICU1 non è importante per l'indice ICU2.

Questo risultato suggerisce che la capacità uditiva dei soggetti analizzati possa essere descritta in modo adeguato mediante una rappresentazione almeno bidimensionale.

Di seguito vengono riportati i valori medi, la deviazione standard per il dataset contenente dati estratti da soggetti sani e con lesione ischemica, suddivisi per classe di appartenenza.

*Global Index (GI)*

<b>GI PER CLASSE 0</b>				
<b>SWEEPS</b>	<b>100</b>	<b>250</b>	<b>500</b>	<b>1000</b>
<b>CA</b>	39,60 ±13,28	36.26 ± 14,19	33,40 ± 13,37	38,96± 11,57
<b>APWF</b>	32.57 ± 21,01	28.95 ± 24,32	38,54 ± 26,31	26,44 ± 18,54
<b>SB</b>	8,28 ± 28,35	29.47 ± 15,51	20,02 ± 8,67	19,86± 8,85

<b>GI PER CLASSE 1</b>				
<b>SWEEPS</b>	<b>100</b>	<b>250</b>	<b>500</b>	<b>1000</b>
<b>CA</b>	53,61 ±25,1	54,75± 23,81	53,26± 24,34	53,89± 25,34
<b>APWF</b>	54,86± 22,7	55,23 ± 24,32	63,01 ± 27,70	58,25 ± 25,26
<b>SB</b>	52,67± 26,12	47,77 ± 19,01	59,52± 29,45	50,29±25,90

*Tabella 26,27: tabelle valori medi e standard deviation per indice globale in soggetti sani e patologici*

Dall'analisi delle tabelle emerge che il GI calcolato per i soggetti con FFR presenti (classe 1) risulta mediamente più elevato rispetto a quello dei soggetti con FFR assenti (classe 0), in accordo con quanto atteso dal punto di vista fisiologico.

In particolare, l'applicazione della PCA ai dati preprocessati con il metodo APWF sembra fornire, in media, le prestazioni migliori. Questo comportamento può essere attribuito alla maggiore capacità di tale metodo di enfatizzare la risposta neurale del soggetto rispetto agli altri metodi di post-elaborazione considerati.

*ICUI*

<b>ICUI PER CLASSE 0</b>				
<b>SWEEPS</b>	<b>100</b>	<b>250</b>	<b>500</b>	<b>1000</b>
<b>CA</b>	28,73 ±7,21	18,41 ± 14,75	14,20 ± 14,58	16,50± 11,3
<b>APWF</b>	17,57 ± 12,73	29,48 ± 20,73	27,79 ± 19,96	18,22 ± 15,60
<b>SB</b>	68,42 ± 22,91	16,45 ± 13,02	15,07 ±12,46	15,53± 8,76

<b>ICUI PER CLASSE 1</b>				
<b>SWEEPS</b>	<b>100</b>	<b>250</b>	<b>500</b>	<b>1000</b>
<b>CA</b>	44,72 ±22,46	52,52± 24,38	48,02± 31	52,65± 25,55
<b>APWF</b>	38,26± 28,18	63,07 ± 21,90	57,28 ± 23,01	59,18 ± 22,24
<b>SB</b>	64,57± 22,09	51,46 ± 19,52	53,89± 28,27	56,02±25,88

*Tabella 28,29: tabelle valori medi e standard deviation per indice ICUI in soggetti sani e patologici*

Analogamente a quanto osservato per GI, i valori di ICU1 risultano mediamente più elevati per i soggetti appartenenti alla classe 1 rispetto alla classe 0.

Questo comportamento suggerisce che la prima componente principale (PC1), da cui è derivato ICU1, catturi una parte rilevante dell'informazione discriminante tra le due classi. In particolare, è plausibile che le proprietà che contribuiscono maggiormente a GI abbiano un peso significativo anche nella definizione di ICU1, in quanto PC1 rappresenta la direzione di massima varianza dei dati. Tale contributo risulta variabile in funzione del numero di registrazioni considerate: la varianza spiegata da PC1 è pari a circa il 30% per un numero ridotto di sweep, e aumenta progressivamente fino a raggiungere circa il 50% per un numero di registrazioni compreso tra 750 e 1000.

Questo andamento evidenzia come, all'aumentare del numero di registrazioni, la struttura dei dati diventi più stabile, permettendo alla PC1 di catturare una quota maggiore della variabilità complessiva.

Tale aspetto verrà approfondito nelle sezioni successive, mediante l'analisi delle correlazioni tra gli indici e le proprietà originali.

## ICU2

<b>ICU2 PER CLASSE 0</b>				
<b>SWEEPS</b>	<b>100</b>	<b>250</b>	<b>500</b>	<b>1000</b>
<b>CA</b>	75,3 ±20,19	71,79 ± 14,03	82,71 ± 5,03	78,30± 10,67
<b>APWF</b>	62 ± 26,17	72,66 ± 14,56	74,90 ± 11,7	81,60 ± 11,56
<b>SB</b>	17,26 ± 22,91	82,42 ± 10,05	68,38 ±6,91	82,07± 7,06

<b>ICU2 CLASSE 1</b>				
<b>SWEEPS</b>	<b>100</b>	<b>250</b>	<b>500</b>	<b>1000</b>
<b>CA</b>	65,67 ±20,21	56,99± 16,19	70,24± 23,88	63,91± 25,43
<b>APWF</b>	57,94± 22,42	60,81 ± 21,15	61,96 ± 25,18	67,73 ± 26,48
<b>SB</b>	41,32± 21,82	70,04 ± 26,56	65,56± 26,61	67,77±25,94

*Tabella 30,31: tabelle valori medi e standard deviation per indice ICU2 in soggetti sani e patologici*

Come si può osservare dalle tabelle precedenti, in questo caso l'indice ICU2 risulta mediamente più elevato per i soggetti appartenenti alla classe 0.

Questo comportamento può essere spiegato dalla correlazione negativa tra la seconda componente principale (PC2) e le proprietà che contribuiscono maggiormente alla sua

definizione. In particolare, valori più elevati di tali proprietà si traducono in valori più bassi della componente, e viceversa, determinando quindi un'inversione dell'andamento dell'indice rispetto a quanto osservato per ICU1 e per GI.

### *Robustezza e confronto degli indici*

A seguito del calcolo degli indici utilizzando l'intero dataset disponibile, è stata valutata la loro robustezza considerando una quantità ridotta di dati mediante 5-fold cv.

Gli indici ottenuti in queste condizioni sono stati quindi confrontati con quelli calcolati sull'intero dataset mediante il calcolo del coefficiente di correlazione.

I risultati ottenuti sono riportati nella tabella seguente.

<b>CORRELAZIONE INDICI CROSS VALIDATION VS INDICE TOTALE</b>												
<b>SWEEP</b>	<b>100</b>			<b>250</b>			<b>500</b>			<b>1000</b>		
<b>INDICE</b>	<b>ICU1</b>	<b>ICU2</b>	<b>GI</b>	<b>ICU1</b>	<b>ICU2</b>	<b>GI</b>	<b>ICU1</b>	<b>ICU2</b>	<b>GI</b>	<b>ICU1</b>	<b>ICU2</b>	<b>GI</b>
<b>CA</b>	0,55	0,45	0,85	0,96	0,56	0,75	0,97	0,93	0,93	0,99	0,21	0,93
<b>APWF</b>	0,33	0,43	0,49	0,98	0,17	0,86	0,97	0,82	0,95	0,99	0,78	0,98
<b>SB</b>	0,45	0,2	0,15	0,97	0,72	0,77	0,99	0,59	0,94	0,99	0,63	0,95

*Tabella 32: correlazione tra indici calcolati nell'intero dataset e indici calcolati con 5-fold cv*

Osservando attentamente i risultati, emerge che la robustezza degli indici aumenta al crescere del numero di sweep considerati. In particolare, per valori elevati di sweep (500 e 1000), si osservano correlazioni generalmente elevate (prossime a 1), indicando una buona stabilità degli indici anche in condizioni di riduzione dei dati.

Al contrario, per un numero ridotto di sweep (100), le correlazioni risultano più basse e variabili tra i diversi metodi di post-elaborazione, suggerendo una minore affidabilità degli indici in presenza di un numero limitato di registrazioni.

Tra i diversi indici, il GI mostra nel complesso le prestazioni più stabili, mantenendo valori di correlazione elevati anche per un numero intermedio di sweep. Questo risultato è coerente con la sua natura di combinazione pesata delle componenti principali, che consente di integrare l'informazione contenuta in più dimensioni e ridurre la variabilità.

Si osserva inoltre come il metodo APWF tenda a garantire una maggiore stabilità rispetto agli altri metodi, soprattutto per valori intermedi di sweep, mentre il metodo SB presenta una maggiore variabilità, in particolare per un numero ridotto di dati.

Al fine di comprendere quali proprietà contribuiscano maggiormente alla costruzione degli indici, è stata anche analizzata la correlazione tra ciascun indice e le proprietà originali. La

Figura 21 rappresenta tali correlazioni per PCA applicata successivamente a CA e APWF per valori crescenti di sweeps.

I risultati ottenuti confermano quanto già osservato dall'analisi dei valori medi degli indici. Ovvero:

- Le caratteristiche maggiormente correlate con ICU1(blu) risultano essere rilevanti anche per il GI (giallo), suggerendo che la PC1 rappresenti la principale fonte di informazione dell'indice complessivo.
- Le caratteristiche che influenzano ICU1 risultano invece poco correlate con ICU2 (rosso), e viceversa, indicando che le due componenti principali descrivono aspetti differenti e complementari della risposta uditiva.
- Valori elevati di ICU2 risultano associati a prestazioni peggiori del soggetto, in quanto la PC2 presenta una correlazione negativa con le proprietà fisiologicamente più rilevanti.

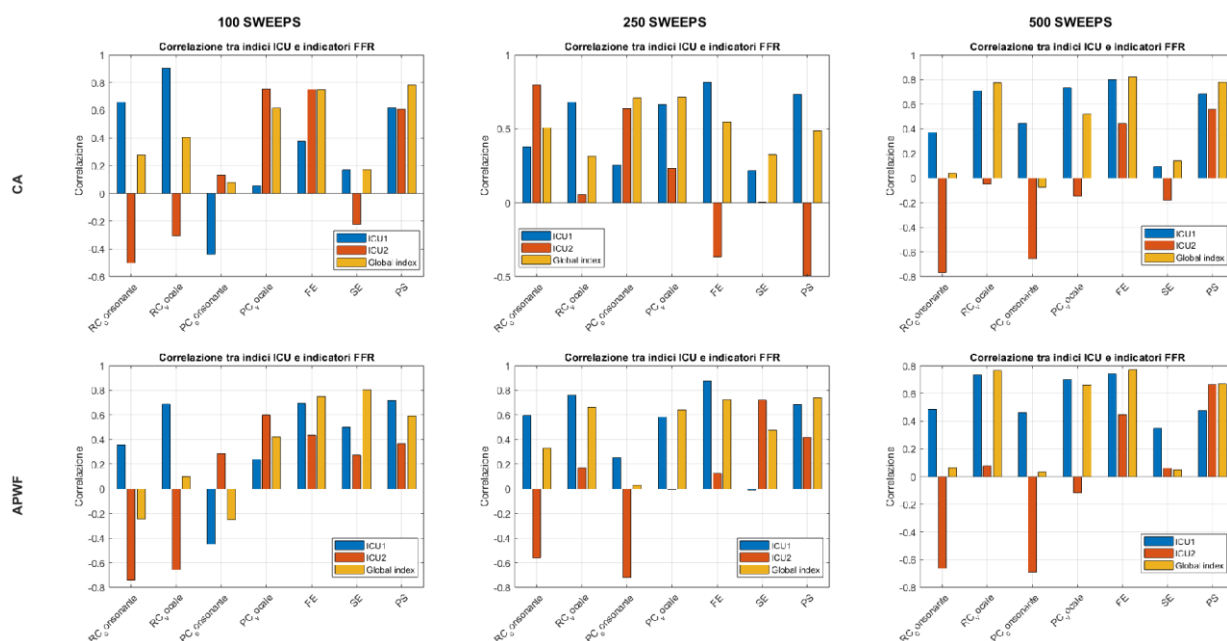


Figura 21: confronto correlazioni degli indici con le proprietà del dataset per PCA applicata successivamente a CA e APWF per valori crescenti di sweeps.

Inoltre, come rappresentato in figura precedente, analizzando l'evoluzione delle proprietà più influenti al variare del numero di registrazioni, si osserva che tali contributi tendono a stabilizzarsi a partire da circa 250 sweep.

In questa condizione, è possibile individuare una chiara separazione tra le caratteristiche associate ai diversi indici:

- ICU1: RC e PC della vocale, FE e PS
- ICU2: RC e PC della consonante

Questi risultati suggeriscono che i due indici possano essere interpretati in termini funzionali distinti ma complementari. In particolare, ICU1 sembra essere principalmente associato alla codifica delle componenti stazionarie del segnale (ad esempio la vocale) della periodicità, mentre ICU2 appare maggiormente legato alla codifica delle componenti transitorie del fonema, come la consonante.

Pertanto, ICU1 potrebbe essere utilizzato come indicatore della capacità di codifica tonale, mentre ICU2 potrebbe fornire informazioni sulla capacità di decodifica degli elementi transitori del segnale acustico.

#### *Test indici su soggetto con ictus pre e post DBS*

Per valutare ulteriormente l'efficacia degli indici sviluppati, è stata condotta un'analisi su un soggetto con lesione ischemica, esaminando le prestazioni prima e dopo l'esecuzione della stimolazione cerebrale profonda (DBS).

