

UNIVERSITÀ
DEGLI STUDI
DI PADOVA

DIPARTIMENTO DI INGEGNERIA DELL'INFORMAZIONE
CORSO DI LAUREA MAGISTRALE IN BIOINGEGNERIA

Metodologie per la predizione delle ipoglicemie post-bariatriche basate su feature estratte da dati di monitoraggio in continua del glucosio

Relatore:

Prof. Andrea Facchinetti

Laureando:

Federico Del Pup

Correlatore:

Ph.D. Giacomo Cappon

M.Sc. Francesco Prendin

28 febbraio 2022

Anno accademico 2021-2022

Abstract

L'obesità è una patologia caratterizzata da un eccessivo accumulo di grasso corporeo tale da compromettere lo stato di salute e la qualità della vita. Coloro che ne soffrono hanno un elevato rischio di sviluppare svariate patologie, ad esempio: endocrine (e.g. diabete); cardiovascolari (e.g. ischemia miocardica); respiratorie (e.g. apnee notturne); renali; a carico delle articolazioni e dell'apparato digerente. Secondo l'Organizzazione Mondiale della Sanità, nel 2016 sono stati registrati più di 650 milioni di individui adulti affetti da tale patologia, con un'incidenza globale in costante crescita, specialmente negli ultimi anni. In risposta all'aumento dei casi di obesità, si è registrato un incremento di interventi di chirurgia bariatrica, quali il bypass gastrico o gastrectomia verticale, mirati a contrastare l'aumento di peso e ridurre il rischio di mortalità e di comorbidità. Nonostante la loro efficacia nel raggiungimento della remissione a lungo termine sia stata ampiamente dimostrata, il grado di invasività di tali interventi è spesso fonte di numerose complicanze post-operatorie, tra cui l'insorgenza di episodi di ipoglicemia. Alcuni studi riportano che questo fenomeno può interessare fino al 25% della popolazione sottoposta a bypass gastrico o gastrectomia verticale. Attualmente, per coloro che ne sono affetti, la terapia prevede una modifica della dieta alimentare e l'assunzione di farmaci per il controllo della glicemia. Strumenti che possono diventare essenziali nell'identificazione e nel monitoraggio di questi eventi critici di ipoglicemia sono i dispositivi per il monitoraggio in continua del glucosio (continuous glucose monitoring - CGM). Infatti, i sensori CGM, grazie al monitoraggio in tempo reale della glicemia ogni 5 minuti per diversi giorni e a sistemi di generazione di allarmi, consentono di attuare azioni correttive per limitare i rischi legati a questi eventi. Tuttavia, dal punto di vista terapeutico, sarebbe ancor più efficace riuscire ad effettuare una predizione degli eventi ipoglicemici 15-20 minuti prima della loro comparsa, possibile attraverso l'utilizzo di modelli predittivi. In questo modo, sarebbe possibile attuare un'azione preventiva piuttosto che correttiva, limitando o addirittura evitando l'evento ipoglicemico. L'obiettivo di questa tesi è proprio lo sviluppo di nuovi algoritmi predittivi basati sull'uso di feature estratte da dati di sensori CGM per predire gli eventi ipoglicemici nella popolazione post-bariatrica. Lo sviluppo e la valutazione dell'efficacia degli algoritmi, che spaziano dalle Support Vector Machine (SVM) ai logistic model tree, e l'analisi di quali feature sono le più informative per lo scopo del progetto, sono stati effettuati su un dataset di dati CGM acquisiti su individui post-bariatrici che sono stati forniti dall'ospedale universitario di Berna (referente Prof. Lia Bally) nell'ambito di un progetto in essere tra il gruppo di bioingegneria del Dipartimento di Ingegneria

dell'informazione e l'ospedale stesso (progetto SID-Networking 2021 DVTDSS). I risultati ottenuti evidenziano come la predizione in tempo reale di tali eventi sia possibile e come i risultati siano migliori di quelli fino ad oggi pubblicati, anche se ancora subottimali, chiamando ulteriori investigazioni. L'elaborato che ne segue riporta quindi le fasi che partono dalla presentazione del problema per arrivare all'analisi dei risultati. Nel primo capitolo verranno sinteticamente presentate procedure e complicanze della chirurgia bariatrica, focalizzandosi sul problema delle ipoglicemie post-bariatriche e sulle strategie proposte per gestirlo. Nel secondo capitolo verrà presentato il dataset fornito e le varie operazioni di pre-processing a cui è stato sottoposto. L'insieme delle metriche considerate verrà descritto nel capitolo tre, mentre nel quarto verranno riportate procedure e risultati per l'identificazione di un loro subset utile alla predizione. Nel capitolo cinque verranno descritte le diverse metodologie scelte per lo sviluppo degli algoritmi di predizione. Infine, le prestazioni degli algoritmi proposti verranno riportate nel capitolo sei per poi essere commentate nel capitolo conclusivo.

Indice

1. Le ipoglicemie post-bariatriche e sensori di monitoraggio in continua del glucosio.....	1
1.1 La chirurgia bariatrica	1
1.2 Il bypass gastrico alla Roux-en-Y	3
1.3 Le ipoglicemie post-bariatriche	4
1.4 Terapie attuali	6
1.5 I sistemi di monitoraggio in continua del glucosio.....	7
1.6 Predizione delle ipoglicemie post-bariatriche	9
1.7 Obiettivi e struttura della tesi.....	10
2. Il dataset	13
2.1 Descrizione generale.....	13
2.2 Pre-processing dei dati.....	16
2.3 Definizione di ipoglicemia	18
3. Le feature considerate	19
3.1 Premessa	19
3.2 Feature basate su proprietà statistiche	19
3.2.1 Media	19
3.2.2 Mediana	20
3.2.3 Deviazione standard (SD).....	20
3.2.4 Coefficiente di variazione (CV)	20
3.2.5 Range	21
3.2.6 Range Interquartile (IQR).....	21
3.2.7 J-Index	21
3.3 Feature basate sulle escursioni glicemiche	22
3.3.1 Mean Amplitude of Glycemic Excursions (MAGE).....	23
3.3.2 Excursion Frequency (EF).....	24

3.3.3 Control Variability Grid Analysis (CVGA).....	24
3.3.4 Altri indici di escursione glicemica	25
3.4 Feature temporali	26
3.5 Feature di trasformazione.....	27
3.5.1 Glycemic Risk Assessment Diabetes Equation (GRADE)	27
3.5.2 M-Value	28
3.5.3 Indici di controllo glicemico	29
3.6 Feature di rischio.....	30
3.6.1 Blood Glucose Risk Index (BGRI)	31
3.6.2 Average Daily Risk Range (ADRR).....	32
3.6.3 Rischio Dinamico (DR)	32
3.7 Feature short-time	33
4. L'analisi di correlazione	35
4.1 Processing dei dati per l'analisi di correlazione.....	35
4.2 Gli indici di correlazione.....	37
4.2.1 Coefficiente di correlazione di Pearson	37
4.2.2 Coefficiente di correlazione di Spearman	39
4.3 Risultati dell'analisi	40
5. Metodologie per la predizione delle ipoglicemie	49
5.1 Logistic Regression.....	49
5.2 Support Vector Machines.....	52
5.3 Decision Tree	56
5.4 Random forest	58
5.5 Logistic model trees	59
6. Procedure e risultati della classificazione	61
6.1 Assegnazione delle classi.....	61
6.1.1 Classi PH esatto	62
6.1.2 Classi PH misto	63

6.2 Processing dei dati per la classificazione	63
6.3 Suddivisione in training e test set	65
6.4 Bilanciamento delle classi	66
6.5 Valutazione delle prestazioni degli algoritmi	68
6.6 Risultati.....	70
6.6.1 Logistic Regression	70
6.6.2 Support Vector Machine.....	77
6.6.3 Decision Trees	79
6.6.4 Random Forest.....	80
6.6.5 Logistic Model Tree	82
6.6.6 Sintesi finale	83
7. Conclusioni.....	85
7.1 Ulteriori sviluppi.....	86
Appendice A	89
A.1 Valutazione preliminare di algoritmi di predizione basati su modello	89
A.2 Performance delle metodologie per differenti soglie di ipoglicemia.....	90
Bibliografia.....	105

CAPITOLO 1

LE IPOGLICEMIE POST BARIATRICHE E I SENSORI DI MONITORAGGIO IN CONTINUA DEL GLUCOSIO

Nel presente capitolo introduttivo verrà fornita una panoramica generale della chirurgia bariatrica, descrivendone scopi, procedure e possibili complicazioni. La trattazione si focalizzerà sul bypass gastrico alla Roux-en-Y e sull'insorgenza delle ipoglicemie post-operatorie. Al termine del capitolo verranno illustrati gli obiettivi della tesi, argomentando le scelte stabilite in relazione all'attuale stato dell'arte.

1.1 LA CHIRURGIA BARIATRICA

La chirurgia bariatrica è quella branca della chirurgia che si occupa del trattamento dei pazienti affetti da obesità, il cui obiettivo fondamentale è quello di garantire, tramite la perdita della massa corporea in eccesso, una riduzione del rischio di mortalità e di comorbidità, con conseguente aumento della qualità della vita [1].

Si tratta di procedure generalmente molto invasive, spesso irreversibili, che operano attraverso una modifica importante del tratto gastrointestinale. Per questo vengono eseguite solo dopo accurati controlli e la verifica di alcuni requisiti minimi [2] tra cui:

- Essere maggiorenni;
- Avere un indice di massa corporea (IMC), o Body Mass Index (BMI) maggiore o uguale a 40 kg/m^2 ;
- Avere un IMC maggiore o uguale di 35 kg/m^2 unito a severe patologie associate all'obesità, come il diabete di tipo 2;
- Avere un IMC compreso tra 30 e 35 kg/m^2 unito al diabete di tipo 2 e ad una inadeguata capacità di controllo della glicemia nonostante uno stile di vita sano combinato ad una adeguata terapia medica;
- Avere dimostrato che un cambiamento dello stile di vita e una adeguata terapia medica non sono sufficienti a garantire la perdita di peso necessaria per prevenire o trattare patologie associate all'obesità.

L'insieme delle procedure di chirurgia bariatrica è molto ampio e può essere suddiviso in tre macrocategorie a seconda dell'azione esercitata sul soggetto per garantire la perdita di peso:

- *Interventi gastrorestrittivi*: procedure volte a ridurre la capacità gastrica e conseguentemente a limitare la quantità di cibo introdotta nell'organismo generando nel soggetto un precoce senso di sazietà. La riduzione della capacità gastrica viene perseguita attraverso la separazione della parte alta dello stomaco, in modo da formare una tasca gastrica di piccole dimensioni. Tra gli interventi appartenenti a questa categoria si citano la gastroplastica verticale, la gastrectomia verticale parziale e il bendaggio gastrico;
- *Interventi di malassorbimento*: tipologia di interventi volta a ridurre l'assorbimento del cibo tramite un cambiamento del processo digestivo combinato a una riduzione irreversibile delle dimensioni dello stomaco. Tale cambiamento avviene attraverso il collegamento dello stomaco con la parte terminale dell'intestino tenue. In questo modo il cibo ha la possibilità di percorrere solo una parte dell'intero tratto gastrointestinale, riducendo notevolmente l'assorbimento delle sostanze nutritive. Appartengono a questa categoria le varianti della divisione bilio-pancreatica;
- *Interventi misti*: tipologia di interventi che combinano le componenti gastrorestrittive e di malassorbimento caratteristiche delle categorie precedenti. Sono chiamati anche ad azione funzionale. Appartengono a questa categoria il bypass gastrico alla Roux-en-Y e le sue varianti.

La perdita di peso è il beneficio primario derivante dall'intervento chirurgico. Rispetto ai valori preoperatori, la diminuzione è molto consistente e varia in maniera proporzionale al grado di invasività della terapia eseguita, oscillando tra valori come il 16% (45 percento del peso corporeo in eccesso) per il bendaggio gastrico a valori come il 30% (56 ÷ 60 percento del peso corporeo in eccesso) per il bypass gastrico [3][4]. Il miglioramento della qualità di vita è quindi assicurato dalla riduzione delle comorbidità, specialmente cardiovascolari e metaboliche. Recenti studi hanno inoltre dimostrato come la chirurgia bariatrica possa indurre la remissione del diabete di tipo 2 [5][6][7]. Va detto infine che la chirurgia bariatrica non è una tecnica perfetta e spesso può causare delle complicazioni sia a breve che a lungo termine. Per questo le linee guida prevedono l'esecuzione di numerosi controlli pre e post-operatori, anche psicologici, per massimizzare l'efficacia della terapia e minimizzare l'insorgenza di eventuali complicanze. Tra le principali complicazioni a lungo termine si citano il recupero del peso corporeo, l'insorgenza di deficit nutrizionali e la dumping sindrome. Quest'ultima si

caratterizza per la manifestazione di sintomatologie come nausea, diarrea, vomito e crampi addominali dovute ad un passaggio eccessivamente veloce del bolo dallo stomaco all'intestino tenue.

