
Università di Verona
Dipartimento di Informatica
Laurea Magistrale in Ingegneria e Scienze Informatiche

**Apprendimento non supervisionato per
l'identificazione di contesti di
Freezing of Gait (FoG)
in pazienti affetti da morbo di Parkinson**

Candidato
Fuser Alessandro

Supervisore
Graziano Pravadelli

Assistente Supervisore
Florenc Demrozi

Sessione di Laurea del 20 Marzo 2018
Anno Accademico 2016/2017

Abstract

Acknowledgements

Indice

Acknowledgements	v
1 Introduction	1
1.1 Thesis Contribution	1
1.2 Outline	1
2 Letteratura	3
3 Background	7
3.1 Freezing Of Gait	7
3.1.1 Le tipologie di Fog	8
3.1.2 Influenza del FoG nella camminata	9
3.2 Machine Learning	10
3.2.1 Cluster Analysis	13
3.2.2 Classificazione	15
3.3 Riduzione della dimensionalità	17
3.3.1 Analisi Delle Componenti Principali	17
3.3.2 Analisi dei Discriminanti Lineari	19
4 Motivations and Goals	21
5 Self-adaptive TLM model	23
6 Software Implementation	25
7 Case Studies	27
8 Conclusions and Future Work	29
Bibliography	31

Elenco delle figure

3.1	Diversi tipi di Freezing esistenti e loro percentuale di incidenza su un certo gruppo di malati di Parkinson. [Fonte: J.M. Shine et al., 2012]	9
3.2	Quadro concettuale relativo al Freezing of Gait (FoG) riguardante aspetti mentali e motori.	10
3.3	Schema delle tipologie di Machine Learning	12
3.4	Esempio di Workflow tramite Machine Learning	13
3.5	Rappresentazione dei dati e dei gruppi ottenuti con la cluster analysis	14
3.6	Misure di Similarità	14
3.7	Rappresentazione dei possibili approcci usando strategie di clustering	16
3.8	Rappresentazione del cambio di dimensionalità tramite PCA	18
3.9	Rappresentazione del cambio di dimensionalità tramite LDA	20

Introduction

1.1 Thesis Contribution

In this thesis, we have developed a methodology

1.2 Outline

The rest of the thesis is organized as follows: Chapter 2

Letteratura

Il problema del FOG è stato analizzato tramite una grande varietà di sistemi e sensori. Alcuni di questi, però, non sono utilizzabili durante la vita quotidiana dei pazienti poiché posso essere disponibili solo in ambienti di laboratorio. Esempi di questi sistemi sono le piattaforme di pressione[18], le quali sono non portatili, l'elettromiografia (EMG)[6], l'elettroencefalogramma (EEG)[21] o la conduttanza della pelle[24], il quale comporta il pizzamento di elettrodi sulla pelle in aggiunta ad sistema di rilevamento per raccogliere i dati. Altri sistemi invasivi sono i goniometri a ginocchio[1] o sistemi che fanno uso di camere e video, i quali hanno una bassa tolleranza del paziente in un'ambiente che non sia di laboratorio[1, 2, 5]. Quindi, dato che il monitoraggio del PD dovrebbe essere deambulatorio e durare diverse ore al fine di ricavare utili informazioni cliniche[9, 4], la maggior parte dei lavori si è basata su sistemi non invasivi come i dispositivi indossabili basati su circuiti microelettromeccanici (MEMS).

Nel 2003, Han et al. hanno usato MEMS basati su sistemi inerziali, come gli accelerometri, per esplorare le caratteristiche collegate agli episodi di FoG. Hanno trovato che la frequenza di risposta nei pazienti che indossavano gli accelerometri nella caviglia era intorno ai 6-8 Hz[17]. Nel 2008, Moore et al. hanno proposto una metodologia per identificare FoG con un'accelerometro posizionato nella caviglia nella quale hanno descritto il Freezing Index (FI), ossia il quoziente del rapporto della densità spettrale di potenza (PSD) tra 3 ed 8 Hz, chiamata Freezing Band (FB), con la PSD tra 0.5 e 3 Hz, denominata Walking Band (WB)[20]. Quando il FI supera una certa soglia (Freezing Threshold (FTH)), si considera che si sia verificato un episodio di FoG. A causa della presenza dei falsi positivi (FP) quando il paziente è a riposo, Bachlin et al. hanno introdotto il concetto di Power Index (PI), definito come la somma della WB e FB, il quale viene comparato con la Power Threshold (PTH) al fine di stabilire se c'era una quantità rilevante di movimento nel momento in cui il FI era alto, ossia oltre la soglia[20]. PI indica la quantità di movimento, perciò situazioni nelle quali il paziente non si stesse muovendo volontariamente sono state eliminate. In quest'ultima versione dell'algoritmo, quindi, un episodio di FoG è occorso se $FI > FTH$ e $PI > PTH$. Questo

metodo è il più avanzato nella detenzione di FoG dato il suo scarso costo computazionale e le sue buone performance[13].

L'algoritmo MBFA è stato ampiamente utilizzato nell'analisi del FoG, anche se di solito in condizioni di laboratorio e molto spesso con pochi pazienti. Jovanov et al. hanno implementato un algoritmo real time, anche se un solo volontario è stato usato per testare l'algoritmo. Inoltre, nessun risultato su sensibilità e specificità è stato riportato[13]. Zabaleta et al. hanno analizzato il FoG per mezzo di accelerometri a tre assi e giroscopi a due assi in differenti locazioni degli arti inferiori. La caratteristica principale ad essere stata analizzata è il FI in congiunzione con i cambiamenti della densità spettrale di potenza. Sono stati capaci di identificare correttamente l'82.7% degli episodi di FoG con i sensori inerziali posizionati su entrambe le caviglie, anche se in soli 2 pazienti[15].

Negli anni più recenti, Niazmand et al. (2011) hanno presentato il Mimed-Pants[19], pantaloni da jogging lavabili con 5 accelerometri integrati. Hanno usato MBFA per identificare FoG, ottenendo un 88.3% in sensibilità e 85.3% in specificità con 6 pazienti in brevi e controllati test focalizzati nell'indurre FoG senza tenere conto dei FP. Nel 2012, Zhao et al.[28] hanno sviluppato un algoritmo embedded basato sull'approccio MBFA all'interno del sistema Mimed-Pants ottenendo un 81% in sensibilità con 8 pazienti usando dei test simili ai precedenti. Più recentemente, Mazilu et al. hanno proposto un nuovo algoritmo online usando 3 accelerometri ed comparando diversi classificatori di machine learning che sfruttavano le caratteristiche del MBFA, aggiungendone di nuovo, in 10 pazienti[25]. I risultati ottenuti sono stati migliori del 95% per specificità e sensibilità con differenti classificatori. Questi test, però, sono stati condotti in situazioni di controllo ed, inoltre, la metodologia di validazione sovrastimava le prestazioni delle misure poiché i classificatori erano allenati, iterativamente, con tutte le finestre del segnale disponibili da un paziente escludendone una, la quale veniva usata per ottenere le prestazioni citate. Inoltre, le sequenze di allenamento e di test erano molto simili, il che è molto diverso da normali situazioni. Quindi, ci si aspetta che le riportate specificità e sensibilità calino drasticamente in situazioni non controllate.