L'obiettivo di questo test è stato verificare se gli indici di capacità uditiva fossero in grado di rilevare variazioni funzionali legate al trattamento, fornendo così un'ulteriore conferma della loro sensibilità e utilità anche in contesti clinici particolarmente complessi.

I test sono stati eseguiti utilizzando 250 e 500 sweep per gli indicatori calcolati dai metodi CA e APWF. I valori medi sono riportati nelle tabelle seguenti.

	ICU1				ICU2				GLOBAL INDEX			
	PRE DBS		POST DBS		PRE DBS		POST DBS		PRE DBS		POST DBS	
	250	500	250	500	250	500	250	500	250	500	250	500
<b>SWP</b>												
<b>CA</b>	71,09± 5,4	91,23± 9,2	100	100	51,40± 13,9	84,51± 21,5	0	0	0,44± 0,44	1,05± 0,38	3,08± 0,22	10,33± 1,67
<b>APW</b>	74,63±	79,10±	100	100	64,76±	74,41±	0	0	0,86±	0,96±	0	11,08±
<b>F</b>	3,9	9,2			8,3	26,1			0,12	0,6		1,78

*Tabella 33: confronto valori medi e standard deviation per indici calcolati in soggetto con lesione ischemica pre e post DBS*

Dall'analisi dei risultati è emerso che la combinazione CA + PCA consente di rilevare miglioramenti funzionali già a partire da 250 registrazioni, con un incremento più evidente a

500 sweep. Il metodo APWF invece risulta più efficace nel rilevare i miglioramenti, ma solo a partire da 500 sweep, suggerendo una maggiore sensibilità alla quantità di dati disponibile.

Questi risultati indicano che gli indici sviluppati sono in grado di catturare variazioni significative nella capacità uditiva anche in soggetti con condizioni cliniche particolarmente complesse, sebbene la sensibilità dipenda dal metodo di post-elaborazione e dal numero di sweep considerati.

## 6.2 Dataset soggetti portatori di impianto cocleare

Gli indici di capacità uditiva sono stati calcolati anche per i soggetti portatori di IC, utilizzando la medesima procedura adottata per i soggetti normoudenti.

In questo contesto, come si può osservare dalle tabelle successive, le prestazioni degli indici risultano differenti rispetto a quelle ottenute nel dataset dei soggetti sani.

In particolare, il GI, sebbene presenti una buona robustezza già a partire da 100 registrazioni, risulta meno discriminante tra le due classi. Infatti, per tutti i metodi di post-elaborazione considerati, i valori dell'indice globale tendono a sovrapporsi tra le classi, rendendo difficile una separazione netta. Una differenza più evidente emerge solo a partire da circa 500 registrazioni, in particolare utilizzando i metodi APWF e SB.

SWEEP	GI PER CLASSE 0				GI PER CLASSE 1			
	100	250	350	500	100	250	350	500
CA	36 ± 21,82	37,13 ± 23,41	41,60 ± 21,50	41,53± 25,61	40,03 ± 19,32	39,24± 24,21	55,80± 22,54	36,5± 28,29
APWF	38,97 ± 21,46	42,66 ± 24,76	48,62 ± 25,70	46,25 ± 21,90	43,60± 25,12	50,52 ± 23,11	45,49 ± 22,60	55,26 ± 26,55
SB	35,40 ± 28,35	38,51 ± 22,57	37,77 ± 20,50	38,02± 20,10	39,18± 20,20	39,14 ± 24,86	53,84± 28,70	63,47± 31,30

Tabella 34: confronto valori medi e standard deviation per GI per numero crescente di sweep e per classe

Gli indici ICU1 e ICU2 invece mostrano una maggiore capacità discriminante già a partire da 100 registrazioni. Tuttavia, una stabilità più elevata si osserva per un numero di sweep superiore (circa 350), in accordo con quanto già evidenziato nell'analisi precedente.

SWEEP	ICU1 CLASSE 0				ICU1 CLASSE 1			
	100	250	350	500	100	250	350	500
CA	32,15 ± 26,08	29,47 ± 20,08	35,56 ± 23,77	37,25± 21,58	33,86 ± 23,70	40,33± 23,99	43,12± 27,67	46,59± 27,97
APWF	31,88 ± 22,33	32,48 ± 20,57	38,93 ± 22,83	24,65 ± 18,81	37,05± 26,31	35,39 ± 22,6	42,61 ± 23,50	37,99 ± 22,44
SB	29,58 ± 19,35	31,20 ± 20,31	34,24 ± 23,48	37,49± 21,81	34,16± 25,66	41,05 ± 24,73	44,92± 28,44	48,90± 28,53

Tabella 35: confronto valori medi e standard deviation per indici ICU1 per numero crescente di sweep e per classe

SWEEP	ICU2 CLASSE 0				ICU2 CLASSE 1			
	100	250	350	500	100	250	350	500
CA	67,35± 25,88	68,67 ± 18,76	28,84 ± 17,90	70,17± 25,61	75,45 ± 19,74	51,58± 16,27	43,78± 14,87	44,23± 28,29
APWF	72,68 ± 18,32	70,28 ± 22,65	81,94 ± 19,03	68,75 ± 17,18	66,82± 13,07	61,37 ± 22,99	67,25 ± 18,4	48,07 ± 26,28
SB	70,37 ± 20,17	71,34 ± 18,96	32,08 ± 18,48	37,27± 12,53	66,68± 16,02	53,82 ± 16,90	46,98± 14,94	57,13± 16,91

Tabella 36: confronto valori medi e standard deviation per indici ICU2 per numero crescente di sweep e per classe

CORRELAZIONE INDICI CROSS VALIDATION VS INDICE TOTALE												
SWEEP	100			250			350			500		
INDICI	ICU1	ICU2	GI	ICU1	ICU2	GI	ICU1	ICU2	GI	ICU1	ICU2	GI
CA	0,997	0,479	0,87	0,976	0,314	0,61	0,99	0,73	0,91	0,986	0,472	0,923
APWF	0,26	-0,014	0,18	0,98	0,87	0,93	0,95	0,28	0,7	0,9	0,36	0,86
SB	0,99	0,22	0,85	0,99	0,49	0,84	0,99	0,71	0,92	0,99	-0,6	0,65

Tabella 37: confronto correlazioni per indici GI, ICU1 e ICU2 per numero crescente di sweep

Analizzando la Figura 22, rappresentante la correlazione tra gli indici e le proprietà originali, si può osservare come nel caso dei soggetti portatori di IC il raggruppamento delle proprietà in ICU1 e ICU2 risulta più variabile rispetto a quanto osservato nei soggetti normoudenti.

In particolare, il raggruppamento delle caratteristiche — FE, PS, RC e PC della vocale per la definizione di ICU1, e RC e PC della consonante per ICU2 — si verifica in modo consistente solo per alcuni metodi di post-elaborazione e per specifici numeri di sweep.

Nel dettaglio, per i metodi SB e CA, tale organizzazione si osserva a partire da circa 250–350 sweep (per SB) e da 350 sweep in poi (per CA). Mentre per il metodo APWF si osserva una configurazione differente, in cui FE e PS risultano maggiormente associati a ICU2, indicando una diversa distribuzione dell'informazione tra le componenti principali.

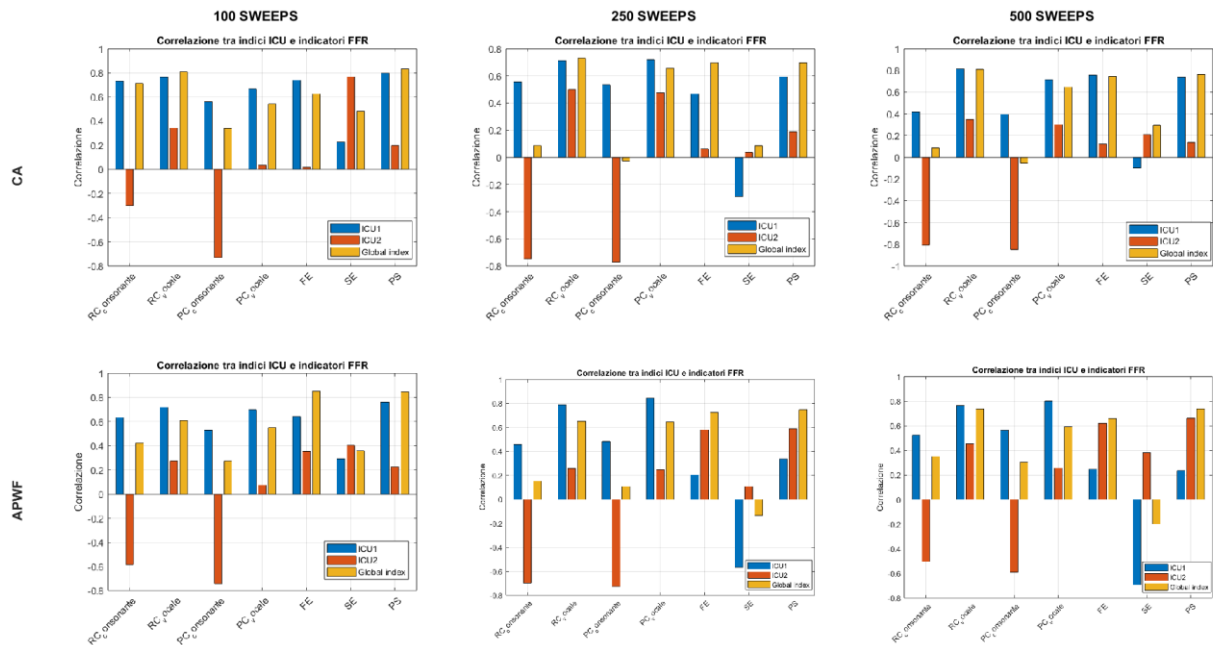


Figura 22: confronto correlazioni degli indici con le proprietà del dataset per PCA applicata successivamente a CA e APWF per valori crescenti di sweeps.

Queste differenze possono essere attribuite alle caratteristiche delle FFR nei soggetti con IC, che risultano significativamente diverse rispetto a quelle dei soggetti normoudenti. In particolare, le registrazioni possono essere influenzate dalla presenza di artefatti elettrici generati dall'impianto, che tendono a sovrapporsi al segnale neurale, rendendo più complessa l'estrazione delle informazioni rilevanti.

Questo comportamento è coerente con quanto osservato anche nei modelli di classificazione, nei quali la separabilità tra le classi risulta ridotta rispetto al caso dei soggetti sani.

Una possibile strategia per migliorare le prestazioni degli indici potrebbe consistere nell'applicazione di tecniche di filtraggio più avanzate, in grado di ridurre l'influenza degli artefatti prima della fase di elaborazione delle FFR per l'ottenimento degli indicatori diagnostici.

# Capitolo 7

## Conclusioni e sviluppi futuri

### 7.1 Bilancio del lavoro svolto

Le FFR rappresentano uno strumento di elevato interesse nell'ambito delle valutazioni otoneurologiche, in quanto consentono un'analisi oggettiva dei meccanismi di codifica del segnale sonoro a livello neurale.

Nel presente lavoro, l'estrazione delle caratteristiche spettrali e temporali delle FFR, integrata con tecniche di post-elaborazione, ha permesso di individuare indicatori affidabili, utilizzabili come input per modelli di machine learning finalizzati allo screening e alla valutazione della capacità uditiva, anche in presenza di dataset di dimensioni limitate.

In particolare, l'impiego degli indicatori estratti mediante CA, in combinazione con un classificatore di tipo LR, si è dimostrato efficace in contesti di screening preliminare, quali l'ambito clinico di primo livello e quello farmaceutico, raggiungendo performance superiori all'80% in tutte le metriche di valutazione già a partire da 100 registrazioni. Inoltre, tale approccio risulta adatto anche al fitting protesico in soggetti portatori di impianto cocleare, con prestazioni significativamente superiori a quelle di uno chance classifier già dopo 250 sweep (~ 80%).

Al contrario, l'utilizzo della RF su indicatori ottenuti tramite APWF risulta maggiormente indicato in contesti riabilitativi o di ricerca, nei quali è richiesta l'individuazione di pattern complessi e relazioni non lineari all'interno del segnale FFR. In tali scenari, il modello raggiunge prestazioni superiori all'80% in tutte le metriche considerate utilizzando un numero ridotto di sweep (pari a 100).

Un ulteriore contributo significativo è rappresentato dall'introduzione di indici sintetici di capacità uditiva basati su PCA, che consentono una rappresentazione compatta e interpretabile delle caratteristiche del segnale. Tali indici permettono non solo di rilevare la presenza della risposta, ma anche di descrivere le modalità di elaborazione del segnale a livello centrale, offrendo un potenziale supporto alla definizione di strategie riabilitative personalizzate. I miglioramenti osservati nel paziente con lesione ischemica, prima e dopo DBS, risultano apprezzabili già a partire da 500 registrazioni.

In conclusione, i risultati ottenuti confermano il potenziale dell'approccio proposto come strumento innovativo e promettente per una valutazione oggettiva della capacità uditiva,

aprendo prospettive rilevanti per future applicazioni cliniche e di ricerca nel campo dell'audiologia.

## **7.2 Margini di miglioramento per sviluppi futuri**

Sebbene i risultati ottenuti evidenzino come l'integrazione tra tecniche di elaborazione del segnale, modelli di machine learning e indici sintetici rappresenti un approccio promettente per una valutazione oggettiva e quantitativa della capacità uditiva, il presente lavoro presenta alcune limitazioni.

In primo luogo, la limitata disponibilità di dati può influenzare le prestazioni dei modelli più complessi, come la RF. Inoltre, nel caso dei soggetti portatori di IC, la presenza di artefatti elettrici può interferire con il segnale FFR, rendendo più complessa l'estrazione di informazioni rilevanti.