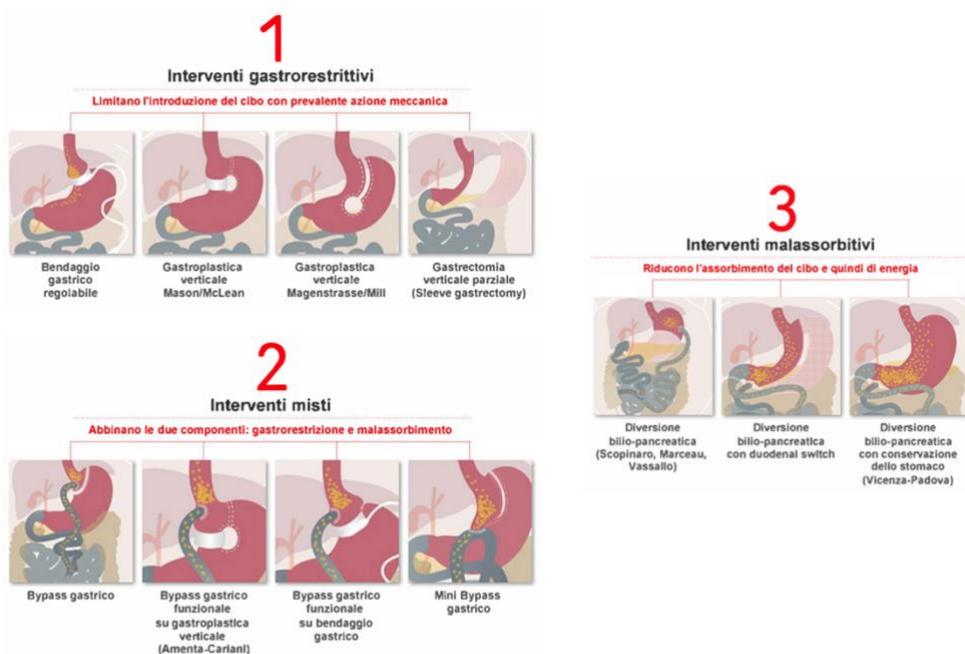


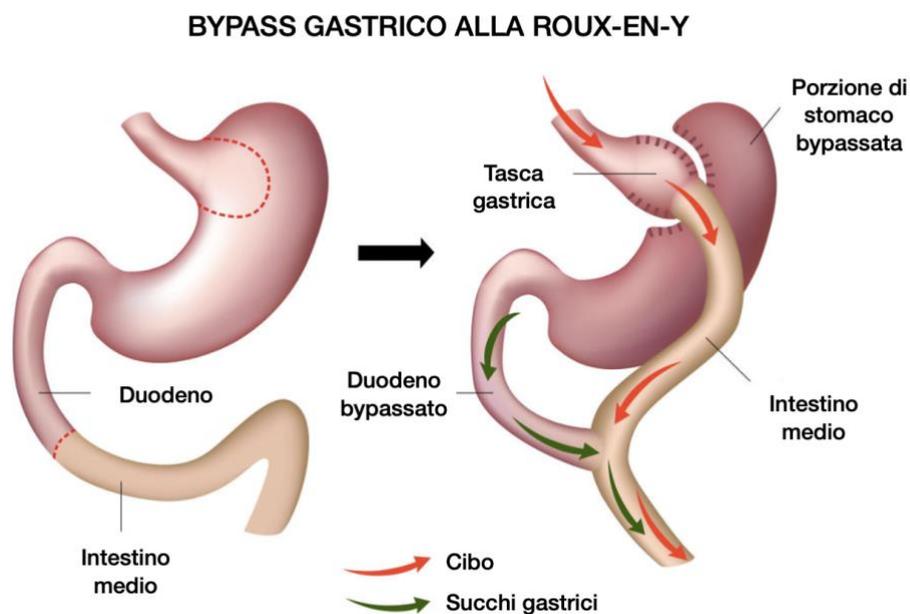
Figura 1.1: Schema degli interventi di chirurgia bariatrica.

1.2 IL BYPASS GASTRICO ALLA ROUX-EN-Y

Il bypass gastrico alla Roux-en-Y è un intervento di chirurgia bariatrica di tipo misto, largamente impiegato da oltre 30 anni in tutto il mondo. L'American Society for Metabolic and Bariatric Surgery ha stimato che solo nell'ultimo decennio, nel solo territorio degli Stati Uniti, sono stati eseguiti oltre 400000 interventi di questo tipo. L'intervento, che può essere eseguito anche per via laparoscopica, consiste nella creazione di una piccola tasca gastrica del volume di circa 30 ml, che viene direttamente collegata all'intestino medio escludendo dal transito del cibo la restante parte dello stomaco e il duodeno.

Oltre alla procedura tradizionale, esistono diverse tipologie di bypass gastrico alla Roux-en-Y, come il RYGB distale, il RYGB prossimale o il mini-bypass gastrico. In sintesi, queste procedure si differenziano per una diversa gestione della componente di malassorbimento tipica di tali interventi, che vede variare la lunghezza del tratto di intestino percorso dal cibo attraverso la modifica del collegamento tra lo stomaco e quest'ultimo.

I benefici, già riportati nel precedente paragrafo, sono molteplici e riguardano principalmente la perdita di peso, tra le più elevate nel complesso delle terapie, e una riduzione del rischio di mortalità e di comorbidità. Tra le possibili complicanze a lungo termine, invece, si rilevano ernie interne, ulcere marginali, ostruzioni intestinali, sindrome di dumping, complicazioni polmonari e deficit nutrizionali [8]. Questi ultimi in particolare, che si manifestano con frequenza maggiore rispetto ad altre tipologie di procedure, coinvolgono vari nutrienti come la vitamina B12, ferro, zinco, proteine.



1.3 LE IPOGLICEMIE POST-BARIATRICHE

Tra le complicanze a lungo termine che possono insorgere a seguito dell'intervento di bypass gastrico (anche a distanza di diversi anni) vi è la manifestazione di episodi di ipoglicemia, da una a tre ore dopo il pasto. La frequenza con cui si manifesta tale complicanza non è definibile, in quanto non esiste una stima realistica della prevalenza di soggetti che ne sono afflitti. Tuttavia, alcuni studi hanno stimato che fino al 25% della popolazione interessata potrebbe soffrire di tale complicanza [9][10].

Per ipoglicemia si intende l'abbassamento dei livelli di concentrazione di glucosio nel sangue al di sotto dei limiti di normalità. Attualmente non esiste una soglia universale per indicare la presenza o meno di un evento di ipoglicemia, quanto piuttosto un insieme di valori distinti per la categoria di soggetti presa in esame. Se, ad esempio, si considerano i soggetti affetti da diabete mellito di tipo 1, il consenso internazionale prevede l'uso di una soglia di 70 mg/dl per le ipoglicemie lievi e 54 mg/dl per quelle severe [11]. Se invece si considerano dei soggetti non diabetici, questa può essere abbassata a valori più bassi nel range dei 55 ÷ 60 mg/dl, sulla base della manifestazione dei primi sintomi legati alla patologia secondo il concetto di ipoglicemia relativa. Anche riguardo alla popolazione post-bariatrica non ci sono delle indicazioni standardizzate in merito alla soglia da utilizzare, anche se alcune pubblicazioni mediche suggeriscono l'uso di una soglia di 60 mg/dl [12]. Le cause che portano alla manifestazione di tale complicanza sono molteplici e ancora non completamente definite. Tuttavia, è noto che tale fenomeno è legato principalmente ad un'alterazione del processo digestivo, con una spiccata variazione dell'attività ormonale associata alla regolazione dei livelli di insulina. In sintesi, escludendo per semplicità i numerosi fattori non insulino-dipendenti che ne contribuiscono alla loro formazione, è possibile sintetizzare il processo di formazione delle ipoglicemie dal momento del pasto secondo quanto segue. Inizialmente il passaggio rapido del cibo dalla tasca gastrica all'intestino porta ad un rapido incremento della concentrazione di glucosio nel plasma fino a valori tipici di eventi di iperglicemia, talvolta molto intensi. In risposta a ciò vi è un innalzamento dell'attività dell'asse entero-insulinare che porta ad una maggiore secrezione e azione dell'ormone GLP-1 [13], che ha la capacità di incrementare la secrezione di insulina agendo sull'attività delle cellule beta nel pancreas. L'innalzamento del livello di insulina porta ad una condizione di iperinsulinemia postprandiale e, di conseguenza, ad un rapido abbassamento della concentrazione di glucosio nel sangue fino a valori al di sotto del range di normalità. In virtù del meccanismo che porta alla loro formazione, le ipoglicemie post-bariatriche sono anche definite ipoglicemie postprandiali iperinsulinemiche, poiché avvengono a distanza di poche ore (1÷3) dal pasto e sono sempre precedute da un innalzamento della concentrazione di glucosio nel sangue.

La sintomatologia indotta dalle ipoglicemie varia in considerazione della severità dell'evento e può causare sia sintomi lievi come sudorazione, tremore, cefalea e difficoltà di concentrazione, che gravi come tachicardia, perdita di conoscenza e, in casi estremi, coma. Per questo è importante trattare prontamente ed efficacemente le ipoglicemie.

1.4 TERAPIE ATTUALI

L'obiettivo principale delle terapie esistenti è quello di ridurre la frequenza e la gravità degli eventi, in modo tale da garantire una ripresa delle normali abitudini di vita. Ciò è reso possibile grazie ad un opportuno bilanciamento tra la terapia medico nutrizionale e la terapia farmacologica:

- *Terapia medico nutrizionale*: è il primo approccio impiegato per la gestione delle ipoglicemie post-bariatriche e prevede un adattamento della dieta sulla base di una corretta gestione dei carboidrati, fonte principale di glucosio. Tra le principali modifiche è previsto un abbandono dei carboidrati semplici in favore di quelli a basso indice glicemico, in quantità non superiore ai 30 g per pasto, e l'introduzione di "pasti misti" caratterizzati dalla combinazione di proteine e grassi per rallentare l'assimilazione dei nutrienti [14];
- *Terapia farmacologica*: viene impiegata in aggiunta alla terapia medico nutrizionale, e prevede l'uso di uno o più farmaci agenti sui meccanismi regolatori dei livelli di glucosio, come l'acarbosio, il diazossido e l'octreotide. Nello specifico, l'acarbosio ritarda e riduce l'assorbimento del glucosio tramite l'inibizione dell'enzima α -glucosidasi, con conseguente diminuzione dell'escursione post-prandiale. Il diazossido riduce la secrezione di insulina tramite inibizione dei canali canali potassio ATP-dipendenti delle cellule β del pancreas mitigando gli effetti dell'iperinsulemia tipica della popolazione post-bariatrica. L'octreotide riduce la secrezione sia di insulina che di GLP-1 attraverso formazione di legami con i sottotipi 2 e 5 dei recettori della somatostatina.

Talvolta l'approccio nutrizionale e farmacologico non producono risultati adeguati. In queste situazioni, soprattutto in presenza di ipoglicemie severe e molto frequenti, può essere necessario ricorrere alla terapia chirurgica. Le procedure sono molto invasive, e per questo vengono eseguite solo quando nessun'altra strada risulta percorribile. Tra le opzioni chirurgiche previste si cita, a titolo di esempio, l'immissione del cibo attraverso un tubo gastrostomico nella regione dello stomaco bypassata o una reversione del RYGB, quando possibile.

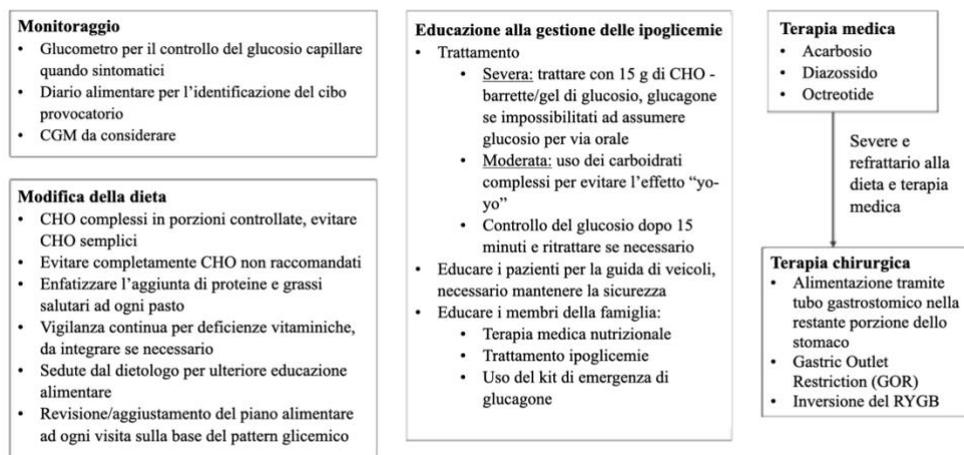


Figura 1.3: Schema delle terapie attuali per le ipoglicemie post-bariatriche (tratto da [12]).

1.5 I SISTEMI DI MONITORAGGIO IN CONTINUA DEL GLUCOSIO

Considerata l'importanza di ridurre la frequenza e la severità delle ipoglicemie nella popolazione sottoposta a bypass gastrico, un ruolo fondamentale può essere giocato dai sistemi di monitoraggio in continua del glucosio (CGM). I sistemi CGM sono dei dispositivi sviluppati per agevolare e migliorare il controllo della glicemia nella popolazione diabetica. Sono in grado di rilevare automaticamente il livello di concentrazione di glucosio nel sangue a intervalli di tempo regolari, tipicamente di 1, 3 o 5 minuti, per diversi giorni.

Sono composti da tre elementi principali:

- un sensore elettrochimico transcutaneo, tipicamente amperometrico, in grado di misurare una corrente proporzionale alla concentrazione di glucosio nel fluido interstiziale tramite un ago inserito nella regione addominale o brachiale;
- un trasmettitore wireless, utilizzato per comunicare i dati rilevati;
- un dispositivo ricevitore, indispensabile per la registrazione e la visualizzazione dei dati raccolti.