Nel 2013, Moore et al. hanno pubblicato il più recente lavoro focalizzato sul MBA. In questo, hanno confrontato differenti configurazioni applicando lo stesso algoritmo in 25 pazienti, dei quali 20 hanno avuto episodi di FoG. Diverse finestre di segnale, posizionamento dei sensori e valori per PTH e FTH sono stati valutati al fine di trovare le condizioni ottimali. I risultati migliori sono stati ottenuti con le finestre di segnale più lunghe, anche se con queste Moore et al. hanno riportato una rilevante perdita di sensibilità negli episodi brevi che, paradossalmente, sono quelli più frequenti nei pazienti affetti da PD[26]. In un test più complesso eseguito precedentemente[27] usando fino a 7 sensori ed un protocollo di test più lungo, sono stati ottenuti una sensibilità e specificità sopra al 70%, anche se, in certe configurazioni (finestra di segnale pari a 7.5s e il sistema installato nella zona lombare), sia

per la sensitività che la specificità hanno raggiunto valori oltre l'80%. In un approccio differente, Tripoli et al. hanno testato diverse configurazioni e locazioni dei sensori al fine di trovare la migliore configurazione[14]. Il lavoro è stato svolto con 5 pazienti ed in condizioni controllate, usando uno specifico protocollo progettato per stimolare il FoG e senza test di FP. In tale lavoro, hanno integrato 2 giroscopi oltre a 6 accelerometri posizionati in posizioni differenti del corpo. Con tutti i sensori indossati, è stata ottenuta un'accuratezza del 96.11%, una specificità del 98.74% e, eseguendo i test su tutti i pazienti tranne uno, una sensitività dell'81.94%. D'altra parte, con una IMU singola nella zona lombare hanno riportato una sensitività del 75% ed una specificità del 95%, anche se l'algoritmo non è stato confrontato con nessun altro metodo usato sotto le stesse condizioni.

In un più recente studio dal gruppo di Mazilu, hanno investigato un approccio di apprendimento non supervisionato per costruire un input ottimale per un classificatore ad albero di decisione con il dataset del progetto DAPHNET (10 pazienti PD). Il loro approccio è stato comparato ad un analogo basato su MBFA nel quale il FI e l'energia della banda spettrale tra 0.5 Hz e 8 Hz sono state valutate. L'allenamento ed i test erano dipendenti dall'utilizzatore e sotto condizioni controllate. I risultati superano l'approccio MBFA simile dell'8.1% in termini di punteggio dell'F1. Un altro approccio è stato presentato da Rodriguez et al., i quali hanno proposto un metodo per contestualizzare gli episodi di FoG tramite un algoritmo di riconoscimento dell'attività, il quale rifiutava i FP quando il paziente era seduto o eseguiva attività quali disegnare o digitare in un laptop. La specificità è stata aumentata in media del 5%, arrivando anche ad un 11.9% in certi casi[10]. Il metodo che aggiungeva la contestualizzazione, però, non ha contribuito a migliorare la sensitività. Altri studi hanno la variabilità della camminata tra un episodio di FoG e condizioni normal. Anche se i risultati sono interessanti, hanno fallito nell'includere i falsi positivi ed un'affidabile classificazione non è stata eseguita[8, 3]. Un paper recente di Zach et al. presenta una nuova metodologia per suscitare FoG in condizioni di laboratorio controllate, le quali sono state valutate con l'algoritmo MBFA ottenendo una sensitività del 75% ed una specificità del 76%[16].

Infine, Alrichs et al., all'interno del progetto REMARK[23], usano una Support Vector Machines (SVM) per rilevare episodi di FoG in 8 pazienti con PD in ambiente casalingo. Il metodo include test in differenti test motori usando un singolo accelerometro nella zona lombare, raggiungendo un'accuratezza del 90%. La specificità, però, è stata calcolata solo non pazienti non FoG, il che può portare a predizioni non affidabili in quanto il modello non è stato testato con pazienti PD con FoG, i quali hanno movimenti molto diversi dai pazienti che non soffrono di FoG. Inoltre, la valutazione è stata eseguita su finestre di un minuto, tempo che è considerato troppo lungo per un'implementazione online[7]. Sempre all'interno del progetto REMARK, Rodriguez et al. hanno presentato un lavoro che utilizza un algoritmo per

rilevare FoG tramite un approccio di machine learning basato su SVM ed un singolo accelerometro a 3 assi indossato nella zona lombare[11]. Il metodo è stato valutato su 21 pazienti affetti da PD in ambienti casalinghi sotto due condizioni: un modello generico testato su tutti i pazienti tranne uno ed un secondo modello personalizzato sull'utente che usa parte del dataset del paziente stesso. I risultati mostrano un significativo vantaggio del modello personalizzato rispetto a quello generico, portando ad un miglioramento in media, sia della sensibilità che della specificità, del 7.2%. Inoltre, l'approccio adottato è stato comparato con i metodi più utilizzati per la detenzione del FoG basati sull'algoritmo MBFA. I risultati del metodo generico mostrano un miglioramento in media dell'11.2% rispetto a metodi MBFA generici, mentre quello personalizzato porta ad un miglioramento del 10% rispetto ad altri metodi specifici sul paziente.

Background

3.1 Freezing Of Gait

Tra i sintomi della malattia di Parkinson, il Freezing può sicuramente essere considerato uno dei più debilitanti. Il Freezing nella malattia di Parkinson, detto anche acinesia paradossa, congelamento o semplicemente blocco motorio, è un'improvvisa, temporanea e involontaria incapacità di iniziare un movimento. È un disturbo che insorge nel corso dell'evoluzione della malattia di cui costituisce un sintomo indipendente e generalmente resistente al trattamento con levodopa. Tale fenomeno si può verificare in ogni momento e i pazienti che lo sperimentano affermano che: *«è come se i piedi rimanessero, per qualche istante, incollati al suolo con la conseguente impossibilità di eseguire il passo successivo»*. In realtà, il Freezing si può verificare anche durante azioni differenti dal cammino come ad esempio l'alzarsi da una sedia o il raccogliere un oggetto. Alcune persone sono più predisposte di altre a subire episodi di congelamento. Tali episodi, si possono verificare sia quando il soggetto è in astinenza da farmaci dopaminergici, in questo caso si parla di "Freezing off", sia quando il soggetto sta assumendo i farmaci, "Freezing on".

È stato dimostrato che il fenomeno del Freezing nella malattia di Parkinson è spesso collegato alle frequenti cadute a cui i soggetti malati vanno incontro. Le cadute nel Parkinson si verificano più spesso quando il soggetto si gira o cambia direzione e sono frequentemente legate a diversi episodi di Freezing. Non tutti i malati di Parkinson subiscono il fenomeno del Freezing, ma si pensa che coloro che lo provano abbiano una più alta probabilità di cadere a terra. L'imprevedibilità del Freezing, accompagnata dallo sforzo inutile a cui il soggetto si sottopone per cercare di muoversi in avanti, possono causare perdita di equilibrio e quindi cadute.