Un ulteriore limite è rappresentato dall'analisi condotta esclusivamente sul fonema /da/ e in condizioni di assenza di rumore, che non consente di esplorare completamente la variabilità delle risposte uditive in scenari più realistici.

Per quanto riguarda gli sviluppi futuri, risulta di particolare interesse estendere l'analisi a diversi fonemi, al fine di valutare la generalizzabilità degli indici proposti. L'introduzione di condizioni di rumore permetterebbe inoltre di simulare situazioni più vicine al contesto reale, rendendo l'analisi più rappresentativa delle condizioni quotidiane di ascolto. Parallelamente, un aumento della numerosità del dataset potrebbe migliorare la robustezza e l'affidabilità dei modelli, in particolare di quelli più complessi. Infine, lo sviluppo di tecniche di filtraggio avanzato potrebbe contribuire a ridurre l'impatto degli artefatti nei soggetti con impianto cocleare, migliorando la qualità del segnale analizzato.



# Appendice A

## A.1 L'impianto cocleare (CI): struttura e funzionamento

L'impianto cocleare (IC) rappresenta un dispositivo medico impiantabile impiegato nella riabilitazione di pazienti adulti e pediatrici affetti da significativa perdita neurosensoriale della funzione uditiva o da limitata capacità di discriminazione del linguaggio non adeguatamente compensata mediante apparecchi acustici tradizionali. Il suo scopo principale consiste nel trasdurre l'energia acustica in segnali elettrici idonei alla stimolazione diretta del nervo uditivo.

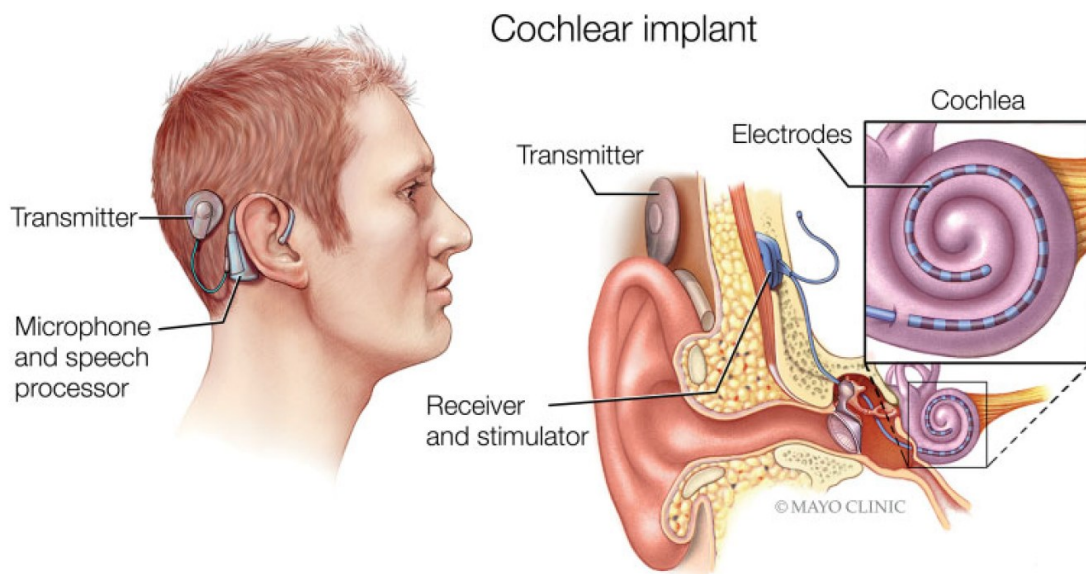


Figura 17: componenti dell'impianto cocleare.[42]

Come illustrato in Figura 16, il sistema IC è costituito da diverse componenti interne ed esterne.

La prima componente a entrare in contatto con il suono è il microfono, il quale registra i segnali acustici ambientali e li converte in segnali elettrici. Successivamente, tali segnali vengono inviati al processore del linguaggio, dove vengono analizzati e trasformati in un codice elettronico. Il codice ottenuto viene trasmesso al ricevitore-stimolatore interno attraverso un trasmettitore posizionato esternamente sul cranio del soggetto.

Una volta ricevuto il segnale, lo stimolatore lo trasforma in impulsi elettrici rapidi distribuiti in elettrodi multipli inseriti nella scala dei timpani della coclea. Gli elettrodi stimolano direttamente le fibre del nervo acustico, che trasmettono l'informazione al cervello, dove il segnale viene ulteriormente elaborato e interpretato come percezione sonora.

L'inserimento delle componenti interne avviene mediante un processo atraumatico, progettato per preservare la residua capacità uditiva a bassa frequenza (<1 kHz) nei pazienti con perdita uditiva parziale, e, al contempo, di beneficiare dei miglioramenti alle medie frequenze.[42]

È importante sottolineare che non tutti i soggetti sono adatti all'inserimento di IC. Per questo motivo è necessario effettuare valutazioni mediche, audiometriche e radiografiche adeguate al fine di definire idoneità del soggetto all'intervento.

Di seguito, viene riportata una tabella delle linee guida generali definite dalla Food and Drug Administration (FDA) per la selezione dei candidati all'impianto cocleare.

	<b>ADULTI</b>	<b>BAMBINI (2-17 ANNI)</b>	<b>BAMBINI (12-24 MESI)</b>
<b>Soglia uditiva</b>	Da moderata a profonda perdita uditiva neurosensoriale in entrambe le orecchie (>40dB)	Da severa a profonda perdita uditiva neurosensoriale (>70 dB)	Profonda perdita uditiva neurosensoriale (> 90dB)
<b>Riconoscimento parole</b>	Beneficio limitato dall'amplificazione binaurale definito da $\leq 50\%$ di riconoscimento delle frasi nell'orecchio da impiantare (oppure $\leq 40\%$ secondo i criteri CMS) e $\leq 60\%$ nell'orecchio controlaterale o in condizioni binaurali.	Beneficio limitato dall'amplificazione binaurale definito da punteggi di riconoscimento delle parole pari al 20–30%.	Beneficio limitato dalla prova di amplificazione binaurale basata sul MAIS.

*Tabella A.1: linee guida utilizzate per la candidatura all'impianto cocleare. Abbreviazioni utilizzate: CMS, Centers for Medicare and Medicaid Services; dB, decibel; MAIS, Meaningful Auditory Integration Scale. Modificata da [42].*

## **A.2 Valutazioni post-operatorie dell'IC**

Successivamente all'intervento chirurgico, gli impianti cocleari devono essere periodicamente regolati e calibrati al fine di garantire un'adeguata adattabilità dell'impianto al paziente e per definire il miglior processo riabilitativo.

Le valutazioni post-operatorie vengono eseguite mediante test oggettivi speech-based e non speech-based che non richiedono la cooperazione dei soggetti, caratteristica fondamentale per l'utilizzo in neonati e soggetti con ridotta capacità di risposta.[43]

In questo contesto, clinici e ricercatori sfruttano i segnali EEG per valutare la risposta neurale a stimoli acustici, focalizzandosi sull'evoluzione della plasticità celebrale durante la riabilitazione. Tra questi segnali, negli ultimi anni, ci si sta focalizzando particolarmente sulle FFR, che per i soggetti con IC sono definite come FFR evocate elettricamente (eFFR).

Tali eFFR sono notevolmente informative. Esse, infatti, come le FFR per i soggetti sani, sono in grado di definire la capacità di sincronizzazione di fase dei neuroni del tronco encefalico in

risposta alla stimolazione elettrica, fornendo indicazioni sui limiti biologici della codifica temporale nel sistema uditivo stimolato elettricamente e sulle differenze interindividuali tra utilizzatori di IC. In aggiunta, studi su modelli animali cronicamente impiantati, hanno confermato la capacità delle eFFR di essere una misura non invasiva affidabile della trasmissione temporale lungo il sistema uditivo, permettendo lo studio delle limitazioni della sincronizzazione di fase a diversi livelli del percorso uditivo.[44][45]

Il limite principale delle eFFR registrate in portatori di IC è la presenza di interferenze elettromagnetiche generate dall'impianto. Tali interferenze sovrastano notevolmente l'attività neurale e possono incidere sulla qualità del segnale registrato.

Per affrontare il problema, numerosi studi hanno applicato diverse tecniche di rimozione degli artefatti. Tra queste tecniche troviamo metodi basati su interpolazione lineare, template subtraction e altre procedure di elaborazione del segnale, che consentono di isolare la risposta neurale e ottenere registrazioni EEG più affidabili permettendo di utilizzare le FFR come misura oggettiva della capacità uditiva del soggetto in analisi. [43][44]

## A.3 Risultati completi

### Soggetti normoudenti classificazione con LR

AUC								
SWEEP	50	100	175	250	350	500	750	1000
CA	83%	100%	78%	94%	78%	89%	100%	100%
APWF	94%	53%	78%	89%	89%	78%	100%	100%
SB	89%	89%	94%	89%	83%	100%	100%	100%

Tabella A.3.1: versione estesa di Tabella 2: risultati AUC per modello LR

ACCURATEZZA								
SWEEP	50	100	175	250	350	500	750	1000
CA	73%	91%	73%	91%	91%	82%	82%	91%
APWF	91%	64%	64%	73%	82%	82%	82%	100%
SB	73%	91%	91%	73%	82%	82%	91%	100%

Tabella A.3.2: versione estesa di Tabella 3: risultati Accuratezza per modello LR

SENSIBILITA'								
SWEEP	50	100	175	250	350	500	750	1000
CA	67%	89%	78%	100%	100%	89%	78%	100%
APWF	89%	67%	67%	78%	89%	89%	78%	100%
SB	78%	89%	89%	78%	89%	78%	89%	100%

Tabella A.3.3: versione estesa di Tabella 4: risultati Sensibilità per modello LR

SPECIFICITA'								
SWEEP	50	100	175	250	350	500	750	1000
CA	100%	100%	50%	50%	50%	50%	100%	50%
APWF	100%	50%	50%	50%	50%	50%	100%	100%
SB	50%	100%	100%	50%	50%	100%	100%	100%

Tabella A.3.4: versione estesa di Tabella 5: risultati Specificità per modello LR

## Soggetti normoudenti classificazione con DT

AUC								
SWEEP	50	100	175	250	350	500	750	1000
CA	64%	64%	72%	72%	75%	89%	72%	69%
APWF	97%	61%	75%	75%	75%	69%	67%	94%
SB	92%	61%	75%	75%	67%	69%	72%	69%

Tabella A.3.5: versione estesa di Tabella 6: risultati AUC per modello DT

ACCURATEZZA								
SWEEP	50	100	175	250	350	500	750	1000
CA	64%	64%	82%	82%	91%	82%	82%	82%
APWF	91%	64%	91%	91%	91%	82%	82%	91%
SB	91%	64%	91%	91%	82%	82%	82%	82%

Tabella A.3.6: versione estesa di Tabella 7: risultati Accuratezza per modello DT

SENSIBILITA'								
SWEEP	50	100	175	250	350	500	750	1000
CA	67%	67%	89%	89%	100%	78%	89%	89%
APWF	89%	67%	100%	100%	100%	89%	89%	89%
SB	89%	67%	100%	100%	89%	89%	89%	89%

Tabella A.3.7: versione estesa di Tabella 8: risultati Sensibilità per modello DT

SPECIFICITA'								
SWEEP	50	100	175	250	350	500	750	1000
CA	50%	50%	50%	50%	50%	100%	50%	50%
APWF	100%	50%	50%	50%	50%	50%	50%	100%
SB	100%	50%	50%	50%	50%	50%	50%	50%

Tabella A.3.8: versione estesa di Tabella 9: risultati Specificità per modello DT

## Soggetti normoudenti classificazione con RF

AUC								
SWEEP	50	100	175	250	350	500	750	1000
CA	58%	75%	100%	56%	92%	69%	89%	92%

<b>APWF</b>	86%	92%	94%	89%	86%	86%	89%	100%
<b>SB</b>	58%	69%	92%	92%	86%	89%	89%	83%

Tabella A.3.9: versione estesa di Tabella 10: risultati AUC per modello RF

<b>ACCURATEZZA</b>								
<b>SWEEP</b>	<b>50</b>	<b>100</b>	<b>175</b>	<b>250</b>	<b>350</b>	<b>500</b>	<b>750</b>	<b>1000</b>
<b>CA</b>	45%	91%	73%	91%	73%	73%	91%	82%
<b>APWF</b>	64%	91%	91%	91%	91%	45%	91%	64%
<b>SB</b>	45%	73%	91%	73%	91%	91%	91%	73%

Tabella A.3.10: versione estesa di Tabella 11: risultati Accuratezza per modello RF

<b>SENSIBILITA'</b>								
<b>SWEEP</b>	<b>50</b>	<b>100</b>	<b>175</b>	<b>250</b>	<b>350</b>	<b>500</b>	<b>750</b>	<b>1000</b>
<b>CA</b>	33%	100%	67%	100%	67%	78%	100%	78%
<b>APWF</b>	56%	100%	89%	100%	100%	33%	100%	56%
<b>SB</b>	33%	78%	100%	67%	100%	100%	100%	67%

Tabella A.3.11: versione estesa di Tabella 12: risultati Sensibilità per modello RF

<b>SPECIFICITA'</b>								
<b>SWEEP</b>	<b>50</b>	<b>100</b>	<b>175</b>	<b>250</b>	<b>350</b>	<b>500</b>	<b>750</b>	<b>1000</b>
<b>CA</b>	100%	50%	100%	50%	100%	50%	50%	100%
<b>APWF</b>	100%	50%	100%	50%	50%	100%	50%	100%
<b>SB</b>	100%	50%	50%	100%	50%	50%	50%	100%