Tali dispositivi rappresentano uno strumento molto utile per la gestione delle ipoglicemie grazie all'elevata frequenza di campionamento e alle funzionalità aggiuntive implementate in quelli

attuali, come ad esempio la possibilità di generare allarmi al raggiungimento di determinati valori di concentrazione del glucosio. Restano tuttavia dei meri strumenti di monitoraggio in quanto, sebbene siano in grado di segnalare prontamente un evento di ipoglicemia così da consentire all'utente di prendere delle contromisure per il ristabilimento della glicemia nei valori del range euglicemico, non sono in grado di prevedere la manifestazione di tale evento. Un'interessante prospettiva legata all'uso dei sensori CGM è legata proprio alla possibilità di integrare nel dispositivo ricevitore un algoritmo per la predizione delle ipoglicemie e la generazione di allarmi preventivi, consentendo di prevenire o perlomeno mitigare l'insorgenza di quest'ultime. Ad oggi i dispositivi in commercio non dispongono di tale funzionalità, sebbene per i soggetti diabetici, categoria maggiormente esposta al fenomeno delle ipoglicemie, siano stati proposti numerosi algoritmi di predizione [15].

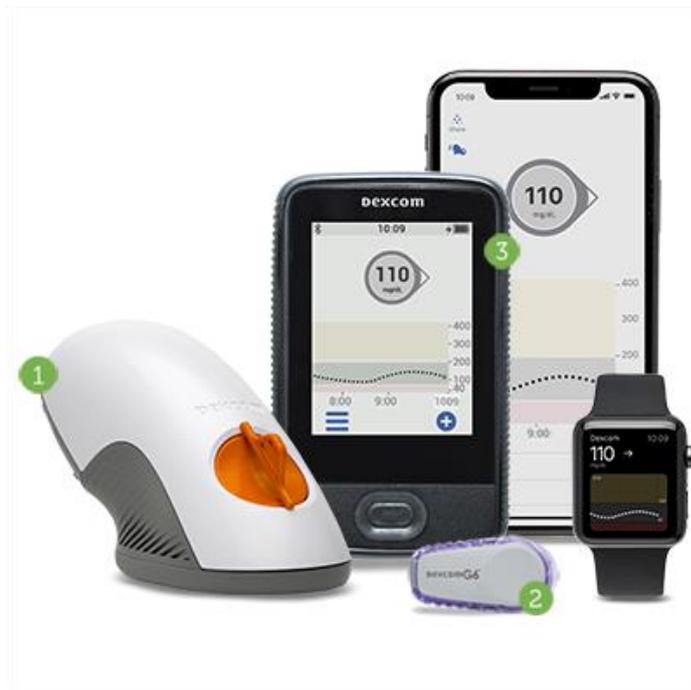


Figura 1.4: Foto di un dispositivo CGM (Dexcom G6®) composto da un applicatore automatico (1), il sensore e il trasmettitore (2) e il dispositivo di visualizzazione (3).

1.6 PREDIZIONE DELLE IPOGLICEMIE POST-BARIATRICHE

Mentre la predizione delle ipoglicemie e dei livelli di glucosio è stata ampiamente investigata nell'ambito del diabete di tipo 1 (T1D), questo argomento è stato solo parzialmente affrontato nei soggetti post-bariatrici. Infatti, finora è stato proposto un solo algoritmo da Laguna Sanz et al [16]. L'algoritmo, il cui diagramma di flusso è riportato nella figura (1.5), prevede l'azione coordinata di due blocchi predittivi per la generazione degli allarmi, i quali attivano un sistema di infusione di glucagone controllato da un tablet computer. I due blocchi predittivi, elementi di maggiore interesse per gli obiettivi della tesi, sono così caratterizzati:

- *Blocco LGP (Low Glucose Predictor)*: viene eseguito ad ogni nuova acquisizione. Sfrutta l'attuale livello di glucosio e la sua derivata temporale per valutare, tramite predizione lineare, il tempo necessario al raggiungimento della soglia ipoglicemica, così da lanciare un allarme nel caso in cui esso sia inferiore ad un valore prefissato. A differenza del modulo LGP originale da cui gli autori si sono ispirati [17], l'allarme viene lanciato solo se due campioni consecutivi soddisfano tale requisito;
- *Blocco PBH (Post-Bariatric Hypoglycemia)*: introdotto dagli autori vista la particolare natura delle ipoglicemie post-bariatriche. Viene eseguito solo dopo il rilevamento di un picco iperglicemico e rimane attivo al massimo per le 2 ore successive. In maniera analoga al blocco LGP sfrutta l'attuale livello di glucosio e la sua derivata temporale per eseguire una predizione lineare, stimare il tempo di raggiungimento della soglia ipoglicemica, e lanciare un allarme nel caso in cui esso sia minore di due valori predefiniti. A differenza del blocco precedente non è necessario che due campioni successivi soddisfino tale requisito.

Tale sistema è stato validato in un trial clinico e in un ambiente altamente controllato [18], situazione che comporta delle dinamiche molto diverse da quelle che gli stessi soggetti si trovano poi a vivere quotidianamente. Sebbene ci siano diverse parti da poter ottimizzare, specialmente nelle porzioni dedicate ai blocchi predittivi, si è scelto di divergere dall'algoritmo descritto per passare ad un approccio di tipo open-loop con predizione basata su tecniche di machine learning. Le ragioni che hanno portato a tale scelta sono da associarsi principalmente alle performance dell'algoritmo e alle scelte implementative prese dagli autori in merito agli elementi predittivi. Una valutazione preliminare delle prestazioni dell'algoritmo ha infatti evidenziato come esso generi un numero considerevole di falsi allarmi, garantendo una sensibilità molto elevata a scapito di una precisione decisamente bassa. Ciò si può associare al fatto che i blocchi predittivi di fatto utilizzano un modello lineare per la generazione degli

allarmi. Come si vedrà in seguito, le dinamiche dei profili CGM della popolazione in esame sono molto rapide e non consentono l'applicazione di metodologie di predizione basate su modelli lineari, in quanto l'ipotesi di stazionarietà del segnale qui non è rispettata. Per una consultazione delle performance dell'algoritmo e di alcuni semplici modelli lineari applicati sul dataset descritto nel capitolo successivo, si rimanda all'appendice dei risultati. In virtù di tali considerazioni, l'obiettivo è quello di superare le attuali limitazioni andando a lavorare su un approccio completamente nuovo e descritto nel paragrafo conclusivo di questo capitolo.

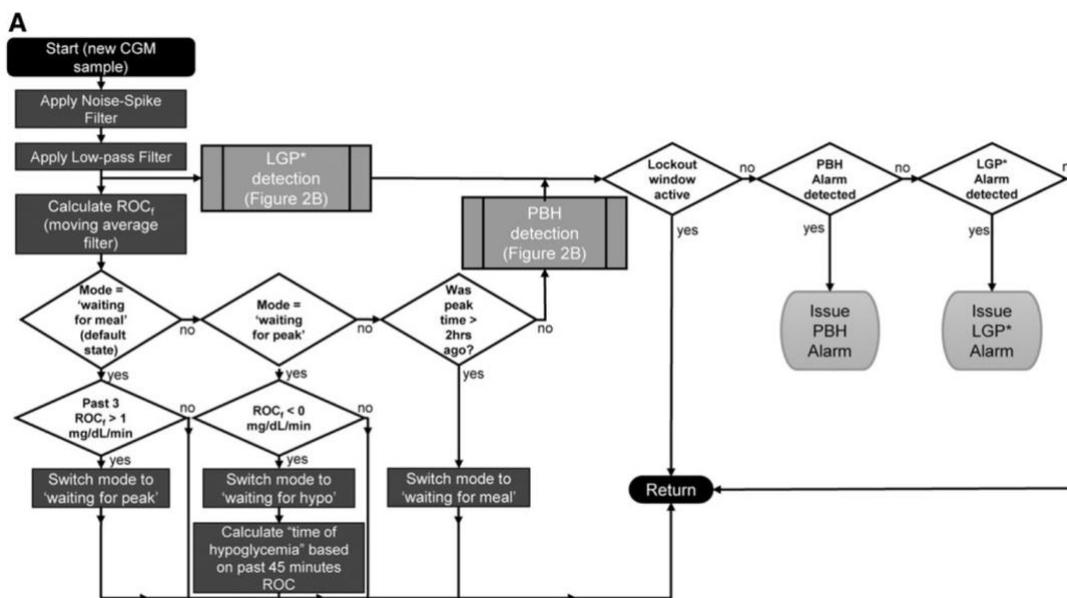


Figura 1.5: Schema dell'algoritmo di predizione di Laguna Sanz et al. (tratto da [15]).

1.7 OBIETTIVI E STRUTTURA DELLA TESI

L'obiettivo del seguente lavoro è quindi quello di sviluppare un algoritmo per la predizione delle ipoglicemie nella popolazione post-bariatrica basato sull'uso di metodologie di machine learning che prevedono l'elaborazione di dati CGM acquisiti durante un contesto di vita reale. L'idea di fondo, rappresentata schematicamente in figura (1.7), è quella di utilizzare le ultime rilevazioni fornite dal sensore CGM per estrarre un insieme di metriche da dare in ingresso ad un classificatore binario addestrato per individuare possibili ipoglicemie future e generare allarmi preventivi.

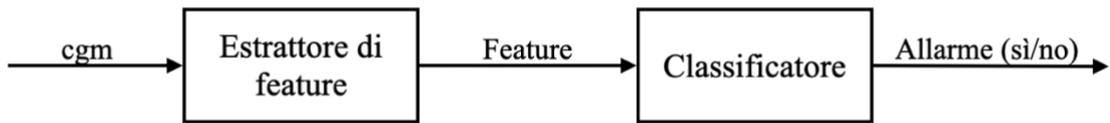


Figura 1.7: Schema semplificato algoritmo di predizione proposto.

Lo sviluppo di tale algoritmo predittivo richiede particolare attenzione in tutti i passaggi del suo sviluppo, in quanto scelte subottime possono inficiare sulle sue prestazioni. Il lavoro prevede la realizzazione e l'implementazione dei seguenti passaggi:

1. definire l'insieme delle feature potenzialmente utili alla predizione dell'evento e capire quali di esse hanno potere predittivo attraverso un'analisi di correlazione;
2. selezionare i possibili classificatori binari, lineari e nonlineari, il cui output sarà 1 in caso di ipoglicemia imminente e 0 viceversa;
3. identificare i classificatori sui dati forniti (training set), prestando attenzione nell'identificazione dei diversi iper-parametri quali l'orizzonte di predizione, cioè quanto temporalmente in avanti si può predire un evento di ipoglicemia, la soglia di ipoglicemia e il numero di valori CGM passati necessari per l'estrazione delle feature;
4. valutare le prestazioni delle metodologie sul test set, andando infine a selezionare la migliore configurazione.

CAPITOLO 2

IL DATASET

2.1 DESCRIZIONE GENERALE

Per lo sviluppo e la valutazione degli algoritmi di predizione delle ipoglicemie sono stati utilizzati dati CGM acquisiti su individui con casi confermati di ipoglicemia post-bariatrica a seguito dell'intervento di chirurgia bariatrica. Nello specifico, i dati provengono da un trial clinico (NCT04330196) diretto dalla Professoressa Lia Bally del dipartimento di diabetologia, endocrinologia, clinica nutrizionale e metabolismo dell'ospedale universitario di Berna (Inselspital), Svizzera, reso disponibile tramite una collaborazione in essere tra il gruppo di bioingegneria del Dipartimento di Ingegneria dell'Informazione e l'ospedale stesso (progetto SID-Networking 2021 DVTDSS). Per questo trial è stata reclutata una popolazione campione sulla base dei seguenti criteri di eleggibilità:

- Età maggiore o uguale ai 18 anni;
- Bypass gastrico con procedura alla Roux-en-Y eseguito almeno un anno prima del reclutamento;
- Presenza confermata di ipoglicemie postprandiali iperinsulinemiche (PHH).

I criteri di esclusione sono stati, invece, i seguenti:

- Cambiamenti rilevanti del peso ($\geq 5\%$) negli ultimi 3 mesi;
- Presenza di diabete basato su valori di HbA1c $\geq 6.5\%$ senza trattamenti ipoglicemizzanti;
- Livelli di emoglobina inferiori 11 g/dl;
- Presenza di patologie a carico di cuore, polmoni, fegato, tratto gastrointestinale oppure reni;
- Trattamenti con farmaci ipoglicemizzanti, anoressizzanti, steroidei o altri farmaci in grado di influenzare la motilità intestinale;
- Gravidanza o allattamento;
- Incapacità di seguire i protocolli;
- Assenza del consenso informato.

Ai soggetti reclutati, in totale 39, è stato fornito il dispositivo di monitoraggio in continua del glucosio Dexcom G6®, in grado di effettuare una acquisizione ogni 5 minuti e di facilitare la raccolta dei dati da parte della struttura ospedaliera tramite i tool implementati nel device. Il monitoraggio è avvenuto in free-living condition, così da permettere sia ai soggetti reclutati un normale svolgimento delle attività lavorative ed extra-lavorative, sia all'ospedale di ottenere acquisizioni quanto più rappresentative delle normali condizioni di vita.

	Rapporto	Media (SD)	Mediana (25 - 75 percentile)
Età	-	41.3 (11.4)	-
Genere (maschile : femminile)	4:35	-	-
IMC [kg/m ²] (37 su 39)	-	27.7 (5.6)	-
Anni Dall'intervento (36 su 39)	-	6.3 (3.7)	-
Giorni di monitoraggio	-	39.2 (75.6)	9.9 (9.3 - 30.2)

Tabella 2.1: Caratteristiche dei soggetti reclutati.

Una prima visione dei tracciati forniti evidenzia come le dinamiche del glucosio nella popolazione sottoposta a bypass gastrico siano molto rapide, specialmente nelle fasce orare tipiche di un pasto, dove l'alterazione dei fattori ormonali descritti nel capitolo precedente causa rapide variazioni della glicemia, con picchi e nadir significativamente al di fuori del range euglicemico [70-180] mg/dl, come mostrato nella figura (2.1).