Nel tentativo di superare questo stato di forzata immobilità, i pazienti, talora con un aiuto esterno, cercano di mettere in atto adeguate strategie che si avvalgono di stimoli sensoriali di diversa natura (tattili, visivi oppure uditivi e verbali). Alcune tecniche di tipo motorio o sensoriale possono aiutare i pazienti a convivere con il problema del Freezing. Ad esempio, un pazien-

te incapace di iniziare il primo passo potrebbe riuscire a superare il blocco motorio adottando una delle seguenti strategie:

- fare un passo in direzione di un bersaglio;
- fare un passo per superare un bastone posto sul pavimento;
- fare il primo passo marciando come un soldato.

L'idea che sta alla base di tali stratagemmi è mettere in atto un programma motorio volontario che sostituisca il programma motorio automatico malfunzionante nei malati di malattia di Parkinson. Episodi frequenti di Freezing possono avere pericolose conseguenze sia sullo stato fisico sia su quello psicologico del malato e compromettono ampiamente la qualità della vita di chi ne soffre privandolo spesso dalla propria indipendenza.

3.1.1 Le tipologie di Fog

Il FoG è un episodio transitorio che usualmente dura pochi secondi e di cui ancora non si conosce la patofisiologia, ossia la causa scatenante, ma è stato dimostrato che esistono più sottotipi di Freezing, che si differenziano per l'evento scatenante il fenomeno:

- necessità del paziente di girare su sè stesso per cambiare direzione (esitazione legata alla svolta);
- attraversamento di spazi stretti, come una porta od un corridoio;
- inizio del movimento di camminata;
- regolazione dei passi in prossimità della destinazione (come ad esempio una sedia su cui sedersi);
- stress, come lo squillo di un telefono o campanello o quando la porta dell'ascensore si apre.

Come la malattia progredisce, però, il FoG può apparire spontaneamente anche in uno spazio aperto, evidenziando così l'aspetto imprevedibile di questo fenomeno. Inoltre, fonti di distrazione, che possono distogliere l'attenzione del soggetto dal cammino, o il compimento contemporaneo di più azioni (dual-tasking) possono aumentare la probabilità che si verifichi un episodio di Freezing.

Dallo studio di Schaafsma et al.[12] emerge che gli episodi di Freezing possono anche essere suddivisi in ulteriori tre sottotipi andando ad osservare i movimenti delle gambe dei pazienti e applicando la classificazione di Thompson e Marsden (1995):

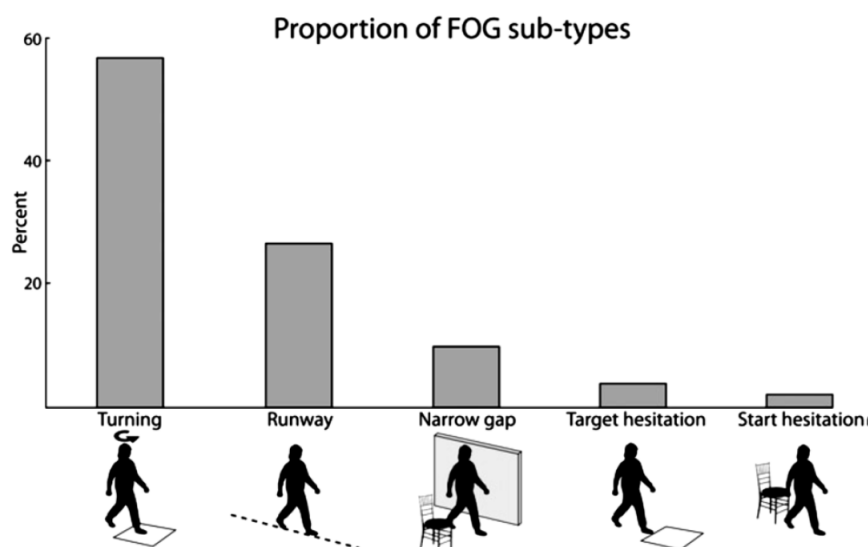


Figura 3.1: Diversi tipi di Freezing esistenti e loro percentuale di incidenza su un certo gruppo di malati di Parkinson. [Fonte: J.M. Shine et al., 2012]

1. FoG associato a passi molto piccoli e strascicati con il minimo movimento in avanti (trascinamento con piccoli passi);
2. FoG caratterizzato da tremore alle gambe, ma nessun movimento in avanti efficace (tremore da fermo);
3. FoG caratterizzato da acinesia completa, vale a dire, nessun movimento osservabile delle gambe.

La necessità di dividere il FoG in sottogruppi dipende dal fatto che questi ultimi potrebbero avere origine differente e quindi essere provocati da cause separate.

3.1.2 Influenza del FoG nella camminata

Il Freezing of Gait influenza il pattern del cammino sia all'inizio della deambulazione sia a regime incrementando o diminuendo in modo evidente i valori dei parametri sopra riportati. Ad esempio, nei parkinsoniani che manifestano frequenti fenomeni di Freezing, la variabilità della durata del passo risulta maggiore e la lunghezza del passo minore rispetto alle situazioni in cui il Freezing è assente. Inoltre, la velocità e la lunghezza dei primi tre passi sono significativamente inferiori nei pazienti con malattia di Parkinson e con Freezing rispetto ai soggetti sani. Anche se il Freezing è tipicamente considerato un problema motorio, il fatto che spesso compaia quando il paziente

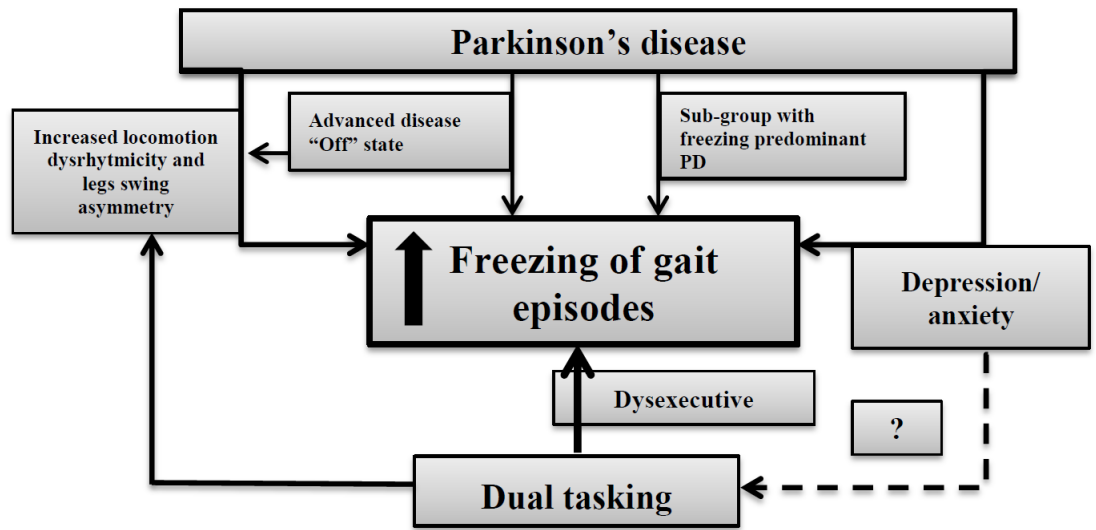


Figura 3.2: Quadro concettuale relativo al Freezing of Gait (FoG) riguardante aspetti mentali e motori.

si trova in spazi ristretti, suggerisce che la percezione dello spazio contribuisce in larga misura a scatenare il sintomo[22]. Inoltre, i pazienti con Freezing hanno velocità media del cammino minore rispetto a pazienti sani e subiscono una riduzione ulteriore della velocità nel momento in cui si trovano a dover attraversare una porta o uno spazio stretto. Depressione e ansia possono comportare un carico cronico sulla salute mentale, e la depressione è associata con i cambiamenti di andatura, tra cui una maggiore variabilità passo-a-passo. Il Dual-Tasking, l'ansia e la depressione possono incrementare la variabilità del passo, la di-sincronizzazione di gamba destra e sinistra e l'asimmetria nei pazienti con Parkinson riducendo così la soglia per il FoG.