Tabella A.3.12: versione estesa di Tabella 13: risultati Specificità per modello RF

## Soggetti portatori di IC classificazione con LR

<b>AUC</b>							
<b>SWEEP</b>	<b>50</b>	<b>100</b>	<b>175</b>	<b>250</b>	<b>350</b>	<b>500</b>	<b>500</b>
<b>CA</b>	29%	65%	73%	74%	75%	95%	
<b>APWF</b>	48%	85%	61%	78%	79%	97%	
<b>SB</b>	33%	85%	71%	81%	73%	97%	

Tabella A.3.13: versione estesa di Tabella 14: risultati AUC per modello LR

<b>ACCURATEZZA</b>							
<b>SWEEP</b>	<b>50</b>	<b>100</b>	<b>175</b>	<b>250</b>	<b>350</b>	<b>500</b>	<b>500</b>
<b>CA</b>	36%	64%	57%	57%	64%	92%	
<b>APWF</b>	50%	78%	57%	50%	71%	84%	
<b>SB</b>	50%	71%	43%	71%	71%	92%	

Tabella A.3.14: versione estesa di Tabella 15: risultati Accuratezza per modello LR

<b>SENSIBILITA'</b>						
<b>SWEEP</b>	<b>50</b>	<b>100</b>	<b>175</b>	<b>250</b>	<b>350</b>	<b>500</b>
<b>CA</b>	50%	62%	38%	38%	63%	100%
<b>APWF</b>	63%	63%	50%	25%	75%	100%
<b>SB</b>	75%	62%	25%	63%	75%	100%

*Tabella A.3.15: versione estesa di Tabella 16: risultati Sensibilità per modello LR*

<b>SPECIFICITA'</b>						
<b>SWEEP</b>	<b>50</b>	<b>100</b>	<b>175</b>	<b>250</b>	<b>350</b>	<b>500</b>
<b>CA</b>	17%	67%	83%	83%	67%	83%
<b>APWF</b>	33%	100%	67%	83%	67%	67%
<b>SB</b>	17%	83%	67%	83%	67%	83%

*Tabella A.3.16: versione estesa di Tabella 17: risultati Specificità per modello LR*

## Soggetti portatori di IC classificazione con DT

<b>AUC</b>						
<b>SWEEP</b>	<b>50</b>	<b>100</b>	<b>175</b>	<b>250</b>	<b>350</b>	<b>500</b>
<b>CA</b>	43%	52%	67%	63%	77%	83%
<b>APWF</b>	63%	79%	68%	50%	71%	91%
<b>SB</b>	43%	83%	50%	71%	77%	92%

*Tabella A.3.17: versione estesa di Tabella 18: risultati AUC per modello DT*

<b>ACCURATEZZA</b>						
<b>SWEEP</b>	<b>50</b>	<b>100</b>	<b>175</b>	<b>250</b>	<b>350</b>	<b>500</b>
<b>CA</b>	43%	57%	64%	64%	71%	84%
<b>APWF</b>	64%	71%	64%	50%	71%	92%
<b>SB</b>	50%	86%	57%	71%	71%	92%

*Tabella A.3.18: versione estesa di Tabella 19: risultati Accuratezza per modello DT*

<b>SENSIBILITA'</b>						
<b>SWEEP</b>	<b>50</b>	<b>100</b>	<b>175</b>	<b>250</b>	<b>350</b>	<b>500</b>
<b>CA</b>	62%	75%	62%	75%	75%	100%
<b>APWF</b>	75%	63%	62%	62%	75%	100%
<b>SB</b>	50%	88%	75%	75%	75%	100%

*Tabella A.3.19: versione estesa di Tabella 20: risultati Sensibilità per modello DT*

<b>SPECIFICITA'</b>						
<b>SWEEP</b>	<b>50</b>	<b>100</b>	<b>175</b>	<b>250</b>	<b>350</b>	<b>500</b>
<b>CA</b>	17%	33%	67%	50%	67%	67%
<b>APWF</b>	50%	83%	67%	33%	67%	83%
<b>SB</b>	50%	83%	33%	67%	67%	83%

Tabella A.3.20: versione estesa di Tabella 21: risultati Specificità per modello DT

## Soggetti portatori di IC classificazione con RF

AUC						
SWEEPS	50	100	175	250	350	500
CA	65%	65%	67%	73%	82%	93%
APWF	51%	38%	54%	75%	67%	81%
SB	51%	50%	57%	65%	74%	86%

Tabella A.3.21: versione estesa di Tabella 22: risultati AUC per modello RF

ACCURATEZZA						
SWEEP	50	100	175	250	350	500
CA	57%	64%	57%	64%	78%	92%
APWF	50%	36%	57%	71%	57%	77%
SB	50%	43%	57%	57%	64%	85%

Tabella A.3.22: versione estesa di Tabella 23: risultati Accuratezza per modello RF

SENSIBILITA'						
SWEEPS	50	100	175	250	350	500
CA	38%	75%	38%	75%	75%	100%
APWF	13%	25%	87%	87%	25%	100%
SB	13%	75%	38%	50%	75%	85%

Tabella A.3.23: versione estesa di Tabella 24: risultati Sensibilità per modello RF

SPECIFICITA'						
SWEEPS	50	100	175	250	350	500
CA	83%	50%	83%	50%	83%	83%
APWF	100%	50%	17%	50%	100%	50%
SB	100%	0%	83%	67%	50%	83%

Tabella A.3.24: versione estesa di Tabella 25: risultati Specificità per modello RF

## Soggetti normoudenti indici capacità uditiva

GLOBAL INDEX PER CLASSE 0								
SWEEP	50	100	175	250	350	500	750	1000
CA	30,23± 22,60	39,60 ± 13,28	37,13± 12,96	36,26 ± 14,19	30,22± 13,72	33,40 ± 13,37	30,41± 12,04	38,96± 11,57
APWF	23,70± 20,98	32,57 ± 21,01	24,70± 17,55	28,95 ± 24,32	40,64± 26,39	38,54 ± 26,31	22,19± 13,88	26,44 ± 18,54
SB	24,70± 6,64	8,28 ± 28,35	28,40± 11,82	29,47 ± 15,51	32,96± 7,20	20,02 ± 8,67	9,41± 8,43	19,86± 8,85

Tabella A.3.25: versione estesa di Tabella 26: risultati medi e standard deviation GI per classe 0

GLOBAL INDEX PER CLASSE 1								
SWEEP	50	100	175	250	350	500	750	1000
CA	42,78± 23,01	53,61 ± 25,1	56,06± 24,37	54,75± 23,81	52,48± 28,43	53,26± 24,34	54,85± 28,43	53,89± 25,34
APWF	55,07± 22,78	54,86± 22,7	52,71± 20,36	55,23 ± 24,32	54,37± 30,56	63,01 ± 27,70	54,37± 30,56	58,25 ± 25,26
SB	47,58± 21,13	52,67± 26,12	48,96± 24,89	47,77 ± 19,01	49,12± 25,45	59,52± 29,45	49,12± 25,45	50,29± 25,90

Tabella A.3.26: versione estesa di Tabella 27: risultati medi e standard deviation GI per classe 1

ICU1 PER CLASSE 0								
SWEEP	50	100	175	250	350	500	750	1000
CA	18,77± 14,26	28,73 ± 7,21	18,43± 11,05	18,41 ± 14,75	19± 16,07	14,20 ± 14,58	14,72± 12,48	16,50± 11,3
APWF	15,61± 13,66	17,57 ± 12,73	26,11± 16,33	29,48 ± 20,73	30,63± 24,41	27,79 ± 19,96	19,17± 18,45	18,22 ± 15,60
SB	11,53± 11,37	68,42 ± 22,91	13,94± 8,79	16,45 ± 13,02	21,68± 14,24	15,07 ± 12,46	18,99± 9,92	15,53± 8,76

Tabella A.3.27: versione estesa di Tabella 28: risultati medi e standard deviation indice ICU1 per classe 0

ICU1 PER CLASSE 1								
SWEEP	50	100	175	250	350	500	750	1000
CA	40,51± 21,73	44,72 ± 22,46	46,74± 24,90	52,52± 24,38	50,51± 26,99	48,02± 31	52,56± 27,63	52,65± 25,55
APWF	40,09± 21,50	38,26± 28,18	49,74± 23,17	63,07 ± 21,90	62,33± 24,37	57,28 ± 23,01	58,87± 23,32	59,18 ± 22,24
SB	37,30± 23,02	64,57± 22,09	41,94± 22,56	51,46 ± 19,52	54,89± 26,82	53,89± 28,27	55,06± 25,25	56,02± 25,88

Tabella A.3.28: versione estesa di Tabella 29: risultati medi e standard deviation indice ICU1 per classe 1

ICU2 PER CLASSE 0								
SWEEP	50	100	175	250	350	500	750	1000
CA	65,91± 26,73	75,3 ± 20,19	41,57± 18,17	71,79 ± 14,03	79,15± 5,12	82,71 ± 5,03	76,65± 12,17	78,30± 10,67
APWF	47,82± 33,25	62 ± 26,17	53,11± 14,07	72,66 ± 14,56	77,15± 8,07	74,90 ± 11,7	24,15± 12,90	81,60 ± 11,56
SB	55,05± 22,88	17,26 ± 22,91	70,90± 22,55	82,42 ± 10,05	75,77± 5,77	68,38 ± 6,91	23,95± 8,99	82,07± 7,06

Tabella A.3.29: versione estesa di Tabella 30: risultati medi e standard deviation indice ICU2 per classe 0

ICU2 CLASSE 1								
SWEEP	50	100	175	250	350	500	750	1000
CA	64,18± 19,83	65,67 ± 20,21	50,54± 26,25	56,99± 16,19	62,37± 27,39	70,24± 23,88	58,16± 26,47	63,91± 25,43
APWF	61,83± 24,66	57,94± 22,42	56,61± 22,76	60,81 ± 21,15	62,97± 26,88	61,96 ± 25,18	37,60± 27,07	67,73 ± 26,48

<b>SB</b>	63,02± 23,06	41,32± 21,82	54,63± 19,24	70,04 ± 26,56	62,40± 26,50	65,56± 26,61	31,96± 27,05	67,77± 25,94
-----------	-----------------	-----------------	-----------------	------------------	-----------------	-----------------	-----------------	-----------------

Tabella A.3.30: versione estesa di Tabella 31: risultati medi e standard deviation indice ICU2 per classe 1

<b>CORRELAZIONE INDICE CROSS VALIDATION VS GI TOT</b>								
<b>SWEEP</b>	<b>50</b>	<b>100</b>	<b>175</b>	<b>250</b>	<b>350</b>	<b>500</b>	<b>750</b>	<b>1000</b>
<b>CA</b>	0,71	0,85	0,77	0,75	0,79	0,93	0,79	0,93
<b>APWF</b>	0,48	0,49	0,42	0,86	0,73	0,95	0,69	0,98
<b>SB</b>	0,89	0,15	0,04	0,77	0,81	0,94	0,87	0,95

Tabella A.3.31: versione estesa valori di correlazione tra indice costruito con dataset totale vs indice costruito con 5-fold cv

<b>CORRELAZIONE ICU1 CROSS VALIDATION VS ICU1 TOT</b>								
<b>SWEEP</b>	<b>50</b>	<b>100</b>	<b>175</b>	<b>250</b>	<b>350</b>	<b>500</b>	<b>750</b>	<b>1000</b>
<b>CA</b>	0,96	0,55	0,98	0,96	0,98	0,97	0,99	0,99
<b>APWF</b>	0,47	0,33	0,58	0,98	0,97	0,97	0,98	0,99
<b>SB</b>	0,98	0,45	0,12	0,97	0,98	0,99	0,99	0,99

Tabella A.3.32: versione estesa valori di correlazione tra indice ICU1 costruito con dataset totale vs indice costruito con 5-fold cv

<b>CORRELAZIONE ICU2 CROSS VALIDATION VS ICU2 TOT</b>								
<b>SWEEP</b>	<b>50</b>	<b>100</b>	<b>175</b>	<b>250</b>	<b>350</b>	<b>500</b>	<b>750</b>	<b>1000</b>
<b>CA</b>	0,25	0,45	0,88	0,56	0,77	0,93	-0,58	0,21
<b>APWF</b>	0,96	0,43	0,2	0,17	-0,13	0,82	-0,85	0,78
<b>SB</b>	0,99	0,2	0,19	0,72	0,6	0,59	-0,37	0,63

Tabella A.3.33: versione estesa valori di correlazione tra indice ICU2 costruito con dataset totale vs indice costruito con 5-fold cv

## Soggetti con IC indici capacità uditiva

<b>GLOBAL INDEX PER CLASSE 0</b>						
<b>SWEEP</b>	<b>50</b>	<b>100</b>	<b>175</b>	<b>250</b>	<b>350</b>	<b>500</b>
<b>CA</b>	30,8± 18,45	36 ± 21,82	31,51± 22,34	37,13 ± 23,41	41,60 ± 21,50	41,53± 25,61
<b>APWF</b>	66,36± 19,74	38,97 ± 21,46	43,44±2 4,79	42,66 ± 24,76	48,62 ± 25,70	46,25 ± 21,90
<b>SB</b>	60,66± 18,86	35,40 ± 28,35	36,06± 22,13	38,51 ± 22,57	37,77 ± 20,50	38,02± 20,10

Tabella A.3.34: versione estesa valori medi e standard deviation del GI per classe 0 IC

<b>GLOBAL INDEX PER CLASSE 1</b>						
<b>SWEEP</b>	<b>50</b>	<b>100</b>	<b>175</b>	<b>250</b>	<b>350</b>	<b>500</b>
<b>CA</b>	34,5± 23,27	40,03 ± 19,32	29,05± 22,74	39,24± 24,21	55,80± 22,54	36,5± 28,29
<b>APWF</b>	57,98± 20,36	43,60± 25,12	46,84± 24,63	50,52 ± 23,11	45,49 ± 22,60	55,26 ± 26,55