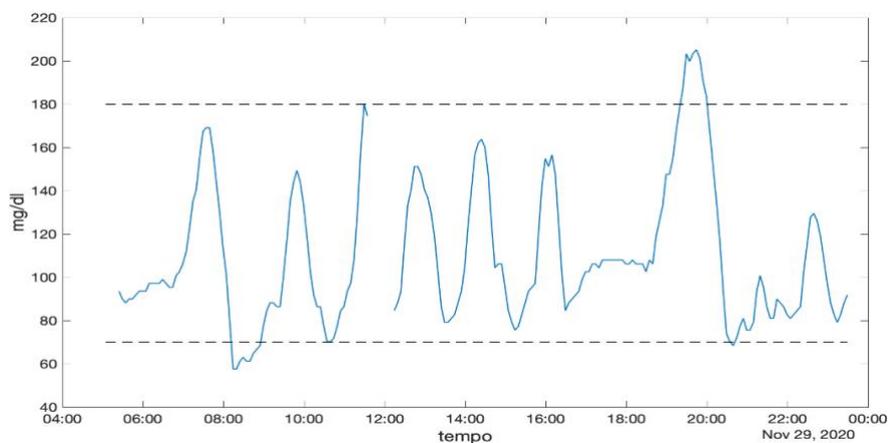


Figura 2.1: Esempio di traccia CGM acquisita.

Sono presenti, inoltre, diversi aspetti negativi legati sia all'alta rumorosità del segnale, derivante da possibili problemi di calibrazione (figura 2.2) e di compressione del sensore, sia alla presenza di numerosi missing value, pari al 13.7 % del tempo totale di acquisizione. Quest'ultimo aspetto rappresenta un problema molto rilevante per lo sviluppo di algoritmi per la predizione delle ipoglicemie dato che da un'analisi preliminare si è notato che la maggior parte dei gap nel segnale CGM è localizzata nei punti di rapida discesa indotta dall'iperinsulinemia postprandiale, che spesso precedono un evento di ipoglicemia. Un possibile motivo per cui sono presenti così tanti missing value è da riscontrarsi nel firmware dei dispositivi CGM: essi sono infatti un prodotto appositamente studiato per il monitoraggio della glicemia nella popolazione diabetica, dove le dinamiche di glucosio sono molto più lente. Perciò, quando vengono rilevate variazioni così rapide tra due acquisizioni ravvicinate, il dispositivo considera l'ultima acquisizione eseguita come non fisiologica, associandola ad un errore di misurazione e bloccandone dunque il suo salvataggio.

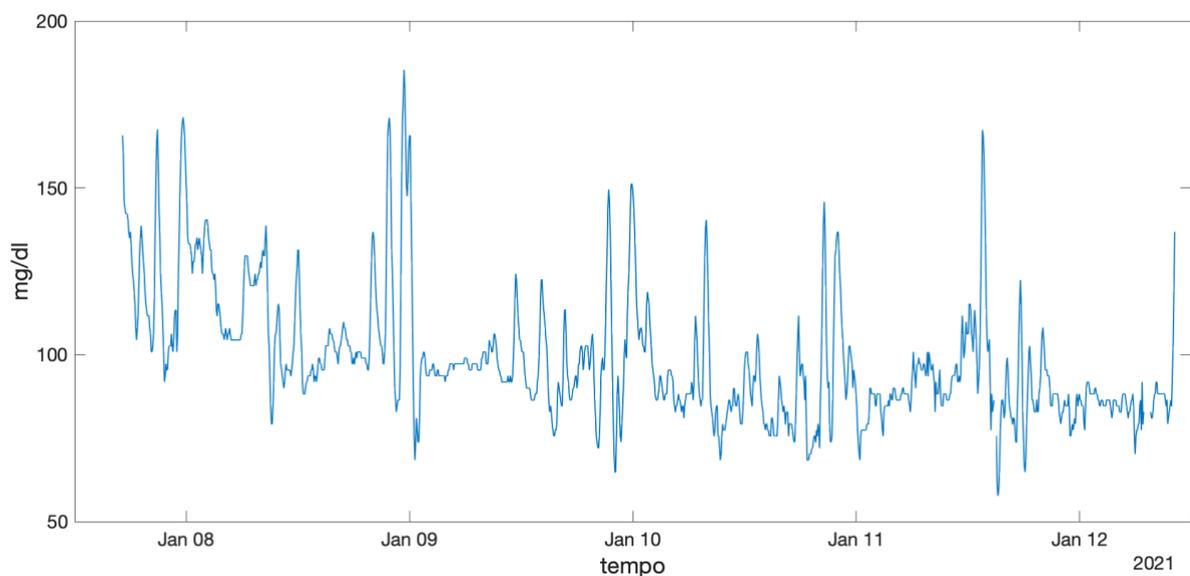


Figura 2.2: Esempio di drift nei tracciati CGM, probabilmente dovuto ad errori di calibrazione del sensore.

2.2 PRE-PROCESSING DEI DATI

In virtù dei problemi riscontrati sui tracciati forniti, si è reso necessario effettuare alcune operazioni di pre-processing dei dati. In primo luogo, per gestire il problema dei missing value, si è deciso di applicare una tecnica di interpolazione denominata sppchip, o shape preserving Piecewise Cubic Hermite Interpolating Polynomial. La sppchip è un metodo di interpolazione composito che prevede l'individuazione di una funzione polinomiale a tratti $P(x)$ che soddisfi le seguenti condizioni:

- In ogni sottointervallo $x_k \leq x \leq x_{k+1}$, la funzione $P(x)$ è un polinomio di interpolazione di Hermite di grado 3 avente specifici valori di derivata nei punti di interpolazione;
- $P(x)$ onora i dati del tracciato originale, per cui vale $P(x_j) = y_j$;
- La derivata prima dP/dx è continua, mentre la derivata seconda può avere salti;
- Il polinomio $P(x)$ è a mantenimento di forma. Le pendenze sono scelte in maniera tale da conservare la forma dei dati e rispettare la monotonicità della funzione. Quindi, negli intervalli in cui i dati sono monotoni anche $P(x)$ lo è, e nei punti in cui i dati hanno estremi locali anche $P(x)$ li avrà.

Il metodo scelto risulta simile all'interpolazione mediante spline cubica, con la differenza che quest'ultima, per preservare la continuità della derivata seconda, non genera un polinomio a mantenimento di forma. Quest'ultimo aspetto è stato decisivo nella scelta della sppchip come metodo di interpolazione, in quanto la spline cubica può enfatizzare eccessivamente le escursioni glicemiche e ricostruire i dati mancanti con valori non fisiologicamente accettabili, soprattutto quando essi sono localizzati nelle porzioni esterne del tracciato.

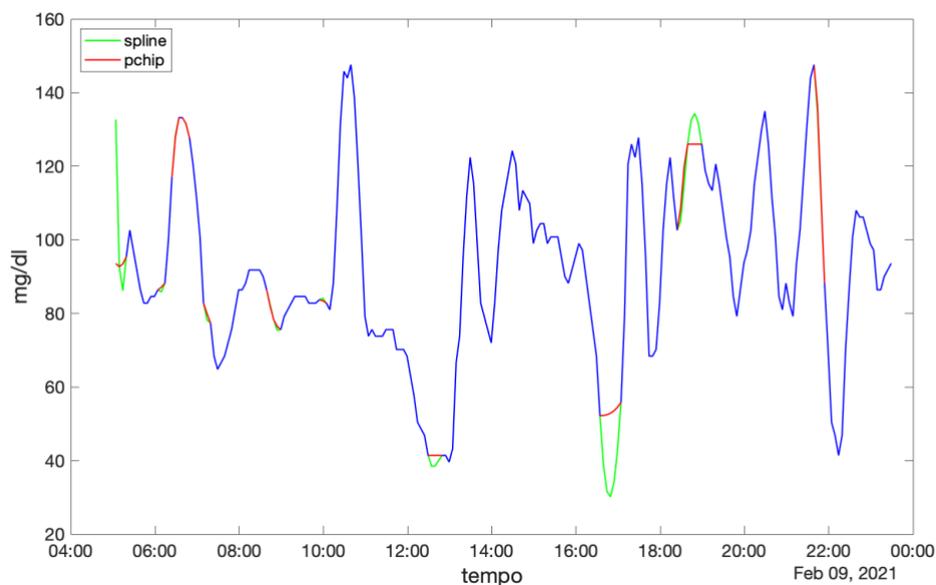


Figura 2.3: Differenza tra interpolazione sppchip (rosso) e spline cubica (verde).

Successivamente alla ricostruzione dei dati mancanti si è proceduto all'eliminazione dei campioni notturni appartenenti alla fascia oraria dalle 23:30 alle 5:00. Questa scelta si basa su due osservazioni pratiche:

1. gli eventi ipoglicemici di interesse si verificano a ridosso dei pasti principali;
2. le ipoglicemie notturne sono spesso associate ad artefatti da compressione. Tali artefatti si manifestano a causa di una pressione prolungata del sensore indotta dalla posizione assunta dal soggetto durante il sonno, che inficia sulla corretta misurazione della concentrazione di glucosio nel fluido interstiziale.

Va precisato che la scelta di escludere i campioni fino alle 5 del mattino e non più tardi deriva dalla necessità di disporre di un numero sufficiente numero di campioni passati da estrarre su cui calcolare il primo set di metriche della giornata. Supponendo, per esempio, di voler utilizzare le rilevazioni degli ultimi 90 minuti per il calcolo delle metriche e di voler predire lo stato glicemico del soggetto 30 minuti in avanti, i campioni dalle 5:00 alle 6:30 servirebbero per estrarre le feature su cui eseguire la prima predizione, da associarsi alla rilevazione delle 7:00. Se invece si eliminassero i campioni CGM fino alle 7:00, la prima predizione fornita sarebbe per le 9:00 e il rischio di perdere la colazione, uno dei tre pasti principali della giornata, sarebbe alto.

2.3 DEFINIZIONE DI IPOGLICEMIA

Prima di passare al prossimo capitolo e alla trattazione delle metriche utilizzate, è utile descrivere come sono state identificate le ipoglicemie nel dataset. Come anticipato nel paragrafo 1.3, attualmente non esistono regole precise per l'identificazione delle ipoglicemie post-bariatriche. Il modus operandi, adottato anche per le altre problematiche del progetto in cui non è stato possibile trovare indicazioni in letteratura, è stato quello di valutare lo stato dell'arte in merito all'argomento nella popolazione diabetica, dove il tema delle ipoglicemie e della loro predizione è stato ampiamente affrontato. Per tale popolazione il consenso internazionale sull'uso dei dispositivi di CGM prevede che un evento di ipoglicemia sia tale solo in caso di soddisfacimento delle seguenti condizioni:

- La soglia di ipoglicemia è posta a 70 mg/dl;
- L'inizio dell'evento presenta almeno 15 minuti consecutivi sotto la soglia ipoglicemica;
- La fine dell'evento è succeduta da almeno 15 minuti consecutivi sopra la soglia ipoglicemica.

Tale definizione, tuttavia, non sembra adattarsi bene alla popolazione post-bariatrica. Considerando le dinamiche molto rapide e l'alta rumorosità del dataset, si è scelto di adeguare le condizioni di identificazione stabilite dal consenso internazionale ponendo i seguenti criteri:

- La soglia di ipoglicemia è posta a 70 mg/dl;
- L'inizio dell'evento presenta almeno 1 campione sotto la soglia ipoglicemica;
- La fine dell'evento è succeduta da almeno 15 minuti consecutivi sopra la soglia ipoglicemica.
- La durata dell'ipoglicemia deve essere maggiore o uguale a 10 minuti.

In questo modo vengono considerati anche gli eventi di durata molto breve, ma vengono esclusi quelli formati da un solo campione sottosoglia, che vengono considerati come errore di misura.

CAPITOLO 3

LE FEATURE CONSIDERATE

3.1 PREMESSA

In questo capitolo verranno descritte tutte le feature inizialmente prese in considerazione per lo sviluppo dell'algoritmo di predizione delle ipoglicemie post-bariatriche. Come per la formulazione dei criteri di identificazione delle ipoglicemie, anche per la scelta delle feature si è fatto riferimento alla vasta letteratura relativa alla popolazione diabetica. Per l'estrazione della maggior parte delle feature considerate, sono state utilizzate porzioni di codice prelevate dal pacchetto AGATA[®], sviluppato e messo a disposizione dal gruppo di ricerca di bioingegneria per ambiente Matlab[®]. Tali porzioni sono state raccolte in un'unica funzione insieme ad altre parti originali per il calcolo delle metriche non presenti nel pacchetto. La trasposizione su un'unica funzione si è voluta per facilitare l'inserimento o la rimozione di feature in fase di progettazione e per alleggerire quanto più possibile il carico computazionale richiesto per il loro calcolo.

3.2 FEATURE BASATE SU PROPRIETÀ STATISTICHE

Si tratta di feature utilizzate per sintetizzare le proprietà statistiche della porzione di traccia CGM. Richiedono l'uso di tutte le rilevazioni CGM considerate e pertanto variano al variare della grandezza di quest'ultima.