3.2 Machine Learning

Il Machine Learning si occupa della realizzazione di sistemi che si basano su osservazioni o esempi come dati per la sintesi di nuova conoscenza (classificazioni, generalizzazioni, riformulazioni). Il Machine Learning (ML) o Apprendimento automatico è una disciplina scientifica che progetta e sviluppa algoritmi che consentono agli elaboratori di evolvere il proprio comportamento basandosi su dati empirici. Il principale obiettivo di ricerca in ambito di Machine Learning è “imparare” a riconoscere automaticamente pattern complessi ed effettuare scelte intelligenti basandosi su dati già analizzati. La necessità di ricorrere al ML nasce dal fatto che prevedere a priori l'intero set

di possibili comportamenti in base all'input, costruendo per esempio manualmente un set di regole, è troppo complesso da descrivere in un linguaggio di programmazione. Parallelamente, la difficoltà di tale metodologia risiede nell'incertezza con cui si individua una corrispondenza tra input e output: essa si basa su un meccanismo parametrico per la generazione dei dati, di cui però non si conoscono a priori valori esatti dei parametri.

Caratteristica del ML è l'induzione, ossia l'estrazione di leggi generali a partire da un insieme di dati osservati. Essa si contrappone alla deduzione in cui, a partire da leggi generali, si prevede il valore di un insieme di variabili. L'induzione parte dall'osservazione per misurare un insieme di variabili e per poi effettuare previsioni su ulteriori dati. Questo processo complessivo nel quale, a partire da un insieme di osservazioni, si vuole effettuare previsioni su nuovi dati prende il nome di inferenza.

Gli algoritmi di apprendimento automatico sono tradizionalmente divisi in tre principali tipologie:

- **Apprendimento supervisionato:** quando l'utente fornisce esempi (e controesempi) di quello che si deve apprendere. E' il problema più studiato nel machine learning. Esso si pone l'obiettivo di prevedere, dato un elemento di cui si conoscono un insieme di parametri (features), il valore di un diverso parametro di output relativo all'elemento stesso;
- **Apprendimento non supervisionato:** parte da osservazioni non preclassificate;
- **Apprendimento con rinforzo:** tecnica di programmazione che si basa sul presupposto che l'algoritmo possa ricevere stimoli dall'esterno a seconda delle scelte fatte.

Il problema del ML è definito a partire da un universo di elementi: ciascun elemento x è descritto dai valori assunti da un insieme di features considerate come input del problema. Ad ogni x è associato un valore y di output (o target). A partire dalla conoscenza di un insieme T di elementi (denominato training set) in cui ogni elemento è descritto da una coppia (x_i, y_i) , con x_i = vettore dei valori delle d features $x_{i1}, x_{i2}, \dots, x_{id}$ e y_i = valore di output, si vuole derivare un modello delle relazioni sconosciute tra features e valori di output, che, dato un nuovo elemento x , consenta di predire il corrispondente valore di output y . Lo scopo dell'apprendimento supervisionato è di costruire un **modello di predizioni** basato su evidenze in presenza di incertezze. Un algoritmo di apprendimento supervisionato prende un insieme conosciuto di dati di input e di risposte ai dati (output) ed allena un modello al fine di generare predizioni ragionevoli a nuovi dati di input. Esistono due tecniche per questo approccio:

- **Classificazione:** tecnica per predire risposte discrete, classificando i dati di input in categorie;

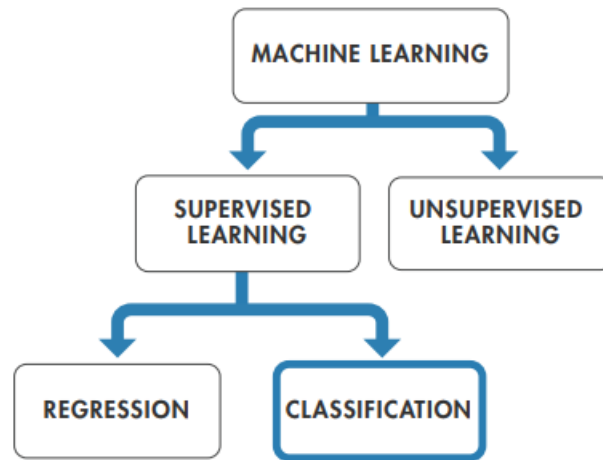


Figura 3.3: Schema delle tipologie di Machine Learning

- **Regression:** tecnica per predire risposte continue.

L'apprendimento non supervisionato, invece, trova pattern nascosti o strutture intrinseche nei dati. Tale tecnica è usata per identificare inferenze da un dataset consistente di dati di input senza classi già definite. Il clustering è l'approccio più diffuso di tale tecnica ed è utilizzato per trovare gruppi nei dati.

Con il ML difficilmente c'è una strada diretta dall'inizio alla fine della progettazione, ma si continua ad iterare e provare idee ed approcci differenti. Per questo è importante definire uno schema di lavoro generale ed evidenziare alcuni punti di decisione chiave lungo il percorso:

1. Accesso e caricamento dei dati: questi possono essere di tutte le forme e i tipi, incompleti o mescolati;
2. Pre-processamento dei dati: applicazione di filtri o ri-campionamento;
3. Derivazione di feature: trovare caratteristiche peculiari a partire dai dati;
4. Allenamento del modello di ML: può essere un procedimento lungo, poichè dipende da molti parametri;
5. Integrazione del modello in un sistema di produzione.