<b>SB</b>	58,06±	39,18±	36,59±	39,14 ±	53,84±	63,47±
	23,28	20,20	25,57	24,86	28,70	31,30

Tabella A.3.35: versione estesa valori medi e standard deviation del GI per classe 1 IC

<b>ICU1 CLASSE 0</b>						
<b>SWEEP</b>	<b>50</b>	<b>100</b>	<b>175</b>	<b>250</b>	<b>350</b>	<b>500</b>
<b>CA</b>	36,16±	32,15 ±	35,42±	29,47 ±	35,56 ±	37,25±
	22,65	26,08	19,85	20,08	23,77	21,58
<b>APWF</b>	65,12±	31,88 ±	37,98±	32,48 ±	38,93 ±	24,65 ±
	24,93	22,33	22,23	20,57	22,83	18,81
<b>SB</b>	71,03±	29,58 ±	36,78±	31,20 ±	34,24 ±	37,49±
	21,05	19,35	18,77	20,31	23,48	21,81

Tabella A.3.36: versione estesa valori medi e standard deviation dell'indice ICU1 per classe 0 IC

<b>ICU1 CLASSE 1</b>						
<b>SWEEP</b>	<b>50</b>	<b>100</b>	<b>175</b>	<b>250</b>	<b>350</b>	<b>500</b>
<b>CA</b>	38,23±	33,86 ±	45,68±	40,33±	43,12±	46,59±
	33,75	23,70	25,48	23,99	27,67	27,97
<b>APWF</b>	60,16±	37,05±	41,29±	35,39 ±	42,61 ±	37,99 ±
	22,41	26,31	22,40	22,6	23,50	22,44
<b>SB</b>	70,43±	34,16±	46,72±	41,05 ±	44,92±	48,90±
	19,82	25,66	24,25	24,73	28,44	28,53

Tabella A.3.37: versione estesa valori medi e standard deviation dell'indice ICU1 per classe 1 IC

<b>ICU2 CLASSE 0</b>						
<b>SWEEP</b>	<b>50</b>	<b>100</b>	<b>175</b>	<b>250</b>	<b>350</b>	<b>500</b>
<b>CA</b>	60,35±	67,35±	66,23±	68,67 ±	28,84 ±	70,17±
	21,31	25,88	17	18,76	17,90	25,61
<b>APWF</b>	34,80±	72,68 ±	76,22±	70,28 ±	81,94 ±	68,75 ±
	19,41	18,32	19,32	22,65	19,03	17,18
<b>SB</b>	47,40±	70,37 ±	65,99±	71,34 ±	32,08 ±	37,27±
	17,97	20,17	17,14	18,96	18,48	12,53

Tabella A.3.38: versione estesa valori medi e standard deviation dell'indice ICU2 per classe 0 IC

<b>ICU2 CLASSE 1</b>						
<b>SWEEP</b>	<b>50</b>	<b>100</b>	<b>175</b>	<b>250</b>	<b>350</b>	<b>500</b>
<b>CA</b>	62,61±	75,45 ±	51,43±	51,58±	43,78±	44,23±
	16,96	19,74	13,5	16,27	14,87	28,29
<b>APWF</b>	29,81±	66,82±	64,47±	61,37 ±	67,25 ±	48,07 ±
	15,51	13,07	22,15	22,99	18,4	26,28
<b>SB</b>	46,05±	66,68±	51,01±	53,82 ±	46,98±	57,13±
	16,92	16,02	13,34	16,90	14,94	16,91

Tabella A.3.39: versione estesa valori medi e standard deviation dell'indice ICU2 per classe 1 IC

<b>CORRELAZIONE INDICE GLOBALE CROSS VALIDATION VS INDICE GLOBALE</b>						
<b>SWEEPS</b>	<b>50</b>	<b>100</b>	<b>175</b>	<b>250</b>	<b>350</b>	<b>500</b>
<b>CA</b>	0,79	0,87	0,6	0,61	0,91	0,923

<b>APWF</b>	-0,88	0,18	0,8	0,93	0,7	0,86
<b>SB</b>	-0,76	0,85	0,71	0,84	0,92	0,65

*Tabella A.3.40: versione estesa valori di correlazione del GI costruito con intero dataset vs indice costruito con 5-fold cv*

<b>CORRELAZIONE ICU1 CROSS VALIDATION VS ICU1 TOT IC</b>						
<b>SWEEP</b>	<b>50</b>	<b>100</b>	<b>175</b>	<b>250</b>	<b>350</b>	<b>500</b>
<b>CA</b>	0,98	0,997	0,98	0,976	0,99	0,986
<b>APWF</b>	-0,96	0,26	0,95	0,98	0,95	0,9
<b>SB</b>	-0,98	0,99	0,99	0,99	0,99	0,99

*Tabella A.3.41: versione estesa valori di correlazione dell'indice ICU1 costruito con intero dataset vs indice costruito con 5-fold cv*

<b>CORRELAZIONE ICU2 CROSS VALIDATION VS ICU2 TOT IC</b>						
<b>SWEEP</b>	<b>50</b>	<b>100</b>	<b>175</b>	<b>250</b>	<b>350</b>	<b>500</b>
<b>CA</b>	0,95	0,479	-0,23	0,314	0,73	0,472
<b>APWF</b>	-0,86	-0,014	0,55	0,87	0,28	0,36
<b>SB</b>	-0,3	0,22	0,37	0,49	0,71	-0,6

*Tabella A.3.42: versione estesa valori di correlazione dell'indice ICU2 costruito con intero dataset vs indice costruito con 5-fold Cv*

# Appendice B

## Codici Matlab

### Algoritmi interni per modelli di rilevazione FFR

```
function [beta] = Newton_Rapson(data,class,lambda,maxiter,toll)
%funzione per ricerca dei beta ottimi mediante il criterio di massima
% verosomiglianza implementato con Newton-Rapshon;

%input:
%data = dati su cui ricercare beta ottimi
%class = classe di appartenenza
%lambda = valore da utilizzare nella Ridge Regression
%maxiter = massimo numero di iterazioni
%toll = tolleranza d'errore

%output:
%beta = beta ottimi

M = size(data,2);
beta_old = zeros(M,1);
iter = 0;
error = Inf;
loglik_old = -Inf;

while error > toll && iter < maxiter
    iter = iter + 1;
    pi = 1./(1+exp(-data*beta_old));
    Win = pi.*(1-pi);
    Lambda = lambda*eye(M);
    Lambda(1,1) = 0;
    penality = lambda*beta_old;
    penality(1) = 0;
    H = data'*(data.*Win) + Lambda;
    g = data'*(class-pi) - penality;
    betanew = beta_old + H\g;

    % log likelihood con beta nuovi
    pi_new = 1./(1+exp(-data*betanew));
    eps = 1e-12;
    pi_new = min(max(pi_new,eps),1-eps);
    loglik = sum(class.*log(pi_new) + (1-class).*log(1-pi_new)) ...
        - (lambda/2)*sum(betanew(2:end).^2);

    error = abs(loglik - loglik_old);
    loglik_old = loglik;

    beta_old = betanew;
end
beta = betanew;
end
```

*Codice 1: funzione Newton\_Rapson.mat per la ricerca dei beta ottimi mediante MLC*

```

function [nodo,iter] = DecisionTree(data,class,iter,maxiter,RF)
%funzione per la creazione di un albero decisionale
%input:
%data = dati di training
%class = classe di appartenenza
%iter = numero di iterazioi
%maxiter = massimo numero di iterazioni
%toll = tolleranza d'errore

%output:
%nodo = nodo formato
%iter = quantità di iterazioni utilizzate

    iter = iter+1;
    nodo = struct();
    class1 = mean(class);
    class0 = 1-class1;

    if class1 == 1
        nodo.foglia = true;
        nodo.p = 1;
        nodo.class = 1;
        return
    elseif class0 == 1
        nodo.foglia = true;
        nodo.p = 0;
        nodo.class = 0;
        return
    end

    %Utile per evitare potatura
    if size(data,1) <= 5
        nodo.foglia = true;
        nodo.p = mean(class);
        nodo.class = round(mean(class));
        return
    elseif iter > maxiter
        nodo.foglia = true;
        nodo.p = mean(class);
        nodo.class = round(mean(class));
        return
    end

    if RF == 1
        d = size(data,2);
        mtry = round(sqrt(d));

        splitting_iesimi = randperm(d,mtry);
        data_sp = data(:,splitting_iesimi);
        bestsc = splitting_criterion(data_sp, class);
        if bestsc(1)<=0 || isnan(bestsc(1)) || bestsc(1)>
            length(splitting_iesimi)

            nodo.foglia = true;
            nodo.p = mean(class);
            nodo.class = round(mean(class));
            return
        else

```

```

        nodo.foglia = false;
        nodo.proprieta = bestsc(1);
    end
else
    bestsc = splitting_criterion(data, class);
    if bestsc(1)<=0 || isnan(bestsc(1))
        nodo.foglia = true;
        nodo.p = mean(class);
        nodo.class = round(mean(class));
        return
    else
        nodo.foglia = false;
        nodo.proprieta = bestsc(1);
    end
end

nodo.soglia = bestsc(2);
trnsplit = data(:,nodo.proprieta);
idsplit = find(trnsplit>= nodo.soglia);

%creazione nodo destro
nododx = data(idsplit,:);
classdx = class(idsplit);

%creazione nodo sinistro
nodosx =data;
nodosx(idsplit,:) = [];
classx = class;
classx(idsplit)= [];

if isempty(classdx) || isempty(classx)
    nodo.foglia = true;
    nodo.p = mean(class);
    nodo.class = round(mean(class));
    return
end

[nodo.destro, ~]= DecisionTree(nododx,classdx,iter,maxiter,RF);
[nodo.sinistro,~] = DecisionTree(nodosx,classx,iter,maxiter,RF);

end

```

**Codice 2:** *funzione DecisionTree.mat utilizzata per formare ricorsivamente gli alberi decisionali sia in DT che in RF*

```

function [deltap] = Gini_index(data,class,splitting_criterion)
%funzione che calcola la variazione di impurità tra il nodo padre e i nodi
%figli creati con il criterio di splitting dato in ingresso

nododx = class(data >= splitting_criterion);
nodosx =class(data < splitting_criterion);
classpadre = mean(class);
Imppadre = (1-classpadre)*classpadre;

if isempty(nododx)
    impnododx = 0;
    pnododx = 0;
else

```

```

        classnododx = mean(nododx);
        impnododx = (1-classnododx)*classnododx;
        pnododx = numel(nododx)/numel(class);
    end

    if isempty(nodosx)
        impnodosx = 0;
        pnodosx = 0;
    else
        classnodosx = mean(nodosx);
        impnodosx = (1-classnodosx)*classnodosx;
        pnodosx = numel(nodosx)/numel(class);
    end
    deltap = Imppadre - (pnodosx*impnodosx + pnododx*impnododx);

end

```

**Codice 3:** *funzione GiniIndex.mat per trovare la variazione di impurità tra nodo padre e nodi figli*

```

function [bestsplitting] = splitting_criterion(data,class)
%funzione trova il miglior criterio di splitting

num_proprieta= size(data,2);
best_gain = -Inf;
best_feature = 0;
best_split = 0;

    for i = 1:num_proprieta
        val = unique(data(:,i));
        for j = 1:length(val)-1
            splitting = (val(j)+val(j+1))/2;
            valforsplit = data(:,i);
            gain = Gini_index(valforsplit,class,splitting);

            if gain>best_gain
                best_gain = gain;
                best_feature = i;
                best_split = splitting;
            end
        end
    end

    bestsplitting = [best_feature best_split];
end

```

**Codice 4:** *funzione splitting\_criterion.mat per trovare il miglior criterio di splitting per ogni nodo*

```

function [Accuratezza, Sensibilita, Specificita] =
metrichediaaccuratezza(pdata,class,positiveclass,negativeclass,alpha)
%funzione che calcola Accuratezza, Sensibilità e specificità del modello

a = length(alpha);
Accuratezza = zeros(a,1);
Sensibilita = zeros(a,1);
Specificita = zeros(a,1);
m = size(pdata,1);
    for i = 1:a

```

```

predclass = zeros(m,1);
for j = 1:m
    soglia = alpha(i);
    if pdata(j,1) >= soglia
        predclass(j) = 1;
    else
        predclass(j) = 0;
    end
end

%confronto con classe reale
ind = find(predclass == class);
Accuratezza(i) = length(ind)/m;
TP = 0;
TN = 0;
FP = 0;
FN = 0;

for g = 1: length(predclass)

    if predclass(g) == positiveclass && class(g) == positiveclass
        TP = TP +1;
    elseif predclass(g) == positiveclass && class(g) == negativeclass
        FP = FP +1;
    elseif predclass(g) == negativeclass && class(g) == negativeclass
        TN = TN+1;
    else
        FN = FN+1;
    end
end

Sensibilita(i) = TP /(TP + FN);
Specificita(i) = TN /(TN + FP);

end

end

```

**Codice 2:** *funzione metrichediaaccuratezza.m per il calcolo delle metriche di valutazione dei modelli LR, RF e DT*

## Algoritmi per la formazione dei modelli descritti nel Capitolo 3

```

function [Beta,Accuratezza,Sensibilita,Specificita] =
binomial_LR(trainingdata,testdata,k)

%funzione Binomial Logistic Regression con metodo di regolarizzazione
%Ridge e Netwon rapson per ottenere i Beta ottimi

%Input:
%trainingdata = matrice NXM contenente i dati utilizzati per addestrare il
%modello(N = numero di istanze considerate, M = numero di caratteristiche per
%istanza)
%testdata = matrice DXM contenente i dati utilizzati per testare il modello
%k = numero di fold da utilizzare per la cross-validation

%Output:
%Beta = beta ottimi
%Accuratezza = matrice di accuratezza del modello al variare della soglia
%Sensibilità = matrice di sensibilità del modello al variare della soglia