3.2.1 Media

Esprime il livello medio del glucosio e si calcola tramite la media aritmetica dei valori di concentrazione di glucosio contenuti nella finestra considerata:

$$m_G = \frac{1}{N} \sum_{k=1}^N G(k)$$

(3.1)

3.2.2 Mediana

La mediana è un indice statistico di posizione corrispondente al cinquantesimo percentile di un insieme ordinabile. Esso si calcola ordinando in ordine crescente il vettore dei campioni CGM estratti e prendendo quello che occupa la posizione centrale. Per questo indice è necessario distinguere il caso in cui il numero di valori estratti (N) sia pari o dispari:

$$m_{50\%} = \begin{cases} G_s\left(\frac{N}{2}\right) + G_s\left(\frac{N+1}{2}\right), & \text{se } N \text{ è pari} \\ G_s\left(\frac{N+1}{2}\right), & \text{se } N \text{ è dispari} \end{cases} \quad (3.2)$$

3.2.3 Deviazione standard (SD)

La deviazione standard è un indice statistico di dispersione utilizzato per stimare la variabilità della sequenza di campioni CGM in esame. Per il suo calcolo è stata applicata la formula associata allo stimatore non polarizzato:

$$\sigma_G = \sqrt{\frac{1}{N-1} \sum_{k=1}^N [G(k) - m_g]^2} \quad (3.3)$$

3.2.4 Coefficiente di variazione (CV)

Il coefficiente di variazione (CV) è una misura della variabilità relativa, calcolato come rapporto tra la media e la deviazione standard dei campioni estratti. Viene definito in percentuale:

$$CV_{\%} = 100 \cdot \frac{\sigma_G}{m_G} \quad (3.4)$$

3.2.5 Range

Il range non è altro che la differenza tra il massimo e il minimo valore dei campioni considerati:

$$\Delta G_r = \max_k[G(k)] - \min_k[G(k)] \quad (3.5)$$

3.2.6 Range Interquartile (IQR)

Il range interquartile è definito come la differenza tra il settantacinquesimo e il venticinquesimo percentile (terzo e primo quartile) dei valori di glucosio estratti dalla finestra:

$$IQR = Q_3 - Q_1 \quad (3.6)$$

dove Q_3 e Q_1 sono il settantacinquesimo e il venticinquesimo percentile, ottenuti seguendo in ordine i seguenti step:

1. Ordinare in ordine crescente gli N valori CGM estratti dalla finestra;
2. Assegnare ad ogni valore n dell'insieme ordinato il corrispettivo percentile tramite la relazione

$$percentile = 100 * \frac{n - 0.5}{N}$$

3. Se tra i percentili assegnati sono presenti il 25° e il 75° si assegna il valore CGM corrispondente. Altrimenti si calcolano i due percentili tramite interpolazione lineare.

3.2.7 J-Index

È un indice statistico proposto come alternativa all'M-value descritto in seguito. Si ottiene tramite combinazione non lineare di media e varianza [19]:

$$J = \frac{(m_g + \sigma_g)^2}{1000} \quad (3.7)$$

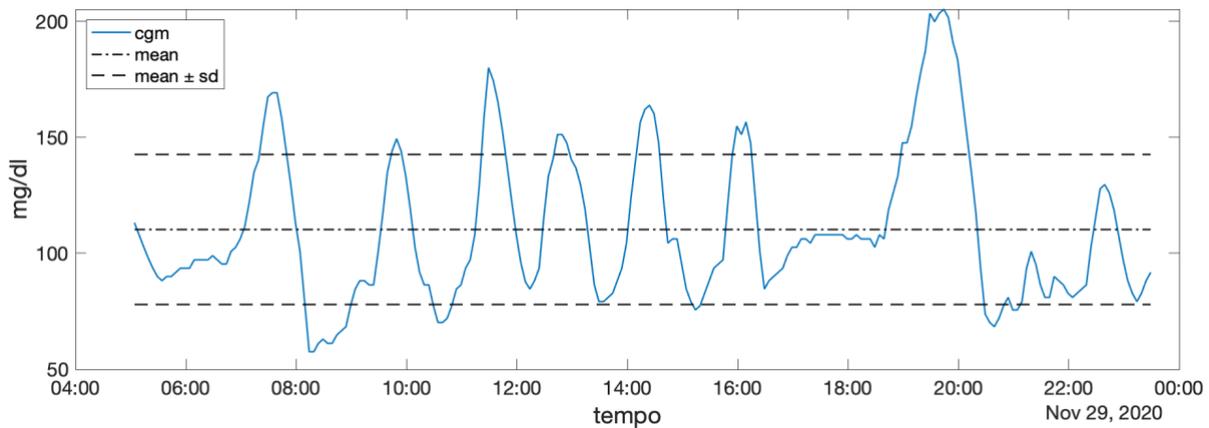


Figura 3.1: Media e deviazione standard di un profilo CGM di riferimento.

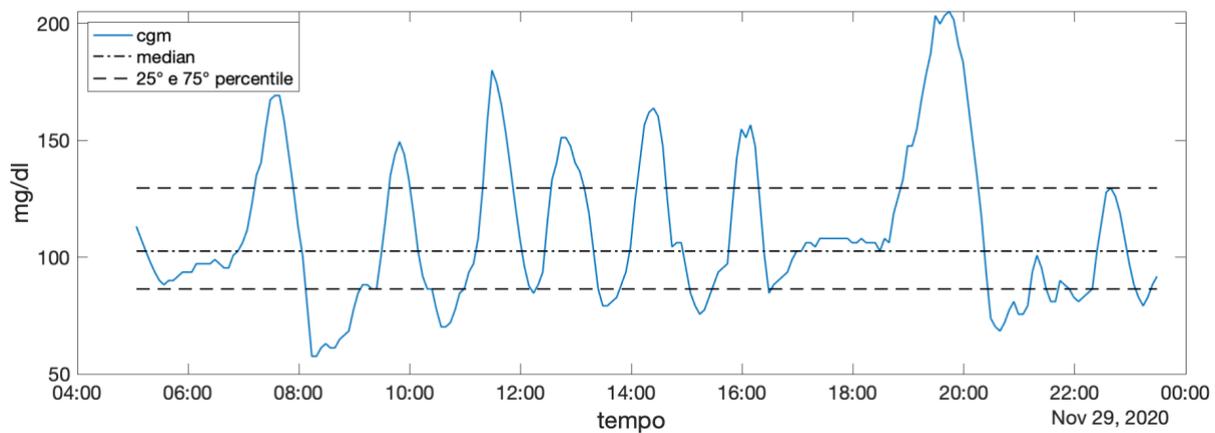


Figura 3.2: Mediana, 25° e 75° percentile di un profilo CGM di riferimento.

3.3 FEATURE BASATE SULLE ESCURZIONI GLICEMICHE

Si tratta di feature utilizzate per sintetizzare le informazioni relative all'intensità e alla frequenza dei picchi glicemici. Alcune tra le metriche più rilevanti di questa categoria sono state sviluppate per i soggetti diabetici, ma trovano un ottimo utilizzo nella popolazione post-bariatrica, vista la natura fisiopatologica delle ipoglicemie.

3.3.1 Mean Amplitude of Glycemic Excursions (MAGE)

Il MAGE è un indice di variabilità ottenuto mediando le escursioni significative del profilo glicemico. La procedura di calcolo utilizzata è quella proposta per i tracciati ottenuti da sensori CGM e descritta in maniera esaustiva nel lavoro di tesi di Enrico Longato [20], a cui si rimanda per maggiore chiarezza. Senza ripercorrere passo per passo il protocollo operativo per la sua estrazione, si riporta di seguito una sintesi dei principali step per il calcolo del MAGE:

1. Individuare tutti gli estremi locali (Turning Point) del profilo glicemico estratto dalla finestra;
2. Eliminare i Turning Point non significativi sulla base dei seguenti criteri:
 - Non sono significativamente diversi da uno o entrambi degli estremi vicini;
 - Sono localizzati all'interno di una sequenza monotona crescente o decrescente.
3. Utilizzando solo i Turning Point rimasti, calcolare il valore delle escursioni glicemiche significative tramite differenza tra il valore CGM di un estremo e quello precedente;
4. Estrarre gli indici MAGE+ e MAGE-, definiti come:

$$MAGE+ = \frac{1}{N_{turn}^+} \sum_{\Delta G_d(k^*) > 0} \Delta G_d(k^*) \quad (3.8)$$

$$MAGE- = \frac{1}{N_{turn}^-} \sum_{\Delta G_d(k^*) < 0} |\Delta G_d(k^*)| \quad (3.9)$$

dove $\Delta G_d(k^*)$ è l'insieme delle escursioni ricavate al punto 3 e N_{turn}^+ e N_{turn}^- sono rispettivamente il numero di escursioni positive e negative;

5. Estrarre il MAGE tramite media aritmetica dei due valori appena ottenuti:

$$MAGE = \frac{MAGE+ + MAGE-}{2} \quad (3.10)$$

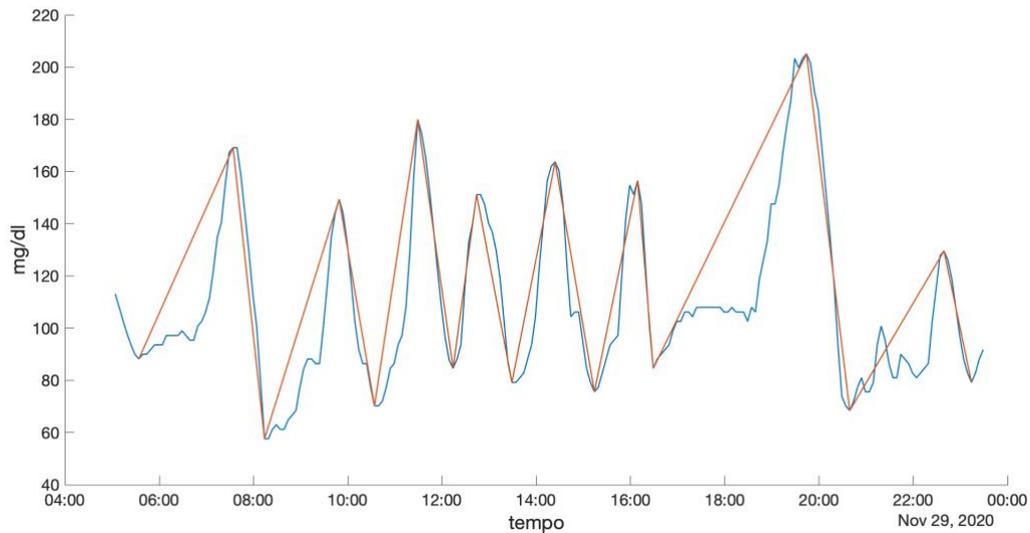


Figura 3.3: Escursioni significative di un profilo CGM di riferimento.

3.3.2 Excursion Frequency (EF)

L'excursion frequency è un indice utilizzato per individuare il numero di escursioni significative presenti all'interno del profilo estratto dalla finestra [21]:

$$EF = \sum_{|\Delta G_d(k^*)| > 75} 1 \quad (3.11)$$

Dove $\Delta G_d(k^*)$ sono le escursioni significative individuate secondo il protocollo operativo descritto in (3.2.1).

3.3.3 Control Variability Grid Analysis (CVGA)

Il Control Variability Grid Analysis (CVGA) è un indice basato su un'applicazione alternativa dell'omonimo metodo grafico proposto da Magni et al. [22] per la valutazione della qualità degli algoritmi di controllo automatico della glicemia. Il suo calcolo prevede l'esecuzione dei seguenti step:

1. Individuare il massimo e il minimo valore di glucosio tra quelli estratti nella finestra;
2. Mappare il profilo estratto nello spazio bidimensionale CVGA tramite le trasformazioni:

$$X = \min\{\max\{(110 - \min BG), 0\}, 60\}$$

$$Y = 2.68 \cdot 10^{-6}(\max BG^3) - 0.0022(\max BG^2) + 0.754(\max BG) - 59.77$$

3. Calcolare la distanza euclidea al quadrato tra il punto di coordinate (X,Y) e il centro dello spazio CVGA (0,0):

$$CVGA = X^2 + Y^2 \tag{3.12}$$

3.3.4 Altri indici di escursione glicemica

Oltre agli indici appena presentati sono stati valutati altri indici di escursione glicemica non presenti nel pacchetto AGATA ed elencati qui di seguito:

- GluPos: La somma di tutti gli incrementi positivi del profilo glicemico estratto dalla finestra;
- GluNeg: La somma di tutti gli incrementi negativi del profilo glicemico estratto dalla finestra;
- MaxDiff: La massima variazione tra un campione e quello successivo, considerando il segno;
- MinDiff: La minima variazione tra un campione e quello successivo, considerando il segno (se sono presenti variazioni di segno negativo verrà presa quella il cui modulo è maggiore);
- Glucose Max: il massimo valore di glucosio estratto dalla finestra;
- Last Local Max: l'ultimo massimo locale presente nella finestra. Nel caso in cui l'ultimo estremo locale del profilo estratto sia un punto di minimo, Last Local Max sarà uguale all'ultimo valore della finestra;
- Prominence: La differenza tra l'ultimo massimo locale del profilo estratto e il minimo locale che lo precede. In caso di profili monotoni esso corrisponderà alla differenza tra il massimo e il minimo valore di glucosio acquisito.

3.4 FEATURE TEMPORALI

Si tratta di feature utilizzate per valutare la permanenza dei valori di glucosio in uno specifico intervallo di valori. Sono definite come la percentuale di campioni localizzata all'interno di una fascia prestabilita. Tra le numerose feature presenti, sono stati valutati:

- Time in Range: percentuale di valori contenuta nell'intervallo [70,180] mg/dl;

$$Time\ in\ Target = 100 \cdot \sum_{k:70 \leq G(k) \leq 180} \frac{1}{N} \quad (3.13)$$

- Time in Hypoglicemia: percentuale di valori sotto i 70 mg/dl;

$$Time\ in\ Hypo = 100 \cdot \sum_{k:G(k) < 70} \frac{1}{N} \quad (3.14)$$

- Time in Hyperglicemia: percentuale di valori sopra i 180 mg/dl;

$$Time\ in\ Hyper = 100 \cdot \sum_{k:G(k) > 180} \frac{1}{N} \quad (3.15)$$

- Time in Tight Target: percentuale di valori contenuta nell'intervallo [90,140] mg/dl;

$$Time\ in\ Tight\ Target = 100 \cdot \sum_{k:90 \leq G(k) \leq 140} \frac{1}{N} \quad (3.16)$$

- Time in Severe Hypoglicemia: percentuale di valori sotto i 54 mg/dl;

$$Time\ in\ Severe\ Hypo = 100 \cdot \sum_{k:G(k) < 54} \frac{1}{N} \quad (3.17)$$

- Time in severe Hyperglycemia: percentuale di valori sopra i 250 mg/dl;

$$Time\ in\ Severe\ Hyper = 100 \cdot \sum_{k:G(k)>250} \frac{1}{N} \quad (3.18)$$

Dove N è il numero di campioni estratti dalla finestra.