La scelta della modalità supervised o unsupervised si basa sui vantaggi e svantaggi di entrambe: la modalità supervised riesce a predire la giusta classe per le istanze appartenenti al test set ma richiede una consistente quantità di istanze annotate e questo può rappresentare un processo costoso se effettuato manualmente. La modalità unsupervised tipica del clustering, invece,

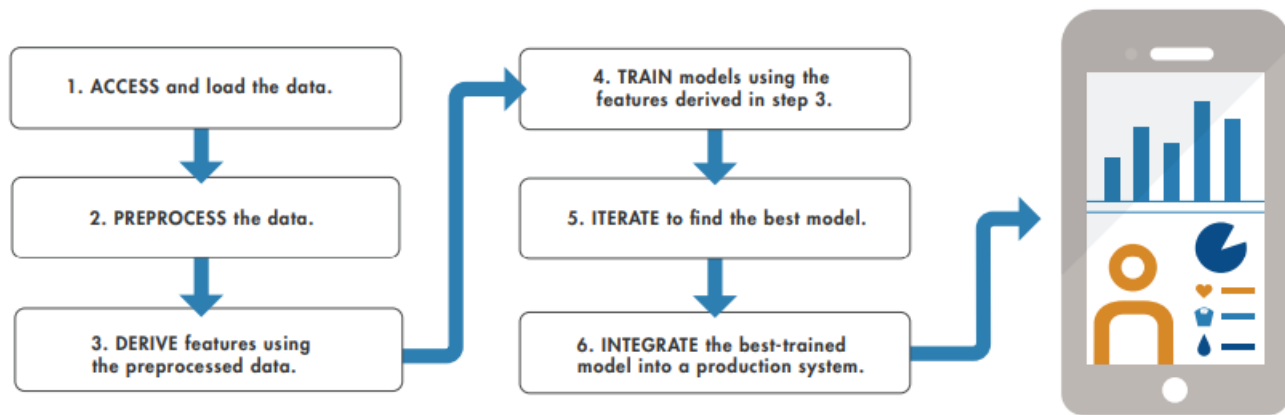


Figura 3.4: Esempio di Workflow tramite Machine Learning

ha il vantaggio di non richiedere un training già annotato (situazione particolarmente frequente quando si ricorre al ML) ma difficilmente etichetta correttamente il cluster e ottiene una precisione più scarsa rispetto al primo metodo nell'associare le istanze ai cluster corretti.

3.2.1 Cluster Analysis

Con il termine Cluster Analysis, o analisi dei gruppi, si intendono le procedure che permettono di individuare, all'interno di un insieme di oggetti di qualsiasi natura, alcuni sottinsiemi, **clusters** appunto, mutuamente esclusivi e tendenzialmente omogenei al loro interno. Le tecniche di Cluster Analysis creano i gruppi in modo tale che ogni osservazione sia molto simile a tutte le altre che appartengono allo stesso gruppo, in funzione di alcuni criteri prestabiliti. Alla fine del procedimento, i cluster finali dovrebbero esibire un'alta omogeneità interna (intra-cluster) ed un'alta eterogeneità esterna (inter-cluster), quindi gli oggetti all'interno dei cluster saranno vicini tra loro, mentre gli oggetti che appartengono a differenti cluster saranno più lontani tra loro.

La cluster analysis rientra tra le tecniche di tipo esplorativo e pertanto non è necessaria alcuna assunzione a priori, anche se impone una serie di decisioni durante l'analisi:

- Scelta delle variabili;
- Criteri di similarità o distanza;
- Tecniche di aggregazione;
- Numero dei gruppi da ottenere;

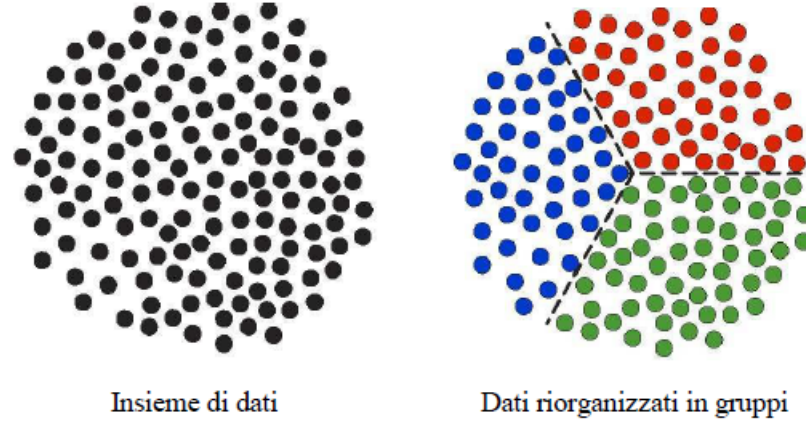


Figura 3.5: Rappresentazione dei dati e dei gruppi ottenuti con la cluster analysis

- Valutazione della qualità della soluzione;
- Scelta fra le diverse possibili soluzioni.

Per classificare e raggruppare gli elementi in gruppi omogenei, è necessario introdurre una nozione di prossimità o similarità. Un indice di prossimità tra due generici elementi x_i e x_j è definito come una funzione dei rispettivi vettori riga nella matrice dei dati: $IP_{ij} = f(x'_i, x'_j)$, $i, j = 1, 2, \dots, n$. Due individui sono vicini quando la loro dissimilarità o distanza è piccola o, equivalentemente, quando la loro similarità è grande. Le principali misure di similarità sono illustrate in tabella 3.6.

Tipo di dato	Misura	Formula
Binario	Coefficiente di Sokal	$S_{ij} = (a + d)/(a + b + c + d)$
Binario	Coefficiente di Jaccard	$S_{ij} = a/(a + b + c)$
Categorici non binari	Media delle variabili	$S_{ij} = (1/p) \sum_{k=1}^p s_{ijk}$
Dati Continui	Distanza Euclidea	$d_{ij} = [\sum_{k=1}^p (x_{ik} - x_{jk})^2]^{1/2}$
Dati Continui	Distanza City Block	$d_{ij} = \sum_{k=1}^p x_{ik} - x_{jk} $

Figura 3.6: Misure di Similarità

I principali algoritmi di clustering sono:

- **K-means:** l'obiettivo che l'algoritmo si prepone è di minimizzare la varianza totale intra-cluster. Ogni cluster viene identificato mediante un centroide o punto medio. L'algoritmo segue una procedura iterativa. Inizialmente crea K partizioni e assegna ad ogni partizione i

punti d'ingresso o casualmente o usando alcune informazioni euristiche. Quindi calcola il centroide di ogni gruppo. Costruisce quindi una nuova partizione associando ogni punto d'ingresso al cluster il cui centroide è più vicino ad esso. Quindi vengono ricalcolati i centroidi per i nuovi cluster e così via, finché l'algoritmo non converge;

- **K-medoids**: è molto simile al k-means, con la sola condizione che il centroide deve corrispondere ad un elemento;
- **Gerarchico**: produce insiemi nidificati di clusters analizzando la similarità tra coppie di oggetti e raggruppando gli elementi in un albero binario gerarchico. Ne esistono di due tipi: agglomerativo, ossia ogni elemento è un cluster e si uniscono tra di loro fino ad arrivare ad unico cluster, e divisivo, nel quale, all'inizio, ho un unico grande cluster che poi viene diviso;
- **Self-Organizing Map**: sono reti neurali a connessioni laterali dove i neuroni di uscita sono organizzati in griglie di bassa dimensione (2D o 3D). Ogni ingresso è connesso a tutti i neuroni di uscita e ad ogni neurone è associato un vettore dei pesi della stessa dimensione dei vettori d'ingresso. La dimensione del vettore d'ingresso è generalmente molto più alta della dimensione della griglia di uscita;
- **Fuzzy C-means**: molto simile anch'esso al K-means, con la sola condizione che un elemento può appartenere a più di un cluster contemporaneamente;
- **Gaussian Mixture Models**: è un modello probabilistico utilizzato per modellare distribuzioni complesse di dati: l'idea base del modello è quella di suddividere un generico dataset di dati (population) in diverse partizioni (subpopulation) e rappresentare queste ultime attraverso una funzione di densità di probabilità. L'obiettivo è quindi rappresentare un segnale complesso attraverso una combinazione lineare di semplici distribuzioni di probabilità (Mixture Model) utilizzando, come si può intuire dal nome stesso del modello, un insieme di gaussiane per la modellazione di distribuzioni complesse di dati.