```

```
%Specificità = matrice di specificità del modello al variare della soglia
```

```
%estrazione della classe
trainingclass = trainingdata(:,end);
testclass = testdata(:,end);
trainingdata = trainingdata(:,1:end-1);
testdata = testdata(:,1:end-1);

%normalizzazione delle caratteristiche
media = mean(trainingdata,1);
sigma = std(trainingdata,0,1);
trainingdata = (trainingdata - media)./sigma;
testdata = (testdata - media)./sigma;

% k-cross validation per scegliere il parametro lambda ottimo
m = length(trainingclass);
xtrain = [ones(m,1) trainingdata];
[N,M] = size(xtrain);
toll = 0.001;
maxiter = 500;
lambdas = [0.0001 0.001 0.01 0.1];
l = length(lambdas);
beta = zeros(M,round(N/k));
accuratezza_media = zeros(1,l);
loss_media = zeros(1,l);

rng(42)
cv = cvpartition(trainingclass,"Kfold",k);

for h = 1:l
    loss_beta = zeros(k,1);
    accuratezza_beta = zeros(k,1);
    for i = 1:cv.NumTestSets
        idxTrain = training(cv,i);
        idxTest = test(cv,i);
        xtrainloo = xtrain(idxTrain,:);
        xclasstrn = trainingclass(idxTrain);
        xtest = xtrain(idxTest,:);
        xclasstst = trainingclass(idxTest);

        % ricerca dei beta ottimi con MLC usando metodo di Newton Raphson

        beta(:,i) = Newton_Rapson(xtrainloo,...
            xclasstrn,lambdas(h),maxiter,toll);

        %%calcolo accuratezza ed entropia
        ptest = zeros(length(xclasstst),1);
        classpred_test = zeros(length(xclasstst),1);
        for t = 1:size(xtest,1)
            ptest(t) = 1./(1+exp(-xtest(t,:)*beta(:,i)));
            if ptest(t) >0.5
                classpred_test(t) = 1;
            else
                classpred_test(t) = 0;
            end
        end

        eps = 1e-12;
        ptest = min(max(ptest,eps),1-eps);
        loss = -mean(xclasstst(:).*log(ptest(:)) + (1-xclasstst(:)).*log(1-
            ptest(:)));
    end
end
```

```

    loss_beta(i) = loss;
    ind = find(classpred_test == xclasstst);
    accuratezza_beta(i) = length(ind)/length(xclasstst);

    end

    accuratezza_media(h) = mean(accuratezza_beta);
    loss_media(h) = mean(loss_beta);
end

%scelgo lambda che minimizza la loss media e calcolo nuovamente i beta
%per intero data set
[~,ind] = min(loss_media);
lambda = lambdas(ind);
betaottimi = Newton_Rapson(xtrain,trainingclass,lambda,maxiter,toll);

%definisco probabilità di appartenenza alle diverse classi
ptrain = zeros(m,2);
for i = 1:m
    ptrain(i,1)= 1./(1+exp(-xtrain(i,:)*betaottimi));
    ptrain(i,2) = 1-ptrain(i,1);
end

n = size(testdata,1);
xtest = [ones(n,1) testdata];
ptest = zeros(n,2);

for i = 1:n
    ptest(i,1)= 1./(1+exp(-xtest(i,:)*betaottimi));
    ptest(i,2) = 1-ptest(i,1);
end

%metriche di accuratezza al variare di alpha
alpha = 1:-0.01:0;
positiveclass = 1;
negativeclass = 0;

[Accuratezza_train, Sensibilita_train, Specificita_train] = ...
    metrichediaccuratezza(ptrain,trainingclass,positiveclass,...
    negativeclass,alpha);

[Accuratezza_test, Sensibilita_test, Specificita_test] = ...
    metrichediaccuratezza(ptest,testclass,positiveclass,...
    negativeclass,alpha);

%output
Beta = betaottimi;
Accuratezza = [Accuratezza_train Accuratezza_test];
Sensibilita = [Sensibilita_train Sensibilita_test];
Specificita = [Specificita_train Specificita_test];

end

```

**Codice 6:** *funzione binomial\_LR.mat utilizzata per implementare modello LR*

```

function[Accuratezza,Sensibilita,Specificita]=binomial_DT(trainingdata,testdata,
    k)
%funzione per albero binomiale implementata con algortimo CART

%Input:

```

```

%trainingdata = matrice NXM contenente i dati utilizzati per addestrare il
%modello(N = numero di istanze considerate, M = numero di caratteristiche per
%istanza)
%testdata = matrice DXM contenente i dati utilizzati per testare il modello
%k = numero di fold da utilizzare per la cross-validation

%Output:
%Accuratezza = matrice di accuratezza del modello al variare della soglia
%Sensibilità = matrice di sensibilità del modello al variare della soglia
%Specificità = matrice di specificità del modello al variare della soglia

% estrazione classe
trainingclass = trainingdata(:,end);
testclass = testdata(:,end);
trainingdata = trainingdata(:,1:end-1);
testdata = testdata(:,1:end-1);
N = size(trainingdata,1);
D = size(testdata,1);
maxiter = 500;
RFflag = 0;

%CROSS VALIDATION
cv = cvpartition(trainingclass,"Kfold",k);
accuratezza_fold = zeros(k,1);

for i =1:cv.NumTestSets

    idxTrain = training(cv,i);
    idxTest = test(cv,i);
    xtrain = trainingdata(idxTrain,:);
    xclasstrn = trainingclass(idxTrain);
    xtest = trainingdata(idxTest,:);
    xclasstst = trainingclass(idxTest);

    iter = 0;
    [nodo,iter_fin] = DecisionTree(xtrain,xclasstrn,iter,maxiter,RFflag);

%calcolo accuratezza
classp = zeros(size(xtest,1),1);
for j = 1:size(xtest,1)
    soggt = xtest(j,:);
    nodonew = nodo;

    while ~nodonew.foglia

        proprietanodo = nodonew.proprieta;
        soglia = nodonew.soglia;
        featuresogg = soggt(proprietanodo);

        if featuresogg >= soglia
            nodonew = nodonew.destro;
        else
            nodonew = nodonew.sinistro;
        end
    end
    classp(j) = nodonew.class;
end
ind = find(classp == xclasstst);
accuratezza_fold(i)= length(ind)/size(xtest,1);

end

```

```

accuratezza_training_media = mean(accuratezza_fold);

%%addestramento su intero dataset di training
iter = 0;
[nodo,iter] = DecisionTree(trainingdata,trainingclass,iter,maxiter,RFflag);

%%definizione percentuali
ptrain = zeros(N,2);
pclasstrain = zeros(N,1);

for i = 1:N
    sogg = trainingdata(i,:);
    nodonew = nodo;

    while ~nodonew.foglia

        proprietanodo = nodonew.proprieta;
        soglia = nodonew.soglia;
        featuresogg = sogg(proprietanodo);

        if featuresogg >= soglia
            nodonew = nodonew.destro;
        else
            nodonew = nodonew.sinistro;
        end

    end
    ptrain(i,1) =nodonew.p;
    ptrain(i,2) =1-nodonew.p;
    pclasstrain(i) = nodonew.class;
end

ptest = zeros(D,2);
pclasstest = zeros(D,1);

for i = 1:D
    sogg = testdata(i,:);
    nodonew = nodo;

    while ~nodonew.foglia

        proprietanodo = nodonew.proprieta;
        soglia = nodonew.soglia;
        featuresogg = sogg(proprietanodo);

        if featuresogg >= soglia
            nodonew = nodonew.destro;
        else
            nodonew = nodonew.sinistro;
        end

    end
    ptest(i,1) =nodonew.p;
    ptest(i,2) =1-nodonew.p;
    pclasstest(i) = nodonew.class;
end

alpha = 1:-0.01:0;
positiveclass = 1;
negativeclass = 0;
[Accuratezza_train, Sensibilita_train, Specificita_train] = ...

```

```

    metricmediaaccuratezza(ptrain,trainingclass,positiveclass,...
    negativeclass,alpha);

[Accuratezza_test, Sensibilita_test, Specificita_test] = ...
    metricmediaaccuratezza(pctest,testclass,positiveclass,...
    negativeclass,alpha);

%output
Accuratezza = [Accuratezza_train Accuratezza_test];
Sensibilita = [Sensibilita_train Sensibilita_test];
Specificita = [Specificita_train Specificita_test];

end

```

**Codice 7:** *funzione binomial\_DT.mat utilizzata per definire il miglior DT*

```

function [n_alberi,Accuratezza,Sensibilita,Specificita]=RF(trainingdata,
    testdata,k,alberi)
%funzione per implementazione RF

%Input:
%trainingdata = matrice NXM contenente i dati utilizzati per addestrare il
%modello(N = numero di istanze considerate, M = numero di caratteristiche per
%istanza)
%testdata = matrice DXM contenente i dati utilizzati per testare il modello
%k = numero di fold da utilizzare per la cross-validation
%alberi = vettore del numero di alberi da formare

%Output:
%n_alberi = numero di alberi ottimo
%Accuratezza = matrice di accuratezza del modello al variare della soglia
%Sensibilità = matrice di sensibilità del modello al variare della soglia
%Specificità = matrice di specificità del modello al variare della soglia

% estrazione classe
trainingclass = trainingdata(:,end);
testclass = testdata(:,end);
trainingdata = trainingdata(:,1:end-1);
testdata = testdata(:,1:end-1);
N = size(trainingdata,1);
D = size(testdata,1);
maxiter = 500;
RFflag = 1;

% CROSS VALIDATION
cv = cvpartition(trainingclass,"Kfold",k);
l = length(alberi);
accuratezza_training_cv = zeros(l,1);

for h = 1:l
    accuratezza_fold = zeros(k,1);
    for i = 1:k

        idxTrain = training(cv,i);
        idxTest = test(cv,i);
        xtrain = trainingdata(idxTrain,:);
        xclasstrn = trainingclass(idxTrain);
        xtest = trainingdata(idxTest,:);
        xclasstst = trainingclass(idxTest);
        n_alberi = alberi(h);
    end
end

```

```

dim_train = size(xtrain,1);
Random_Forest = struct();

    rng(42)
    for j = 1:n_alberi
        r = randi([1,dim_train],dim_train,1);
        training_i = xtrain(r,:);
        class_i = xclasstrn(r);
        iter = 0;
        [Random_Forest(j).albero,~]=DecisionTree(training_i,class_i,...
            iter,maxiter,RFflag);
    end

classrf = zeros(size(xtest,1),1);

for t = 1:size(xtest,1)

    sogg = xtest(t,:);
    classp = zeros(n_alberi,1);

    for j = 1:n_alberi
        nodo = Random_Forest(j).albero;

        while ~nodo.foglia
            proprieta = nodo.proprieta;
            soglia = nodo.soglia;
            feature = sogg(proprieta);
            if feature >= soglia
                nodo = nodo.destro;
            else
                nodo = nodo.sinistro;
            end
        end

        classp(j) = nodo.class;

    end
    classrf(t) = round(mean(classp));
end

ind = find(classrf == xclasstst);
accuratezza_fold(i) = length(ind)/size(xtest,1);

end

accuratezza_training_cv(h) = mean(accuratezza_fold);
end

% TRAINING FINALE SU TUTTO IL TRAINING SET

[~,ind] = max(accuratezza_training_cv);
n_alberi = alberi(ind);
dim_train = size(trainingdata,1);
Random_Forest = struct();

rng(42)
for i = 1:n_alberi

    r = randi([1,dim_train],dim_train,1);
    training_i = trainingdata(r,:);
    class_i = trainingclass(r);

```

```

    iter = 0;
    [Random_Forest(i).albero,~]=DecisionTree(training_i,class_i,...
                                             iter,maxiter,RFflag);

end

% probabilità classi TRAIN
ptrain = zeros(N,n_alberi,2);

for i = 1:N

    sogg = trainingdata(i,:);

    for j = 1:n_alberi

        nodo = Random_Forest(j).albero;

        while ~nodo.foglia

            proprieta = nodo.proprieta;
            soglia = nodo.soglia;
            feature = sogg(proprieta);

            if feature >= soglia
                nodo = nodo.destro;
            else
                nodo = nodo.sinistro;
            end

        end

        ptrain(i,j,1) = nodo.p;
        ptrain(i,j,2) = 1 - nodo.p;

    end

end

% probabilità classi TEST
ptest = zeros(D,n_alberi,2);

for i = 1:D

    sogg = testdata(i,:);

    for j = 1:n_alberi

        nodo = Random_Forest(j).albero;

        while ~nodo.foglia

            proprieta = nodo.proprieta;
            soglia = nodo.soglia;
            feature = sogg(proprieta);

            if feature >= soglia
                nodo = nodo.destro;
            else
                nodo = nodo.sinistro;
            end

        end

    end
end

```

```

        ptest(i,j,1) = nodo.p;
        ptest(i,j,2) = 1 - nodo.p;

    end

end

%score medi forest
score_train = squeeze(mean(ptrain(:, :, 1), 2));
score_test = squeeze(mean(ptest(:, :, 1), 2));

% metriche
alpha = 1:-0.01:0;
positiveclass = 1;
negativeclass = 0;

[Accuratezza_train, Sensibilita_train, Specificita_train] = ...
    metrichediaccuratezza(ptrain, trainingclass, positiveclass, ...
        negativeclass, alpha);

[Accuratezza_test, Sensibilita_test, Specificita_test] = ...
    metrichediaccuratezza(ptest, testclass, positiveclass, ...
        negativeclass, alpha);

% output
n_alberi = alberi(ind);
Accuratezza = [Accuratezza_train Accuratezza_test];
Sensibilita = [Sensibilita_train Sensibilita_test];
Specificita = [Specificita_train Specificita_test];

end