3.5 FEATURE DI TRASFORMAZIONE

Si tratta di feature ottenute tramite trasformazioni non lineari del profilo glicemico. Anche in questo caso, le metriche proposte sono state sviluppate per la popolazione diabetica. Tuttavia, visto che il loro intento è quello di amplificare eventuali picchi uscenti dal range euglicemico e che le ipoglicemie post-bariatriche sono spesso precedute da eventi di iperglicemia postprandiale, si è deciso di testarle sulla popolazione in esame.

3.5.1 Glycemic Risk Assessment Diabetes Equation (GRADE)

È un indice di trasformazione ottenuto tramite media aritmetica dei valori di glucosio estratti dalla finestra a cui è stata applicata la seguente trasformazione [23]:

$$GRADE(k) = 425 \cdot \left[\log_{10} \left(\log_{10} \left(\frac{G(k)}{18} \right) + 0.16 \right) \right]^2 \quad (3.19)$$

$$GRADE = \frac{1}{N} \sum_{k=1}^N GRADE(k) \quad (3.20)$$

Oltre al GRADE sono stati valutati altri tre indici basati sul rapporto tra la somma delle trasformazioni dei valori di glucosio contenuti in un range di valori prefissato e la somma delle trasformazioni di tutti i campioni CGM estratti:

$$GRADE_{eu} = \frac{\sum_{\Omega_{eu}} GRADE(k)}{\sum_{\Omega} GRADE(k)} \quad (3.21)$$

$$GRADE_{hypo} = \frac{\sum_{\Omega_{hypo}} GRADE(k)}{\sum_{\Omega} GRADE(k)} \quad (3.21)$$

$$GRADE_{hyper} = \frac{\sum_{\Omega_{hyper}} GRADE(k)}{\sum_{\Omega} GRADE(k)} \quad (3.21)$$

Dove

$$\Omega_{eu} = \{k: 70 \leq G(k) \leq 180\}$$

$$\Omega_{hypo} = \{k: G(k) < 70\}$$

$$\Omega_{eu} = \{k: G(k) > 180\}$$

$$\Omega_{eu} = \Omega_{eu} \cup \Omega_{hypo} \cup \Omega_{hyper}$$

3.5.2 M-Value

L'M-value è un altro indice di trasformazione glicemico molto simile al GRADE appena descritto. Si ricava tramite media aritmetica dei valori di glucosio del profilo estratto sottoposti alla seguente trasformazione [24]:

$$M_R(k) = 1000 \cdot \left| \log_{10} \left(\frac{G(k)}{R} \right) \right|^3 \quad (3.24)$$

$$M_R = \frac{1}{N} \sum_{k=1}^N M_R(k) \quad (3.25)$$

dove R è un parametro da fissare liberamente nell'intervallo di valori 80÷120 [25]. Per la popolazione post-bariatrica è stato utilizzato il valore standard 100.

3.5.3 Indici di controllo glicemico

Gli ultimi indici di trasformazione considerati sono gli indici di valutazione della qualità del controllo glicemico, definiti come:

$$Hypo\ Index = \frac{1}{30 \cdot N} \sum_{\Omega_{hypo}} [70 - G(k)]^2 \quad (3.26)$$

$$Hyper\ Index = \frac{1}{30 \cdot N} \sum_{\Omega_{hyper}} [G(k) - 180]^{1.1} \quad (3.27)$$

dove Ω_{hypo} e Ω_{hyper} sono gli insiemi già definiti nella sezione dedicata al GRADE.

Come si può notare sia dalle formule che dalla figura 3.4, le trasformazioni sopra definite portano a 0 tutti i valori contenuti nel range euglicemico amplificando invece i picchi in maniera maggiore rispetto al GRADE e all'M-value.

I due indici ottenuti vengono poi sintetizzati nell'Index of Glycemic Control:

$$IGC = Hypo\ Index + Hyper\ Index \quad (3.28)$$

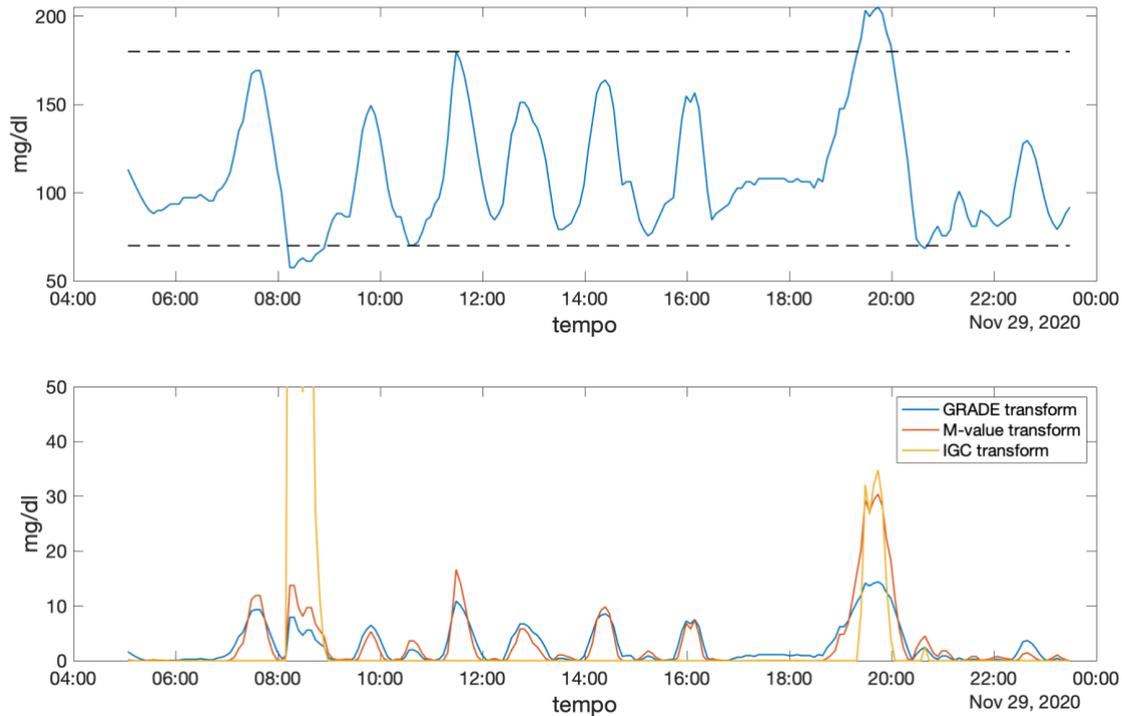


Figura 3.4: Confronto tra le diverse trasformazioni del profilo glicemico

3.6 FEATURE DI RISCHIO

Si tratta di feature molto simili a quelle presentate nel paragrafo precedente in quanto il loro calcolo è di nuovo preceduto da una trasformazione del profilo glicemico, definito dalla relazione:

$$f(G(k)) = \gamma \cdot [(\ln(G(k)))^\alpha - \beta] \quad (3.29)$$

dove α, β, γ sono quantità scalari rispettivamente uguali a 1.084, 5.381, 1.509 e $G(k)$ è un generico valore di glucosio espresso in mg/dl.

La funzione appena presentata pone come obiettivo quello di mappare la scala glicemica in uno spazio in cui essa risulti simmetrica rispetto ad un valore standard contenuto nel range euglicemico. Infatti, si può notare che la funzione:

- È in grado di mappare il range glicemico [20,600] mg/dl in $[-\sqrt{10},\sqrt{10}]$;
- Vale +0.88 nella soglia di iperglicemia (180 mg/dl) e -0.88 nella soglia di ipoglicemia (70 mg/dl);
- È nulla per $G(k)=112.5$ mg/dl.

Successivamente i valori ottenuti da (3.29) vengono ulteriormente trasformati da una ulteriore funzione, individuando un nuovo indice definito rischio statico:

$$r(G(k)) = 10 \cdot [f(G(k))]^2 \quad (3.30)$$

In sintesi, la funzione di rischio statico mappa il range glicemico [20,600] mg/dl nel range [0,100], fornendo valori molto alti quando $|f(G(k))| > 1$, cioè in presenza di valori di concentrazione del glucosio nettamente al di fuori della fascia euglicemica.

Per il calcolo delle feature appartenenti a questa categoria, il rischio statico viene ridefinito separando i casi sulla base delle condizioni di ipoglicemia e iperglicemia:

$$r_l(G(k)) = \begin{cases} r(G(k)), & \text{se } f(G(k)) < 0 \\ 0, & \text{altrimenti} \end{cases} \quad (3.31)$$

$$r_h(G(k)) = \begin{cases} r(G(k)), & \text{se } f(G(k)) > 0 \\ 0, & \text{altrimenti} \end{cases} \quad (3.32)$$

3.6.1 Blood Glucose Risk Index (BGRI)

Il BGRI [26] è un indice di rischio statico ottenuto dalla somma di altri due indici, il Low Blood Glucose Index (LBGI) e l'High Blood Glucose Index (HBGI), ricavati semplicemente dalla media di $r_l(G(k))$ e $r_h(G(k))$:

$$LBGI = \frac{1}{N} \sum_{k=1}^N r_l(G(k)) \quad (3.33)$$

$$HBGI = \frac{1}{N} \sum_{k=1}^N r_h(G(k))$$

(3.34)

$$BGRI = LBGI + HBGI$$

(3.35)

3.6.2 Average Daily Risk Range (ADRR)

L'ADRR si ottiene mediando la somma dei massimi di $r_l(G(k))$ e $r_h(G(k))$ di ogni giornata acquisita. Ovviamente le finestre usate per estrarre le metriche su cui fare predizione sono molto più piccole di una giornata. Pertanto, l'ADRR verrà definito semplicemente come:

$$ADRR = \max_k r_l(G(k)) + \max_k r_h(G(k))$$

(3.36)

3.6.3 Rischio Dinamico (DR)

Il rischio dinamico è un indice di rischio ottenuto modificando la funzione di rischio statico in modo che essa in considerazione non solo l'attuale livello di glucosio, ma anche il suo andamento, misurato tramite calcolo della derivata di primo ordine [27]. Esso è definito come:

$$DR\left(g, \frac{dg}{dt}\right) = \begin{cases} SR(g) \cdot e^{+\mu \frac{dr}{dt}}, & \text{se } SR(g) > 0 \\ SR(g) \cdot e^{-\mu \frac{dr}{dt}}, & \text{se } SR(g) < 0 \end{cases}$$

(3.37)

con

$$SR = R_h(g) - R_l(g)$$

$$\frac{dr}{dt} = \frac{dr}{dg} \frac{dg}{dt} = \left\{ 10\gamma^2 \cdot \left[(\ln(g))^{2\alpha-1} - \beta(\ln(g))^{\alpha-1} \right] \cdot 2\alpha \frac{1}{g} \right\} \cdot \frac{dg}{dt}$$

dove α, β, γ sono le quantità scalari definite per il rischio statico, μ è un fattore di amplificazione, in questo caso posto a 2.5, e g è un generico valore di glucosio espresso in mg/dl.

L'uso del rischio dinamico nell'ambito dello sviluppo di algoritmi di predizione delle ipoglicemie si è rilevato particolarmente efficace [28] in virtù della sua capacità di amplificare il rischio statico quando la glicemia si muove al di fuori del range euglicemico e di attenuarlo quando la glicemia si muove verso il range euglicemico, segnalando più accuratamente possibili ipoglicemie imminenti.

Per la valutazione delle metriche sono stati considerati solo l'ultimo rischio dinamico dell'ultimo campione CGM estratto dalla finestra e la media degli ultimi 3.

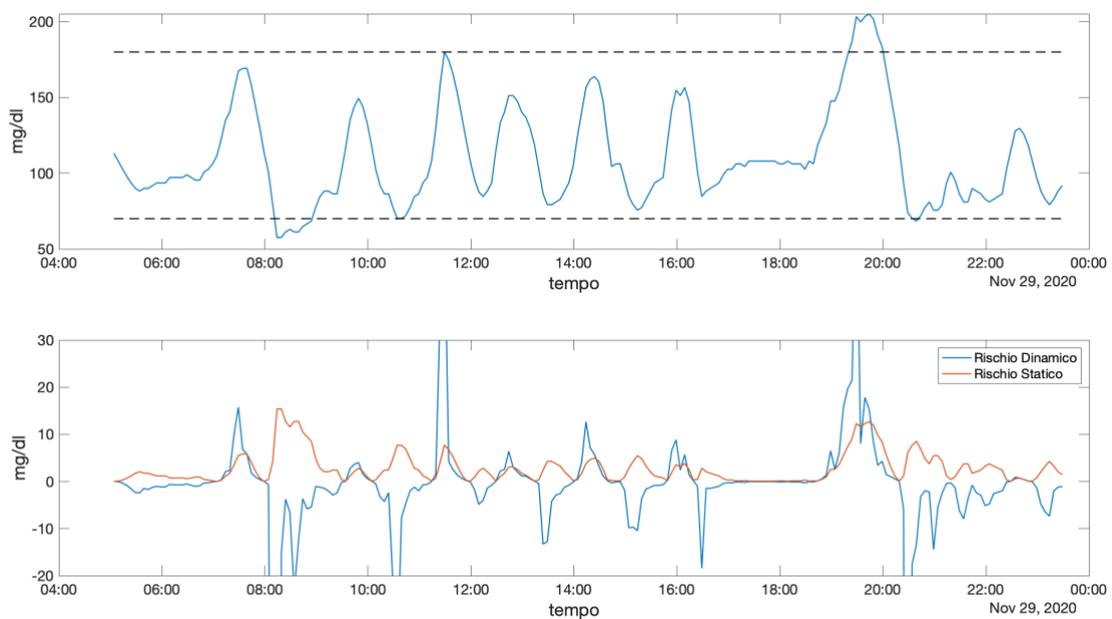


Figura 3.5: Confronto tra rischio statico e rischio dinamico ottenuto da un profilo di riferimento.