3.2.2 Classificazione

I principali approcci di classificazione sono due. In un modello parametrico, il modello stesso è preventivamente caratterizzato da un vettore X di parametri: si suppone quindi che esista una relazione tra features e input e che tale relazione sia rappresentabile all'interno di una famiglia di relazioni parametriche rispetto a X che costituisce un modello; in altre parole, un'assegnazione di valori al vettore X definisce una specifica relazione della

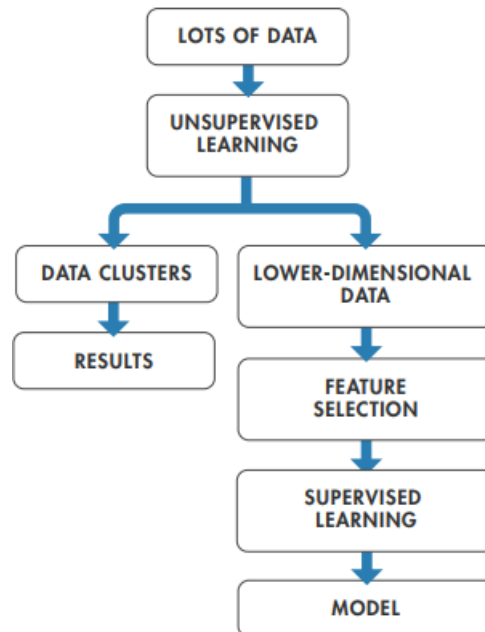


Figura 3.7: Rappresentazione dei possibili approcci usando strategie di clustering

famiglia. Gli elementi nel training set sono utilizzati proprio per derivare tale assegnazione di valori ai parametri (o una distribuzione di probabilità per tali valori), dopo di che non sono più utilizzati. In un modello non parametrico, invece, il numero di parametri cresce con la dimensione del training set: sostanzialmente, ogni singola previsione, in questo caso, richiede l'utilizzo dell'intero training set. Un esempio di approccio non parametrico sono i classificatori di tipo nearest neighbor, in cui si determina l'elemento x_i del training set più vicino a x e si impone il valore della nuova previsione y associata a x uguale al valore y_i dell'elemento x_i .

I principali algoritmi di classificazione sono:

- **Decisione Tree:** si tratta di un classificatore con struttura ad albero, in cui ogni nodo può essere o foglia o nodo interno: se foglia, indica il valore della classe assegnata all'istanza; se nodo interno, specifica il test effettuato su un attributo. Per ciascun valore assunto da un attributo in un test, l'algoritmo crea un ramo ed il relativo sottoalbero. Il focus principale dell'algoritmo di crescita del decision tree è come scegliere gli attributi da testare in ciascun nodo interno dell'albero. L'obiettivo è selezionare gli attributi più utili per classificare le istanze di training attraverso una strategia top down, che consiste in una ricerca greedy degli attributi senza tornare a riconsiderare le precedenti

scelte. Il criterio di split (suddivisione) con cui crea un nuovo nodo si basa sul massimo guadagno di informazione (info gain). In pratica sceglie l'attributo che riesce a dividere "meglio" le istanze appartenenti a classi diverse (detto anche criterio di massima riduzione di incertezza). Quando tutti gli elementi in un nodo hanno la medesima classe, l'algoritmo non procede con ulteriore split (criterio di stopping).

- **K-Nearest-Neighbors:** esso memorizza le istanze del training, poi, basandosi su un criterio di vicinanza, mette in relazione l'istanza da classificare con alcune istanze del training set presenti nello spazio delle feature. In pratica, l'istanza è classificata "a maggioranza" in base alla classe più comune tra le k istanze più vicine del training. Tutto il lavoro è fatto dal classificatore in runtime. Data una nuova istanza x da classificare, il classificatore cerca i k esempi del training che sono più "simili" a x e guarda le loro etichette. Qualsiasi label ricorra più frequentemente tra le k label più vicine è scelta per assegnare la classe a x.
- **Support Vector Machine:** l'idea principale di questo classificatore consiste nel rappresentare gli esempi del training come punti nello spazio mappati in modo tale che punti appartenenti a classi differenti siano separati dal più ampio gap possibile. I punti che mappano il test set saranno assegnati ad una categoria o all'altra in base al lato del gap su cui cadono. Più specificatamente, SVM costruisce un iperpiano ed esegue una buona separazione quando l'iperpiano ha la più ampia distanza dai punti di training più vicini di ciascuna classe. Ci sono molti iperpiani che potrebbero classificare il dato. La miglior scelta è quella di selezionare l'iperpiano che rappresenta la più ampia separazione, o margine, tra due classi, ossia l'iperpiano tale che la distanza tra esso e il punto più vicino su ciascun lato sia massima.

3.3 Riduzione della dimensionalità

3.3.1 Analisi Delle Componenti Principali

L'analisi in componenti principali o **PCA** è una tecnica per la semplificazione dei dati, con lo scopo primario di ridurre un numero più o meno elevato di variabili in alcune caratteristiche latenti. Questo procedimento prende il nome di **feature reduction**. Ciò avviene tramite una trasformazione lineare delle variabili che proietta quelle originarie in un nuovo sistema cartesiano nel quale la nuova variabile con la maggiore varianza viene proiettata sul primo asse, la seconda variabile per dimensione della varianza sul secondo asse e così via. La riduzione della complessità avviene quindi limitandosi ad analizzare le principali (per varianza) tra le nuove variabili. Diversamente

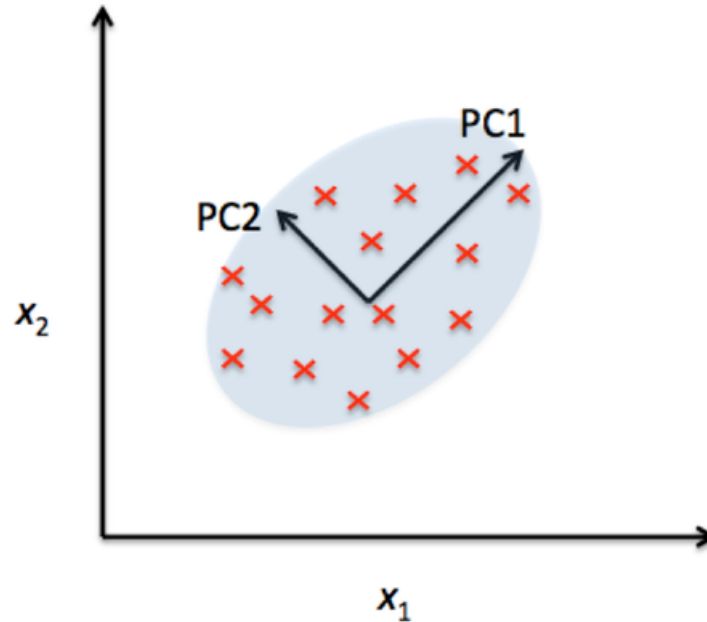


Figura 3.8: Rappresentazione del cambio di dimensionalità tramite PCA

da altre trasformazioni lineari, in questa tecnica sono gli stessi dati che determinano i vettori di trasformazione.