```

**Codice 8:** *funzione RF.mat utilizzata per implementare RF con algoritmo CART*

```

%% MAIN per confronto tra modelli

close all
clear
clc

%% carico gli indicatori ottenuti per i diversi pazienti
indicatori= readtable('CA_IC.xlsx', "VariableNamingRule", "preserve", 'Sheet',
    'sw100');

%% divisione della classe dalle proprieta

%% IC
classe = indicatori{:,end};
proprieta= [indicatori{:,3:6} indicatori{:,14:16} ];

nomi = indicatori.Properties.VariableNames;
nomiP = [nomi(2:6) nomi(14:17)];

tipo_stimolazione = string(indicatori{:,2});
tipo_stimolazione(tipo_stimolazione == 'rarefazione')= 0;
tipo_stimolazione(tipo_stimolazione == 'condensazione') = 1;
tipo_stimolazione = str2double(tipo_stimolazione);

proprieta = [tipo_stimolazione proprieta];

```

```

%% visualizzazione relazione tra variabili con relativo p-value

proprietà1 = [proprietà classe];
figure('Name','Relazione tra le proprietà')
corrplot(proprietà1,'varNames',nomiP,'TestR','on')

%% divisione dataset in training set e test set

probclasse1 = mean(classe);
probclasse0 = 1-mean(classe);
n = length(classe);
m = round(n*0.75);
class1training = 1;
class1test = 0;

iter = 0;
maxiter = 200;

rng(10)
while abs(class1training-class1test) > 0.05 && iter<maxiter
    iter = iter+1;
    id = randperm(n,m);
    trainingclass = classe(id);
    testclass = classe;
    testclass(id) = [];

    trainingdata = proprietà(id,:);
    testdata = proprietà;
    testdata(id,:) = [];

    class1training = mean(trainingclass);
    class1test = mean(testclass);
end

%% modelli
training= [trainingdata trainingclass];
test = [testdata testclass];
k = 5;
alberi = 10:10:100;

[Beta,Accuratezza_LR,Sensibilita_LR,Specificita_LR] =
binomial_LR(training,test,k);
[Accuratezza_DT,Sensibilita_DT,Specificita_DT] = binomial_DT(training,test,k);
[n_alberi,Accuratezza_RF,Sensibilita_RF,Specificita_RF] =
RF(training,test,k,alberi);

%% ROC curve della Logistic Regression

FPR_train_LR = 1-Specificita_LR(:,1);
FPR_test_LR = 1-Specificita_LR(:,2);

[FPR_sort_train_LR,idx] = sort(FPR_train_LR);
TPR_sort_train_LR = Sensibilita_LR(idx,1);
FPR_sort_train_LR= [0 FPR_sort_train_LR' 1];
TPR_sort_train_LR= [0 TPR_sort_train_LR' 1];
AUC_train_LR = trapz(FPR_sort_train_LR,TPR_sort_train_LR);

[FPR_sort_test_LR,idx] = sort(FPR_test_LR);
TPR_sort_test_LR = Sensibilita_LR(idx,2);
FPR_sort_test_LR= [0 FPR_sort_test_LR' 1];

```

```

TPR_sort_test_LR= [0 TPR_sort_test_LR' 1];
AUC_test_LR = trapz(FPR_sort_test_LR,TPR_sort_test_LR);

%% ROC curve Decision Tree

FPR_train_DT = 1-Specificita_DT(:,1);
FPR_test_DT = 1-Specificita_DT(:,2);

[FPR_sort_train_DT,idx] = sort(FPR_train_DT);
TPR_sort_train_DT = Sensibilita_DT(idx,1);
FPR_sort_train_DT= [0 FPR_sort_train_DT' 1];
TPR_sort_train_DT= [0 TPR_sort_train_DT' 1];
AUC_train_DT = trapz(FPR_sort_train_DT,TPR_sort_train_DT);

[FPR_sort_test_DT,idx] = sort(FPR_test_DT);
TPR_sort_test_DT = Sensibilita_DT(idx,2);
FPR_sort_test_DT= [0 FPR_sort_test_DT' 1];
TPR_sort_test_DT= [0 TPR_sort_test_DT' 1];
AUC_test_DT = trapz(FPR_sort_test_DT,TPR_sort_test_DT);

%ROC RF

FPR_train_RF = 1-Specificita_RF(:,1);
FPR_test_RF = 1-Specificita_RF(:,2);

[FPR_sort_train_RF,idx] = sort(FPR_train_RF);
TPR_sort_train_RF = Sensibilita_RF(idx,1);
FPR_sort_train_RF= [0 FPR_sort_train_RF' 1];
TPR_sort_train_RF= [0 TPR_sort_train_RF' 1];
AUC_train_RF = trapz(FPR_sort_train_RF,TPR_sort_train_RF);

[FPR_sort_test_RF,idx] = sort(FPR_test_RF);
TPR_sort_test_RF = Sensibilita_RF(idx,2);
FPR_sort_test_RF= [0 FPR_sort_test_RF' 1];
TPR_sort_test_RF= [0 TPR_sort_test_RF' 1];
AUC_test_RF = trapz(FPR_sort_test_RF,TPR_sort_test_RF);

%% ROC DI CONFRONTO TRA MODELLI

figure(5)
plot([0 1],[0 1],'r','LineWidth',2)
hold on
plot(FPR_sort_test_LR,TPR_sort_test_LR,'b','LineWidth',2)
plot(FPR_sort_test_DT,TPR_sort_test_DT,'k','LineWidth',2)
plot(FPR_sort_test_RF,TPR_sort_test_RF,'g','LineWidth',2)

ylim([-0.05 1.05])
xlim([-0.01 1])

xlabel('False Positive Ratio (FPR)')
ylabel('True Positive Ratio (TPR)')
title('ROC a 100 sweeps')
legend('Chance classifier', ...
      ['LR AUC = ' num2str(AUC_test_LR)], ...
      ['DT AUC = ' num2str(AUC_test_DT)], ...
      ['RF AUC = ' num2str(AUC_test_RF)])

%% trovare soglia migliore
sensibilita_train = [Sensibilita_LR(:,1) Sensibilita_DT(:,1)
Sensibilita_RF(:,1)];

```

```

specificita_train = [Specificita_LR(:,1) Specificita_DT(:,1)
Specificita_RF(:,1)];

indice_youden = (sensibilita_train+specificita_train)-1;
[val] = max(indice_youden,[],1);

for i = 1:length(val)
    [idsoglia(i),~]= find(val(i) == indice_youden(:,i),1,'first');
end

accuratezza_finale = [Accuratezza_LR(:,2) Accuratezza_DT(:,2)
Accuratezza_RF(:,2)];
sensibilita_finale = [Sensibilita_LR(:,2) Sensibilita_DT(:,2)
Sensibilita_RF(:,2)];
specificita_finale = [Specificita_LR(:,2) Specificita_DT(:,2)
Specificita_RF(:,2)];

alpha = 1:-0.01:0;
bestsoglia = zeros(1,3);
for i = 1:length(idsoglia)
    bestsoglia(i) = alpha(idsoglia(i));

    if bestsoglia(i) >0.90 || bestsoglia(i) < 0.1
        bestsoglia(i) = probclasse1;
        [~,idsoglia(i)] = min(abs(alpha - probclasse1));
    end
    Accuratezza(i) = accuratezza_finale(idsoglia(i),i);
    Sensibilita(i) = sensibilita_finale(idsoglia(i),i);
    Specificita(i) = specificita_finale(idsoglia(i),i);
end

AUC = [AUC_test_LR AUC_test_DT AUC_test_RF];

LR = [val(1); bestsoglia(1); AUC(1); Accuratezza(1); Sensibilita(1);
Specificita(1)];
DT = [val(2); bestsoglia(2); AUC(2); Accuratezza(2); Sensibilita(2);
Specificita(2)];
RF = [val(3); bestsoglia(3); AUC(3); Accuratezza(3); Sensibilita(3);
Specificita(3)];

metriche = {'Indice Youden'; 'Best soglia'; 'AUC'; 'Accuratezza'; 'Sensibilita';
'Specificita'};

tabella_risultati = table(metriche, LR, DT, RF, ...
    'VariableNames', {'Metrica','LR','DT','RF'});
disp(tabella_risultati)

```

**Codice 9:** main per confronto tra modelli di rilevazione FFR

## Codici per indici di capacità uditiva

```

function [data_new,index,U,pc,variabilitatot,pesi,erroremedio] =
principal_component_analysis(data,nomiP,variabilita_richiesta,normalizzazione,
plot)
%funzione implementante PCA

%Input
%data = matrice delle istanze
%variabilità_richiesta = percentuale di variabilità che devono spiegare i dati
%normalizzazione= flag che indica se i dati sono stati normalizzati o meno

```

```

%Output
%data_new = dati rappresentati con solo le componenti principali
%index = indici non normalizzati
%U = matrice degli autovettori delle componenti principali considerate
%pc = quantità di componenti considerate
%erroremedio = errore medio ottenuto considerando solo le componenti
%principali per rappresentare i dati

%normalizzazione
if normalizzazione ==1
    P = zscore(data);
else
    P = data;
end

% calcolo matrice di covarianza ,i suoi autovalori e autovettori
SIGMA = cov(P);
[autovet_SIGMA, autoval_SIGMA] = eig(SIGMA);

% ordinare gli autovalori dal più grande al più piccolo
lambda= diag(autoval_SIGMA);
[val_sort,ind] = sort(lambda,'descend');
LAMBDA = diag(val_sort);
U = autovet_SIGMA(:,ind);

% definire percentuale di variabilità espressa da ogni autovalore
tot = sum(val_sort);
variabilita = (val_sort/tot)*100;
variabilitatot = 0;
pc = 0;

while variabilitatot < variabilita_richiesta && pc < length(variabilita)
    pc= pc+1;
    variabilitatot = variabilitatot+variabilita(pc);
end

if plot == 0
    for i = 1:pc

        propr = U(:,i);
        importanza = 100*(propr.^2)/sum(propr.^2);
        [imp_sort,idx] = sort(importanza,'descend');
        nomip = nomiP(idx);
        figure(i+1)
        bar(imp_sort)
        set(gca,'XTickLabel',nomip)
        ylabel('importanza variabili')
        title(['Importanza delle variabili su PC ' num2str(i)])

    end
end

nuova_rapp = P*U;
index = P*U(:,1:pc);
pesi = val_sort(1:pc)/sum(val_sort(1:pc));

% riduco la dimansione di r considerando solo le componenti per cui la
% variabilità cumulativa è maggiore della variabilità richiesta
p = nuova_rapp;
p(:,pc+1:end) = 0;

```

```

erroremedio = sum(val_sort(pc+1:end));

% valori ricostruiti
X_std = p*U';

if normalizzazione == 1
    data_new = X_std.*std(data) + mean(data);
else
    data_new = X_std + mean(data);
end

end

```

**Codice 10:** *funzione principal\_component\_analysis.mat per rappresentazione dei dati attraverso le componenti principali*

```

%% FORMAZIONE INDICE DI CAPACITA' UDITIVA

%% caricamento dei dati, selezione delle proprietà e normalizzazione
indicatori= readtable('APWF_IC.xlsx', 'Sheet', 'sw500');

%% ic
classe = indicatori(:,17);
proprieta =[indicatori(:,3:6) -indicatori(:,14:15) indicatori(:,16)];
nomi = indicatori.Properties.VariableNames;
nomiP = [nomi(3:6) nomi(14:16)];

%% normalizzazione delle proprietà
P = zscore(proprieta);

%% Barlett-test per definire correlazioni tra variabili e giustificazione PCA

[N,M] = size(P);
R = corr(P);
D = det(R);
Bartlett_test = -(N-1-((2*M+5)/6))*log(D);
df = M*(M-1)/2;
p_value = 1 - chi2cdf(Bartlett_test,df);

disp(['Bartlett chi-square = ' num2str(Bartlett_test)])
disp(['df = ' num2str(df)])
disp(['p-value = ' num2str(p_value)])

%% applicazione PCA

variabilita_richiesta = 90;
normalizzazione = 0;
[~,pc_score,U,pc,variabilitatot,pesi,~] =
principal_component_analysis(P,nomiP,variabilita_richiesta,normalizzazione,1);

disp(['Numero componenti principali: ' num2str(pc)])
disp(['variabilità spiegata: ' num2str(variabilitatot)])
disp('Pesi delle componenti:')
disp(pesi)

%% ricavo gli indici
%tracking encoding
ICU1 = pc_score(:,1);
index_periodic_encoding = 100*(ICU1 - min(ICU1)) / (max(ICU1) - min(ICU1));