3.7 FEATURE SHORT-TIME

Le feature short-time sono chiamate così perché tengono in considerazione soltanto la porzione finale del profilo glicemico estratto all'interno della finestra. Pertanto, non sono influenzate dalla grandezza di quest'ultima. Il loro obiettivo è quello di quantificare il trend del profilo glicemico. A questa categoria appartengono:

- *Derivata di primo ordine o rate of change (ROC)*: definita come

$$G'(k) = \frac{G(k) - G(k - 1)}{T_s} \quad [mg/dl/min] \quad (3.38)$$

con T_s intervallo di campionamento. Sono state considerate l'ultimo ROC e la media degli ultimi 3;

- *DiffX*: definita come la differenza tra l'ultimo valore CGM estratto dalla finestra e quello acquisito X minuti prima. In questo lavoro si sono considerate le differenze di 10, 20 e 30 minuti prima della fine della finestra mobile, indicati rispettivamente come Diff10, Diff20, Diff30.

CAPITOLO 4

L'ANALISI DI CORRELAZIONE

In questo capitolo verranno presentate modalità e risultati dell'analisi di correlazione svolta per l'identificazione del feature set da utilizzare nello sviluppo degli algoritmi di predizione delle ipoglicemie. In particolare, si descriverà il processo di estrazione delle metriche basato sull'uso di una "finestra mobile" e si illustreranno i diversi indici di correlazione utilizzati. L'obiettivo finale è individuare un subset delle oltre quaranta feature analizzate da utilizzare nelle successive fasi di training dell'algoritmo.

4.1 PROCESSING DEI DATI PER L'ANALISI DI CORRELAZIONE

Una volta riempito i missing value ed eliminato i periodi notturni dai tracciati, è richiesta una ulteriore procedura di processing dei dati prima di passare al calcolo effettivo dei coefficienti di correlazione. Lo scopo di quest'ulteriore passaggio è quello di estrarre una serie di vettori di rilevazioni CGM su cui calcolare le feature e di associare ad ognuno di essi il corrispettivo valore di glucosio al tempo futuro. Prima di fornire una descrizione accurata della procedura applicata è utile definire alcuni dei termini che verranno utilizzati in seguito. Si definisce orizzonte di predizione (PH) la distanza temporale tra il campione acquisito e quello di cui si vuole anticiparne il valore; grandezza della finestra il prodotto tra il numero delle ultime rilevazioni da estrarre e l'intervallo di campionamento.

Di seguito si riportano in maniera sequenziale i vari passaggi per l'estrazione dei vettori di rilevazioni CGM:

1. Fissare la grandezza della finestra di estrazione e l'orizzonte di predizione da usare;
2. Scegliere una delle giornate ottenute dopo la fase di pre-processing;
3. Salvare i primi N_w campioni di una giornata nella prima riga di una matrice X;
4. Salvare il campione CGM acquisito al tempo $t = t_{windowend} + PH$ al primo indice di un vettore Y;
5. Traslare la finestra un campione in avanti ed estrarre nuovamente gli N_w campioni contenuti al suo interno da salvare in X e l'acquisizione futura da salvare in Y;

6. Ripetere lo step cinque fino a quando il campione da inserire in Y non corrisponde all'ultima rilevazione della giornata;
7. Ripetere i passaggi dal tre al sei per ognuna delle restanti giornate presenti nel dataset.

Come controllo della corretta esecuzione di questa fase, è utile verificare che la matrice X ottenuta abbia dimensioni $(N_{day} \cdot (L_{day} - (N_w + PH/T_s - 1))) \times N_w$, con N_{day} numero di giornate, L_{day} numero di campioni in una giornata e N_w numero di campioni contenuti nella finestra di estrazione e che il vettore Y abbia dimensione pari al numero di righe di X.

A questo punto, prima di procedere con l'estrazione delle metriche, si è scelto di ridurre la dimensionalità dell'insieme ottenuto tramite l'esecuzione di diverse procedure di sottocampionamento, ognuna via via meno selettiva dell'altra. L'obiettivo è quello di conservare le osservazioni associate a specifici momenti della giornata, come ad esempio i periodi antecedenti ad un evento di ipoglicemia, in modo da individuare una possibile correlazione tra le metriche e le porzioni del tracciato di interesse.

Le procedure di sottocampionamento applicate sono state scelte in modo da soddisfare le seguenti proprietà:

- Non devono contenere doppioni;
- Non devono contenere osservazioni in cui il campione estratto alla fine della finestra si trova all'interno di un evento di ipoglicemia;
- Le procedure meno selettive devono sempre contenere tutte le osservazioni di quelle più selettive.

Definite le proprietà si descrivono le tre diverse procedure applicate, partendo da quella più selettiva:

1. Vengono mantenute solo le osservazioni il cui l'acquisizione futura è contenuta in una ipoglicemia;
2. Vengono mantenute anche le ultime quattro osservazioni antecedenti a quella associata all'inizio della ipoglicemia;
3. Vengono mantenute anche, delle restanti osservazioni, una ogni quattro, scorrendo le osservazioni.

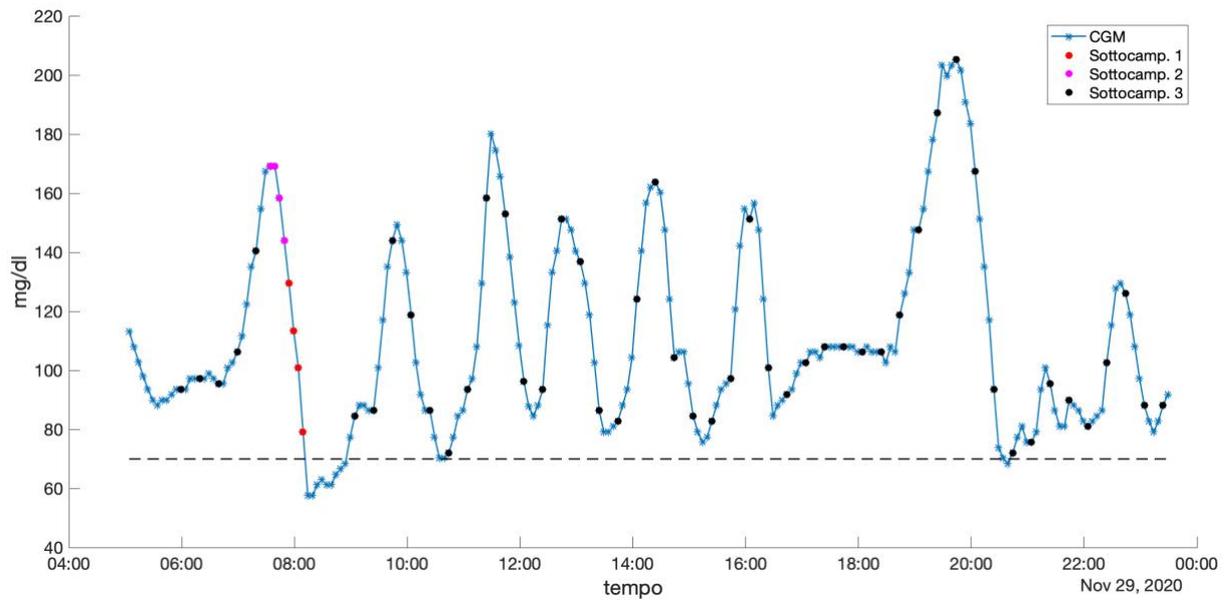


Figura 4.1: Osservazioni rimanenti dopo i vari sottocampionamenti utilizzando una finestra di 60 min e PH di 20 min. I pallini si riferiscono alla posizione della fine della finestra per le osservazioni conservate.

Terminato il sottocampionamento, l'ultimo step del processing dei dati prevede l'estrazione delle metriche, con cui verrà calcolato il grado di correlazione rispetto alle acquisizioni contenute in Y.

4.2 GLI INDICI DI CORRELAZIONE

Per misurare il grado di correlazione tra le metriche e i valori di glucosio futuri sono stati utilizzati due diversi indici: il coefficiente di correlazione di Pearson e il coefficiente di correlazione di Spearman.

4.2.1 Coefficiente di correlazione di Pearson

Il coefficiente di correlazione di Pearson è un indice usato per valutare l'esistenza di una relazione lineare tra due variabili statistiche o insiemi di dati. Può essere applicato solo nel caso in cui le due variabili soddisfino le seguenti assunzioni:

1. Level of measurement: le due variabili devono essere quantitative;
2. Related pairs: I dati devono essere appaiati, cioè per ogni osservazione devono esserci sia la misura di una variabile che dell'altra;
3. Linearity: la relazione tra le due osservazioni deve essere lineare;
4. Absence of outliers: non devono esserci outliers, poiché la loro presenza può alterare considerevolmente il risultato ottenuto.

Il coefficiente di correlazione di Pearson è definito come il rapporto tra la covarianza delle due variabili e il prodotto delle loro deviazioni standard. Nello specifico se l'indice viene calcolato su una popolazione viene denominato con la lettera greca ρ e si definisce:

$$\rho_{X,Y} = \frac{\text{cov}(X, Y)}{\sigma_X \sigma_Y} \quad (4.1)$$

Se invece viene calcolato su campione estratto dalla popolazione viene denominato con la lettera r e si definisce:

$$r_{x,y} = \frac{\sum_{i=1}^N (x_i - \bar{x})(y_i - \bar{y})}{\sqrt{\sum_{i=1}^N (x_i - \bar{x})^2} \sqrt{\sum_{i=1}^N (y_i - \bar{y})^2}} \quad (4.2)$$

Si noti che la relazione in (4.2) non è altro che la (4.1) a cui sono stati sostituiti i vari termini con le loro stime campionarie.

Il coefficiente può assumere valori compresi tra -1 e 1, dove:

- Se $\rho_{X,Y} > 0$, le due variabili sono direttamente correlate, cioè all'aumentare dell'una corrisponde un aumento dell'altra;
- Se $\rho_{X,Y} = 0$, le due variabili sono incorrelate;
- Se $\rho_{X,Y} < 0$, le due variabili sono inversamente correlate, cioè all'aumentare dell'una corrisponde una diminuzione dell'altra.

Inoltre, di maggiore interesse per l'individuazione delle feature rilevanti, tanto più grande è $|\rho_{X,Y}|$ tanto più forte è la relazione tra le due variabili. Per fornire una scala qualitativa del grado di relazione si riporta quanto indicato da Evans nella sua guida del 1996 [29]:

- Se .00 - .19 correlazione molto debole;
- Se .20 - .39 correlazione debole;
- Se .40 - .59 correlazione moderata;
- Se .60 - .79 correlazione forte;
- Se .80 - 1.0 correlazione molto forte.

4.2.2 Coefficiente di correlazione di Spearman

Il coefficiente di correlazione di Spearman è una misura non parametrica di correlazione per ranghi ideata da Charles Spearman [29]. Può essere applicato a qualsiasi insieme di dati, purché essi siano almeno ordinabili, ed è in grado di stabilire quanto bene una relazione tra due variabili può essere descritta da una funzione monotona. Matematicamente si definisce come:

$$r_s = \rho_{R(X),R(Y)} = \frac{\text{cov}(R(X), R(Y))}{\sigma_{R(X)}\sigma_{R(Y)}} \quad (4.3)$$

Osservando la relazione (4.1), si può notare che il coefficiente r_s di Spearman non è altro che il coefficiente di correlazione di Pearson applicato ai ranghi delle due variabili statistiche. Per questo motivo, tutto ciò che è stato detto nel paragrafo precedente in merito a segno e modulo del coefficiente di correlazione di Pearson si applica in maniera eguale anche per quello di Spearman e pertanto non verrà ripetuto.

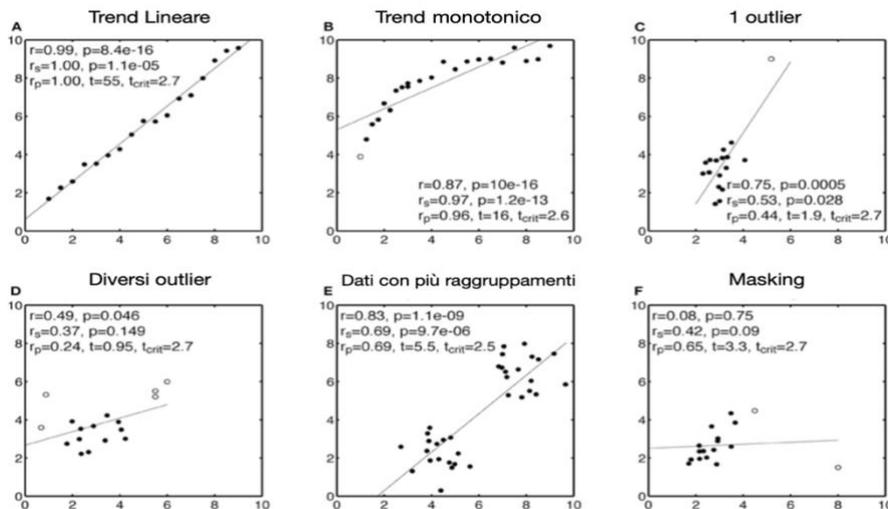


Figura 4.2: Coefficienti di correlazione di Pearson e Spearman su dati aventi diverse tendenze.