Un metodo per calcolare la componente w_i (ossia quella che effettua la trasformazione per la variabile i) utilizza la **matrice delle covarianze** di x , mentre un altro metodo possibile usa la matrice dei coefficienti di correlazione. Innanzitutto, si devono trovare gli autovalori della matrice di covarianza: Si ottengono così tanti autovalori quante sono le variabili x . L'autovalore con il maggiore valore corrisponde alla dimensione w che ha la maggior varianza: esso sarà dunque la varianza della componente principale numero 1. Per ciascun autovalore viene calcolato il corrispondente autovettore, ossia la matrice dei coefficienti che moltiplicano le vecchie variabili x nella combinazione lineare per l'ottenimento delle nuove variabili w . La matrice degli autovettori degli autovettori viene definita anche matrice di rotazione V e, eseguendo l'operazione $W = V * X$, dove W è il vettore colonna avente come elementi le nuove variabili w_1, w_2, \dots e X è il vettore colonna avente come elementi le vecchie variabili x_1, x_2, \dots , si possono trovare le coordinate di ciascun punto nel nuovo spazio vettoriale. Alla fine, quindi, si tengono solo le componenti le quali, sommate tra di loro, esprimono una certa varianza (es. 90% della varianza dei dati), mentre le altre vengono ignorate. In questo modo, partendo da n variabili iniziali, posso arrivare a $(n-k)$ nuove variabili, dove k è il numero di componenti che non mi servono per raggiungere la soglia di

varianza prefissata.

3.3.2 Analisi dei Discriminanti Lineari

Mentre la PCA è utile per la rappresentazione dei pattern (es. riconoscimento di immagini), l'analisi dei discriminanti lineari o **LDA** viene usata per discriminare tali pattern, ossia per trovare delle misure che mi permettano di dividere in più classi i miei dati. Entrambe vengono usate per la riduzione della dimensionalità delle variabili di input).

L'obiettivo della LDA è quello di trovare un vettore w , di trasformazione, tale per cui classi differenti siano ben separate, mentre la diffusione di ogni classe sia ridotta il più possibile. In pratica, si tratta di trovare una soluzione al cosiddetto criterio di Fisher:

$$J_F = \frac{w^T S_B w}{w^T S_w w} \quad (3.1)$$

, nella quale S_B ed S_W sono rispettivamente la matrice di dispersione tra le classi e la matrice di dispersione all'interno della classe. Nel caso di due classi, la S_B si calcola come:

$$S_B = (m_1 - m_2)(m_1 - m_2)^T \quad (3.2)$$

, dove m_1 rappresenta la media della classe 1 e m_2 quella della classe 2. S_w , invece, si calcola come

$$S_W = \sum_{i=1}^c \sum_{x \in w_i} (x - m_i)(x - m_i)^T \quad (3.3)$$

con m_i = media della classe i e c = numero di classi.

La soluzione del criterio di Fisher viene chiamata anche **Problema dell'Autovalore Generalizzato** e viene rappresentata da tale equazione:

$$S_B w = \lambda S_w w \quad (3.4)$$

e, se S_w è invertibile, diventa

$$S_w^{-1} S_B w = \lambda w \quad (3.5)$$

, corrispondente al problema degli autovalori regolari che coinvolge $S_w^{-1} S_B$. Una volta che w viene trovato, le feature cercate vengono calcolate:

$$y = w^T x \quad (3.6)$$

le quali possono essere usate per allenare l'algoritmo di classificazione scelto e procedere alla predizione su nuovi dati.

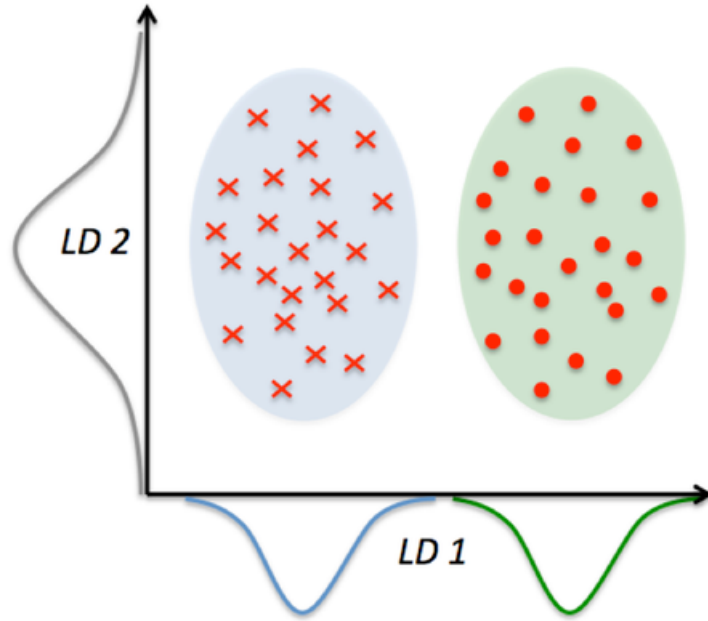


Figura 3.9: Rappresentazione del cambio di dimensionalità tramite LDA

Nel caso si abbiano più di 2 classi, la matrice di dispersione tra le classi, la quali misura la separazione tra le classi, diventa:

$$S_B = \sum_{i=1}^c n_i (m_i - m)(m_i - m)^T \quad (3.7)$$

con n_i = numero di campioni di training appartenenti alla classe i ed m = media di tutti i campioni di training. Ci saranno quindi $C-1$ autovettori, ognuno proveniente da una soluzione di (3.5), che potrebbero non essere ortogonali tra loro ma formano uno sottospazio lineare tale per cui il criterio di Fisher è massimizzato. Questi vengono inseriti in una matrice W e si calcolano le feature come in (3.6).

Motivations and Goals

Chapter 5

Automatic generation of a self-adaptive TLM model

qui ci va la spiegazione del tuo lavoro

Chapter 6

Software Implementation of the proposed methodology

qui invece ci puoi mettere a livello implementativo il tuo lavoro

Case Studies

Conclusions and Future Work

This thesis proposes a methodology

Bibliography

Riferimenti bibliografici

- [1] Delval A, Snijders AH, Weerdesteyn V, Duysens JE, Defebvre L, Giladi N e et al. “Objective detection of subtle freezing of gait episodes in Parkinson’s disease.” In: *Mov Disord.* (2010), 25: 1684–1693.
- [2] Nieuwboer A, De Weerd W, Desloovere K, Fieuws S e Broens kaucsik E. “Abnormalities of the Spatiotemporal Characteristics of Gait at the Onset of Freezing in Parkinson ‘ s Disease.” In: *Mov Disord.* (2001), 16: 1066–1075.
- [3] Tay A, Yen SC, Lee PY, Wang CY, Neo A, Phan SW e et al. “Freezing of Gait (FoG) detection for Parkinson Disease.” In: *10th Asian Control Conference (ASCC)* (2015), pp. 1–6.
- [4] Weiss A, Herman T, Giladi N e Hausdorff JM. “Objective assessment of fall risk in Parkinson’s disease using a body-fixed sensor worn for 3 days.” In: *PLoS One* (2014).
- [5] Takač B, Català A, Rodriguez Martin D, van der Aa N, Chen W e Rauterberg M. “Position and Orientation Tracking in a Ubiquitous Monitoring System for Parkinson Disease Patients With Freezing of Gait Symptom.” In: *JMIR mhealth uhealth* (2013), 1: e14.
- [6] Cole BT, Roy SH e Nawab SH. “Detecting Freezing-of-Gait During Unscripted and UNconstrained Activity”. In: *33rd Annual International Conference of the IEEE EMBS* (2011), pp. 5649–5652.
- [7] Ahlrichs C, Samà A, Lawo M, Cabestany J, Rodriguez-Martin D, Pérez-López C e et al. “Detecting freezing of gait with a tri-axial accelerometer in Parkinson’s disease patients.” In: *Med Biol Eng Comput.* (2016), 54: 223–233.
- [8] Coste C, Sijobert B, Pissard-Gibollet R, Pasquier M, Espiau B e Geny C. “Detection of Freezing of Gait in Parkinson Disease: Preliminary Results.” In: *Sensors* (2014), 14: 6819–6827.