```

```

%transitory encoding
ICU2 = pc_score(:,2);
transitory_encoding = 100*(ICU2 - min(ICU2)) / (max(ICU2) - min(ICU2));

%global index
ICU_TOT = pc_score*pesi;
ICU_TOT = 100*(ICU_TOT - min(ICU_TOT)) / (max(ICU_TOT) - min(ICU_TOT));

min_ICU1 = min(ICU1);
max_ICU1 = max(ICU1);
min_ICU2 = min(ICU2);
max_ICU2 = max(ICU2);
min_ICUTOT = min(ICU_TOT);
max_ICUTOT = max(ICU_TOT);

% Indice medio per classe
classi = unique(classe);
mean_ICU = zeros(length(classi),1);
std_ICU = zeros(length(classi),1);
mean_ICU1 = zeros(length(classi),1);
std_ICU1 = zeros(length(classi),1);
mean_ICU2 = zeros(length(classi),1);
std_ICU2 = zeros(length(classi),1);

for i = 1:length(classi)
    idx = classe == classi(i);
    mean_ICU(i) = mean(ICU_TOT(idx));
    std_ICU(i) = std(ICU_TOT(idx));

    mean_ICU1(i) = mean(index_periodic_encoding(idx));
    std_ICU1(i) = std(index_periodic_encoding(idx));

    mean_ICU2(i) = mean(transitory_encoding(idx));
    std_ICU2(i) = std(transitory_encoding(idx));
end

%% Test affidabilità

%% correlazione tra indice e proprietà originali
correlazioni_ICU_1= corr(index_periodic_encoding, P)';
correlazioni_ICU_2= corr(transitory_encoding, P)';
correlazioni_ICU_tot = corr(ICU_TOT,P)';

%% test robustezza
k = 5;
rng(42)
cv = cvpartition(classe, "KFold",k);
pc_k = zeros(k,1);

for i =1:cv.NumTestSets

    idxTrain = training(cv,i);
    idxTest = test(cv,i);
    xtrain = P(idxTrain,:);
    xclasstrn = classe(idxTrain);
    xtest = P(idxTest,:);
    xclasstst = classe(idxTest);

    [~,pc_score,U,pc1,~,pesi_t,~] = principal_component_analysis(xtrain,nomiP,
        variabilita_richiesta,normalizzazione,1);

```

```

ICU_test = xtest*U(:,1:pc);

%tracking encoding
ICU1_test(idxTest,1) = ICU_test(:,1);

%transitory encoding
ICU2_test(idxTest,1) = ICU_test(:,2);
ICU_TOT_test(idxTest,1) = ICU_test*pesi_t;

pc_k(i) = pc1;
end

pc_medie = mean(pc_k);
pc_std = std(pc_k);

index_periodic_encoding_test = 100*(ICU1_test- min(ICU1_test)) /
(max(ICU1_test) - min(ICU1_test));
transitory_encoding_test= 100*(ICU2_test - min(ICU2_test)) / (max(ICU2_test) -
min(ICU2_test));
ICU_TOT_test = 100*(ICU_TOT_test - min(ICU_TOT_test)) / (max(ICU_TOT_test) -
min(ICU_TOT_test));

corr_pe= corr(index_periodic_encoding,index_periodic_encoding_test);
corr_te = corr(transitory_encoding,transitory_encoding_test);
corr_gi = corr(ICU_TOT,ICU_TOT_test);

disp(['correlazione indice 1: ' num2str(corr_pe)])
disp(['correlazione indice 2: ' num2str(corr_te)])
disp(['correlazione indcie globale: ' num2str(corr_gi)])

```

**Codice 11:** main per formazione indici di capacità uditiva e valutazione robustezza

# Bibliografia

- [1] <https://www.blumedical.com/come-funziona-il-nostro-orecchio/>
- [2] Sheikh A, Bint-e-Zainab, Shabbir K, Imtiaz A. *Structure and Physiology of Human Ear Involved in Hearing*. In: Naz S, editor. *Auditory System - Function and Disorders*. IntechOpen; 2022. DOI: 10.5772/intechopen.105466.
- [3] Hackett T. A. *Information flow in the auditory cortical network*. *Hearing Research*. 2011;271(1–2):133–146. DOI:10.1016/j.heares.2010.
- [4] Kraus N., Nicol T. *Aggregate neural responses to speech sounds in the central auditory system*. *Speech Communication* 41(1–2), 35–47, 2003
- [5] Smith, J. C., Marsh, J. T. & Brown, W. S. *Far-field recorded frequency following responses: evidence for the locus of brainstem sources*. *Electroencephalogr. Clin. Neurophysiol.* 39, 465–472, 1975
- [6] Sohmer, H., Pratt, H. & Kinarti, R. *Sources of frequency following responses (FFR) in man*. *Electroencephalogr. Clin. Neurophysiol.* 42, 656–664, 1977.
- [7] Coffey B.J. E, Nicol T, White-Schwoch T, Chandrasekaran B, Krizman J, Skoe E, Zatorre R. J, Kraus N. *Evolving perspectives on the sources of the frequency-following response*. *Nature Communications*, 2019. <https://doi.org/10.1038/s41467-019-13003-w>
- [8] G. Nike Gnanateja et al. *Frequency-Following Responses to Speech Sounds Are Highly Conserved across Species and Contain Cortical Contributions*. *eNeuro*, 2021. <https://doi.org/10.1523/ENEURO.0451-21.2021>
- [9] Kraus N., Krizman J. *Analyzing the FFR: A tutorial for decoding the richness of auditory function*. *Hearing Research*, 382, 2019
- [10] Skoe, E. & Kraus, N. *Auditory brain stem response to complex sounds: a tutorial*. *Ear Hear.* 31, 302–324, 2010.
- [11] Tallal P, Stark RE. *Speech acoustic-cue discrimination abilities of normally developing and language impaired children*. *The Journal of the Acoustical Society of America* 1981;69:568–574.[PubMed: 7462479]
- [12] Turner CW, Fabry DA, Barrett S, et al. *Detection and recognition of stop consonants by normal hearing and hearing impaired listeners*. *J Speech Hear Res* 1992;35:942–949. [PubMed:1405550]

- [13] Banai, K., Hornickel, J., Skoe, E., Nicol, T., Zecker, S., Kraus, N., 2009. *Reading and subcortical auditory function*. *Cerebr. Cortex* 19, 2699e2707.
- [14] Hornickel, J., Kraus, N., 2013. *Unstable representation of sound: a biological marker of dyslexia*. *J. Neurosci.* 33, 3500e3504.
- [15] Musacchia, G., Hu, J., Bhutani, V.K., Wong, R.J., Tong, M.-L., Han, S., Blevins, N.H., Fitzgerald, M.B., 2019. *Frequency-following response among neonates with progressive moderate hyperbilirubinemia*. *J. Perinatol.* <https://doi.org/10.1038/s41372-019-0421-y>.
- [16] Anderson, S., Skoe, E., Chandrasekaran, B., Kraus, N., 2010a. *Neural timing is linked to speech perception in noise*. *J. Neurosci.* 30, 4922e4926.
- [17] Musacchia, G., Sams, M., Skoe, E., Kraus, N., 2007. *Musicians have enhanced subcortical auditory and audiovisual processing of speech and music*. *Proc. Natl. Acad. Sci. U.S.A.* 104, 15894e15898.
- [18] Song, J.H., Skoe, E., Wong, P.C.M., Kraus, N., 2008. *Plasticity in the adult human auditory brainstem following short-term linguistic training*. *J. Cogn. Neurosci.* 20, 1892e1902.
- [19] Krizman, J., Marian, V., Shook, A., Skoe, E., Kraus, N., 2012b. *Subcortical encoding of sound is enhanced in bilinguals and relates to executive function advantages*. *Proc. Natl. Acad. Sci.* 109, 7877e7881
- [20] Skoe, E., Krizman, J., Kraus, N., 2013a. *The impoverished brain: disparities in maternal education affect the neural response to sound*. *J. Neurosci.* 33, 17221e17231.
- [21] Russo, N.M., Skoe, E., Trommer, B., Nicol, T., Zecker, S., Bradlow, A., Kraus, N., 2008. *Deficient brainstem encoding of pitch in children with Autism Spectrum Disorders*. *Clin. Neurophysiol.* 119, 1720e1731.
- [22] J. P. C. de Weerd and W. L. J. Martens, 'Theory and practice of a posteriori "Wiener" filtering of average evoked potentials', *Biol. Cybernetics*, vol. 30, no. 2, pp. 81–94, June 1978, doi: 10.1007/BF00337321.
- [23] M. Hoke, B. Ross, R. Wickesberg, and B. Lütkenhöner, 'Weighted averaging — theory and application to electric response audiometry', *Electroencephalography and Clinical Neurophysiology*, vol. 57, no. 5, pp. 484–489, May 1984, doi: 10.1016/0013-4694(84)90078-6.
- [24] Q. Xu and D. Ye, 'Evaluation of a posteriori Wiener filtering applied to frequency-following response extraction in the auditory brainstem', *Biomedical Signal Processing and Control*, vol. 14, pp. 206–216, Nov. 2014, doi: 10.1016/j.bspc.2014.08.003.

- [25] Galassi C, *Metodi di elaborazione del segnale per l'ottimizzazione della misura di Frequency Following Response nella diagnosi differenziale audiologica*, 2025
- [26] Zhu, L., Bharadwaj, H., Xia, J., Shinn-Cunningham, B., 2013. *A comparison of spectral magnitude and phase-locking value analyses of the frequency-following response to complex tones*. J. Acoust. Soc. Am. 134, 384e395.
- [27] Tierney, A., Kraus, N., 2013. *The ability to move to a beat is linked to the consistency of neural responses to sound*. J. Neurosci. 33, 14981e14988.
- [28] Franguelli M, *Auditory frequency following response: pilot evaluation of novel stimulation and signal processing strategies in the clinical diagnosis of developmental disorders and cognitive decline*, 2024
- [29] P. C. M. Wong, E. Skoe, N. M. Russo, T. Dees, and N. Kraus, 'Musical experience shapes human brainstem encoding of linguistic pitch patterns', *Nat Neurosci*, vol. 10, no. 4, pp. 420–422, Apr. 2007, doi: 10.1038/nm1872.
- [30] A. Krishnan, Y. Xu, J. Gandour, and P. Cariani, 'Encoding of pitch in the human brainstem is sensitive to language experience', *Cognitive Brain Research*, vol. 25, no. 1, pp. 161–168, Sept. 2005, doi: 10.1016/j.cogbrainres.2005.05.004.
- [31] N. Kraus and T. Nicol, 'The Cognitive Auditory System: The Role of Learning in Shaping the Biology of the Auditory System', in *Perspectives on Auditory Research*, New York, NY: Springer New York, 2014, pp. 299–319. doi: 10.1007/978-1-4614-9102-6\_17.
- [32] A. Krishnan, Y. Xu, J. T. Gandour, and P. A. Cariani, 'Human frequency-following response: representation of pitch contours in Chinese tones', *Hearing Research*, vol. 189, no. 1–2, pp. 1–12, Mar. 2004, doi: 10.1016/s0378-5955(03)00402-7.
- [33] OpenAI. (2026). *ChatGPT* (modello GPT-5.3) [Strumento di intelligenza artificiale]. <https://chat.openai.com/>
- [34] S. W. Bauer, F. C. Jeng e A. Carriero, *Machine Learning Recognizes Frequency-Following Responses in American Adults: Effects of Reference Spectrogram and Stimulus Token*, in *Perceptual and Motor Skills*, Vol. 131(5) 1584–1602, 2024. DOI: 10.1177/00315125241273993
- [35] [https://datapeaker.com/it/big-data/regresion-logistica-guia-para-principiantes-de-regresion-logistica-con-python/#google\\_vignette](https://datapeaker.com/it/big-data/regresion-logistica-guia-para-principiantes-de-regresion-logistica-con-python/#google_vignette)
- [36] Shai Shalev-Shwartz and Shai Ben-David, *Understanding Machine Learning: From Theory to Algorithms*, 2014. <http://www.cs.huji.ac.il/~shais/UnderstandingMachineLearning>

- [37] Fuh-Cherng Jeng, M.D., Ph.D. and Yu-Shiang Jeng, B.S, *Implementation of Machine Learning on Human Frequency-Following Responses: A Tutorial*, Seminars in Hearing/Volume 43, Number 3, 2022
- [38] G. James, D. Witten, T. Hastie, R. Tibshirani, *An Introduction to Statistical Learning with Applications in R*, Springer, 2013. DOI 10.1007/978-1-4614-7138-7
- [39] <https://datafx.it/blog/decision-tree-p1/>
- [40] S. Marsland, *Machine Learning: An Algorithmic Perspective*, second edition, 2014
- [41] B. Talekar and S. Agrawal, *A Detailed Review on Decision Tree and Random Forest*, Biosc.Biotech.Res.Comm. Special Issue Vol 13 No 14 (2020) Pp-245-248.  
DOI: <http://dx.doi.org/10.21786/bbrc/13.14/57>
- [42] N. L. Deep, E. M. Dowling, D. Jethanamest e M. L. Carlson, *Cochlear Implantation: An Overview*, J Neurol Surg B 2019;80:169–177, 2018. DOI <https://doi.org/10.1055/s-0038-1669411>
- [43] Zheng et al.: *CI Artifacts Removal In EEG-Based Objective Auditory Rehabilitation Assessment*, IEEE Transactions on Neural Systems and Rehabilitation Engineering, Vol. 32, 2024
- [44] Gransier et al.: *Artifact removal by template subtraction enables recordings of the frequency following response in cochlear-implant users*, Scientific Reports, 2024.  
<https://doi.org/10.1038/s41598-024-56047-9>
- [45] J. C. Middlebrooks et al.: *Chronic Auditory-Nerve Implant Enhances Brainstem Phase Locking to Electric Pulse Trains*, Journal of the Association for Research in Otolaryngology (2025) 26:601–620 <https://doi.org/10.1007/s10162-025-01003-1>
- [46] <https://blog.alliedoffsets.com/boost-your-binary-classification-game-auc-roc-vs-auc-pr-which-one-should-you-use>
- [47] X. Zhang and Q. Gong, *Correlation between the frequency difference limen and an index based on principal component analysis of the frequency-following response of normal hearing listeners*, Hearing Research 344 (2017) 255-264