- [9] Pérez-López C, Samà A, Rodríguez-Martin D, Català A, Cabestany J e et al. de Mingo E. “Monitoring Motor Fluctuations in Parkinson’s Disease Using a Waist-Worn Inertial Sensor.” In: *International Work Conference on Artificial Neural Networks Advances in Computational Intelligence Lecture Notes on Computer Science*. (2015), pp. 461–474.
- [10] Rodríguez-Martin D, Samà A, Pérez-López C, Català A, Cabestany J e Rodríguez-Molinero A. “A. Posture Detection with waist-worn Accelerometer : An application to improve Freezing of Gait detection in Parkinson ‘ s disease patients.” In: *European Conference on Ambient Assisted Living*. (2014), pp. 1–6.
- [11] Rodríguez-Martin D, Samà A, Pérez López C, Català A, Moreno ArosteGUI JM, Cabestany J e et al. “Home detection of freezing of gait using support vector machines through a single waist-worn triaxial accelerometer”. In: *PLoS ONE* (2017).
- [12] Schaafsma J. D., Gurevich T. Balash Y and, Bartels A. L., Hausdorff J. M. e Giladi N. “Characterization of freezing of gait subtypes and the response of each to levodopa in Parkinson’s disease”. In: *European Journal of Neurology* (2003), 18: 25–29.
- [13] Jovanov E, Wang E, Verhagen L, Fredrickson M e Fratangelo R. “de-FOG: A real time system for detection and unfreezing of gait of Parkinson’s patients.” In: *31st Annual International Conference of the IEEE EMBS*. (2009), pp. 5151–5154.
- [14] Tripoliti EE, Tzallas AT, Tsipouras MG, Rigas G, Bougia P, Leontiou M e et al. “Automatic detection of freezing of gait events in patients with Parkinson’s disease.” In: *Comput Methods Programs Biomed. Elsevier Ireland Ltd* (2013), 110: 12–26.
- [15] Zabaleta H, Keller T e Fimbel EJ. “Gait analysis in frequency domain for freezing detection in patients with Parkinson’s disease.” In: *Gerontechnology*. (2008).
- [16] Zach H, Janssen AM, Snijders AH, Delval A, Ferraye MU, Auff E e et al. “Identifying freezing of gait in Parkinson’s disease during freezing provoking tasks using waist-mounted accelerometry.” In: *Parkinsonism Relat Disord. Elsevier Ltd* (2015), 21: 1362–1366.
- [17] Han JH, Lee WJ, Tae Beom A, Beom Suk J e Kwang Suk P. “Gait analysis for freezing detection in patients with movement disorder using three dimensional acceleration system.” In: *25th Annual International Conference of the IEEE EMBS*. (2003), pp. 1863–1865.
- [18] Hausdorff JM, Balash Y e Giladi N. “Time series analysis of leg movements during freezing of gait in Parkinson’s disease:akinesia,rhyme or reason?” In: *Physica A*. (2003), 321:656–570.

- [19] Niazmand K, Tonn K, Zhao Y, Fietzek UM, Schroeteler F, Ziegler K e et al. “Freezing of Gait detection in Parkinson’s disease using accelerometer based smart clothes.” In: *IEEE Biomedical Circuits and Systems Conference (BioCAS)*. (2011), pp. 201–204.
- [20] Bachlin M, Roggen D, Plotnik M, Hausdorff JM, Giladi N e Troster G. “Online Detection of Freezing of Gait in Parkinson’s Disease Patients: A Performance Characterization.” In: *Proc 4th Int ICST Conf Body Area, Networks*. (2009).
- [21] Handojoseno a MA, Shine JM, Nguyen TN, Tran Y, Lewis SJG e Nguyen HT. “The detection of Freezing of Gait in Parkinson’s disease patients using EEG signals based on Wavelet decomposition”. In: *34th AnnuInt Conf IEEE Eng Med Biol Soc*. (2012), 2012:69–72.
- [22] Almeida Q.J. e Lebold C. A. “Freezing of Gait in Parkinson’s Disease: A Perceptual Cause for a Motor Impairment?” In: *Neurol Neurosurg Psychiatry* (2010), 81: 513–518.
- [23] REMPARK. “Personal Health Device for the Remote and Autonomous Management of Parkinson’s Disease.” In: *FP7-ICT-2011-7-287677*. (2011-2014).
- [24] Mazilu S, Calatroni A, Gazit E, Mirelman A, Hausdorff J e Troester G. “Prediction of Freezing of Gait in Parkinson’s from Physiological Wearables: An Exploratory Study.” In: *IEEE J Biomed Heal Informatics*. (2015), 6:1–1.
- [25] Mazilu S, Hardegger M, Zhu Z, Roggen D, Troster G, Plotnik M e et al. “Online Detection of Freezing of Gait with Smartphones and Machine Learning Techniques.” In: *Proc 6th Int ICST Conf Pervasive Comput Technol Healthc*. (2012), pp. 123–130.
- [26] Moore ST, Yungher Da, Morris TR, Dilda V, MacDougall HGa, Shine JM e et al. “Autonomous identification of freezing of gait in Parkinson’s disease from lower-body segmental accelerometry”. In: *Neuroeng Rehabil. Journal of NeuroEngineering and Rehabilitation* (2013).
- [27] Moore ST, MacDougall HG e Ondo WG. “Ambulatory monitoring of freezing of gait in Parkinson’s disease”. In: *J Neurosci Methods* (2008), 167: 340–348.
- [28] Zhao Y, Tonn K, Niazmand K, Fietzek UM, D’Angelo LT, Ceballos-Baumann a. e et al. “Online FOG Identification in Parkinson’s disease with a time-frequency combined Algorithm.” In: *Proc 2012 IEEE-EMBS Int Conf Biomed Heal Informatics*. (2012), pp. 192–195.

Acronyms

FI	Freezing Index
MEMS	MicroElectroMechanical Systems
PSD	Power Spectral Density
FB	Freezing Band
WB	Walking Band
EMG	Elettromiografia
EEG	Elettroencefalogramma
FTH	Freezing Threshold
FP	False Positive
PI	Power Index
PT	Power Threshold
MBFA	Moore-Bächlin FoG Algorithm
PD	Parkinson's disease