4.3 RISULTATI DELL'ANALISI DI CORRELAZIONE

Nel primo paragrafo del capitolo sono state descritte dettagliatamente le varie procedure di processing a cui il dataset è stato sottoposto in preparazione al calcolo dei coefficienti di correlazione appena presentati. In tale sezione, tuttavia, si ommesso di dire come la scelta di alcuni parametri, indispensabili per il completamento del processing stesso, influenzi i valori dei coefficienti ottenuti. Nello specifico, questi parametri sono:

- La grandezza della finestra di estrazione, che influisce sul calcolo della maggior parte delle metriche considerate, ad eccezione del rischio dinamico e delle feature short-time;
- Il PH, che influisce sulla creazione delle coppie (metrica, glucosio) da correlare;
- Il tipo di sottocampionamento, che circoscrive il risultato alle sole porzioni di tracciato considerate.

Per garantire un'analisi quanto più completa possibile, si è quindi stabilito una serie di grandezze della finestra (30, 45, 60, 90, 120 minuti) e di PH (15, 20, 30 minuti) da combinare con ogni tipo di sottocampionamento implementato e per ogni terna di combinazioni sono stati calcolati i coefficienti di correlazione tra le metriche e i valori di glucosio futuri. Per facilitare la consultazione dell'enorme mole di dati ottenuta si è scelto di suddividere i risultati in base alla tipologia di sottocampionamento applicata, creando delle tabelle in cui sono state riportate

solo le metriche con almeno una correlazione debole per diverse coppie di PH e grandezza della finestra. Le tabelle create, invece di riportare i valori dei coefficienti ottenuti, si servono di colori per segnalare il grado di correlazione, sulla base della seguente legenda:

LEGENDA	
	$0.2 \leq \rho < 0.3$
	$0.3 \leq \rho < 0.4$
	$0.4 \leq \rho < 0.5$
	$0.5 \leq \rho < 0.6$
	$0.6 \leq \rho < 0.7$
	$0.7 \leq \rho < 0.8$

Tabella 4.1: Legenda dell'associazione colore grado coefficiente

L'impatto visivo dato dalla colorazione dovrebbe facilitare l'individuazione delle metriche più rilevanti e l'identificazione di eventuali tendenze al variare dei parametri utilizzati.

RISULTATI CON METODO DI SOTTOCAMPIONAMENTO 1

Coefficiente	Pearson									Spearman									
	120			90			60			120			90			60			
	15	20	30	15	20	30	15	20	30	15	20	30	15	20	30	15	20	30	
J-index																			
MAGE																			
MAGE+																			
mean																			
median																			
range																			
std																			
hbgi																			
lbgi																			
GRADE																			
last 3 ROC mean																			
last ROC																			
Last DR																			
Last 3 DR Mean																			
ROCstd																			
Diff10																			
Diff20																			
glupos																			
MaxDiff																			
MinDiff																			
maxglu																			
Last Local Max																			

Tabella 4.2: Risultati analisi correlazione con metodo di sottocampionamento 1
(conservazione delle osservazioni il cui l'acquisizione futura è contenuta in una ipoglicemia).

La tabella mostra i risultati ottenuti dell'analisi di correlazione per il sottocampionamento 1. Le correlazioni individuate sono tutte deboli già a partire dal PH=15 [min] e non sembra esserci alcun miglioramento al variare della grandezza della finestra di estrazione. Nonostante siano negativi, questi risultati sono in linea con le aspettative e rappresentano un'ulteriore prova della

complessità del problema affrontato. Infatti, individuare una forte correlazione con alcune delle metriche delle osservazioni conservate, significherebbe disporre di un indicatore dell'imminente discesa dei valori glucosio nella fascia ipoglicemica.

RISULTATI CON METODO DI SOTTOCAMPIONAMENTO 2

Coefficiente	Pearson									Spearman								
	120			90			60			120			90			60		
Finestra [min]	15	20	30	15	20	30	15	20	30	15	20	30	15	20	30	15	20	30
mage+				Y	Y		R	R								R	Y	
Mage-	Y	Y		Y	Y		R	R	Y	Y	Y		R	R	Y	R	R	Y
last 3 ROC mean	R	R	R	R	R		R	R		R	R		R	R		R	R	
last ROC		R	R	Y	R		Y	R		Y	R		Y	R		Y	R	
last 3 DR mean										R	R		R	R		R	R	
Last DR										R	R	Y	R	R	Y	R	R	Y
Diff10	R	R	R	R	R		R	R		R	R		R	R		R	R	
Diff20	R	R	R	R	R		R	R		R	R		R	R		R	R	
Diff30	R	R	R	R	R		R	R		R	R	Y	R	R	Y	R	R	Y
glupos	Y			Y	Y		R	R	Y							Y	Y	
Gluneg	Y	Y		Y	Y		Y	Y		Y	Y		R	R		R	R	
MaxDiff							R	Y								Y		
MinDiff				Y	Y		Y	Y		Y	Y		Y	Y		Y	Y	
MaxGlu							Y											
Last Local Max							Y											
Prominence				Y			Y	Y										

Tabella 4.3: Risultati analisi correlazione con metodo di sottocampionamento 2 (Osservazioni sottocampionamento 1 più le quattro antecedenti a quella associata all'inizio dell'ipoglicemia).

Appurato che non esistono indicatori diretti che permettano di individuare con anticipo un evento di ipoglicemia, si è investigata l'esistenza di un pattern comune nei periodi immediatamente precedenti all'inizio dell'ipoglicemia predicibile dalle metriche considerate. I risultati sono decisamente più soddisfacenti e individuano alcune correlazioni moderate tra le feature basate sulle escursioni e quelle short-time, in aggiunta ad una forte correlazione con il rischio dinamico. Tali correlazioni calano vistosamente all'aumentare del PH, evidenziando

come le dinamiche estremamente rapide dei tracciati rendano la predizione su orizzonti di predizione ampi un obiettivo difficilmente raggiungibile con questo tipo approccio. Relativamente alle metriche individuate, anche in questo caso i risultati sono in linea con le aspettative. È noto, infatti, che prima di un evento di ipoglicemia il glucosio tenda ad assumere valori molto alti, con picchi iperglicemici talvolta piuttosto moderati, per poi calare bruscamente a causa dell'iperinsulinemia. E questa tendenza viene evidenziata dalla variazione delle metriche sopra individuate.

Osservando attentamente la tabella (4.3) si può notare inoltre come una diminuzione della finestra porti ad un aumento dei valori dei coefficienti di correlazione per alcune delle metriche dipendenti da tale parametro (si veda MAGE+ o GluNeg). Per questo motivo l'analisi è stata arricchita calcolando il grado di correlazione anche per le finestre di 45 e 30 minuti, riportati nella tabella (4.4).

Coefficiente	Pearson						Spearman					
	45			30			45			30		
Finestra [min]	15	20	30	15	20	30	15	20	30	15	20	30
PH [min]	15	20	30	15	20	30	15	20	30	15	20	30
J-Index	Yellow			Orange	Yellow					Yellow		
MAGE+	Red	Light Red	Orange	Red	Red	Orange	Orange	Orange	Yellow	Orange	Orange	Yellow
MAGE-	Orange	Orange	Yellow	Yellow	Orange	Yellow	Orange	Orange	Yellow	Orange	Orange	Yellow
mean	Yellow			Light Red	Orange					Yellow		
Median	Yellow			Orange	Yellow					Yellow		
HBGI	Yellow			Orange	Orange					Yellow		
LBGI							Yellow			Orange	Yellow	
Time in Hyper	Yellow			Orange	Yellow							
GRADE	Yellow			Light Red	Orange					Yellow		
GRADE _{hyper}				Orange	Yellow							
M-value	Yellow			Orange	Yellow							
Hyper Index	Yellow			Orange	Yellow							
last 3 ROC mean	Orange	Light Red	Orange	Orange	Light Red	Orange	Orange	Orange	Orange	Orange	Orange	Orange
last ROC	Yellow	Orange	Orange	Yellow	Orange	Orange	Yellow	Orange	Orange	Yellow	Orange	Orange
Last 3 DR mean							Dark Red	Light Red		Dark Red	Light Red	
Last DR							Dark Red	Red	Yellow	Dark Red	Red	Yellow
Diff10	Orange	Light Red	Orange	Orange	Light Red	Orange	Orange	Orange	Orange	Orange	Orange	Orange
Diff20	Light Red	Light Red	Orange	Light Red	Light Red	Orange	Orange	Orange	Orange	Orange	Orange	Orange

Diff30	Red	Red	Orange	White	White	White	Red	Red	Yellow	White	White	White
GluPos	Red	Red	Orange	Red	Red	Orange	Orange	Orange	White	Orange	Orange	Yellow
GluNeg	Yellow	Yellow	Yellow	Yellow	Yellow	Yellow	Orange	Orange	Yellow	Orange	Orange	Yellow
MaxDiff	Red	Orange	Yellow	Red	Red	Orange	Orange	Yellow	White	Orange	Orange	Yellow
MinDiff	Yellow	Yellow	Yellow	Yellow	Orange	Yellow	Yellow	Orange	Yellow	Yellow	Orange	Yellow
MaxGlu	Yellow	White	White	Orange	Yellow	White	White	White	White	Yellow	White	White
Last Local Max	Yellow	White	White	Orange	Yellow	White	White	White	White	Yellow	White	White

Tabella 4.4: Risultati analisi correlazione con metodo di sottocampionamento 2

La riduzione della finestra ha introdotto nuove correlazioni, sia deboli che moderate, e ha intensificato quelle già esistenti. Questo evidenzia come non sia necessario calcolare le metriche su una finestra ampia, ma sia sufficiente una più piccola.

RISULTATI CON METODO DI SOTTOCAMPIONAMENTO 3

Coefficiente	Pearson						Spearman					
	Finestra [min] 45			30			45			30		
PH [min]	15	20	30	15	20	30	15	20	30	15	20	30
CVGA	Yellow	Yellow	Yellow	Yellow	Yellow	Yellow	Red	Orange	Orange	Red	Orange	Orange
J-index	Orange	Yellow	White	Red	Orange	Yellow	Orange	Orange	White	Red	Orange	Yellow
MAGE+	Red	Red	Yellow	Red	Red	Orange	Orange	Yellow	White	Red	Orange	Yellow
MAGE-	Yellow	Yellow	Yellow	Yellow	Yellow	Yellow	Orange	Orange	Yellow	Orange	Orange	Orange
mean	Red	Orange	White	Red	Red	Yellow	Red	Orange	Yellow	Red	Orange	Yellow
median	Orange	White	White	Red	Orange	Yellow	Orange	Yellow	White	Red	Orange	Yellow
HBGI	Orange	Yellow	White	Red	Orange	White	Orange	Yellow	White	Orange	Yellow	White
LBGI	Yellow	Yellow	Yellow	Orange	Yellow	Yellow	Red	Red	Orange	Red	Red	Orange
Time In Tight	White	White	White	White	White	White	Yellow	Yellow	Yellow	Yellow	Yellow	Yellow
Time In Hyper	Orange	Yellow	White	Orange	Orange	White	White	White	White	Yellow	White	White
GRADE	Orange	Yellow	White	Red	Orange	White	Yellow	White	White	Orange	Yellow	White
GRADE _{hyper}	Orange	Yellow	White	Orange	Orange	White	White	White	White	Yellow	White	White
last 3 ROC mean	Red	Red	Red	Red	Red	Red	Red	Red	Orange	Red	Red	Orange
last ROC	Red	Red	Red	Red	Red	Red	Orange	Red	Orange	Red	Red	Red

MAGE-	MAGE+
Mean	Median
Diff10	CVGA
Diff20	Hbgi
Diff30	Lbgi
Media degli ultimi 3 ROC	Last Local Max
GluPos	GluNeg
MaxDiff	MinDiff
Last DR	Glucose Max
Range	GRADE

Tabella 4.5: Elenco delle feature selezionate.

Le feature selezionate coprono quasi tutte le diverse categorie analizzate nel capitolo precedente, anche se si evidenzia una preferenza per quelle basate sulle escursioni e short-time. Ciò non sorprende in quanto per ognuna di esse è possibile individuare una diretta correlazione tra la variazione dei loro valori e una delle due principali caratteristiche della dinamica del segnale CGM antecedente ad un evento di ipoglicemia (rapido innalzamento e abbassamento dei valori di glucosio).

Come ulteriore conferma delle feature selezionate, è stata eseguita un'ulteriore analisi basata sulla creazione di diversi boxplot. Nello specifico, dall'insieme delle osservazioni ottenuto prima dell'esecuzione di alcun sottocampionamento, sono state escluse tutte quelle ottenute negli istanti in cui la fine della finestra era localizzata all'interno di un evento di ipoglicemia. Successivamente le osservazioni rimaste sono state separate in due categorie. La prima contenente tutte le osservazioni ricavate quando la fine della finestra era posizionata dai cinque ai trenta minuti prima dell'inizio di un'ipoglicemia, la seconda contenente tutte le restanti. Successivamente sono stati generati e confrontati i boxplot per ogni metrica delle due classi, individuando quelle per cui la mediana di una categoria era esterna al box dell'altra. Molte delle metriche individuate, in particolar modo quelle basate sulle escursioni e quelle short-time, sono presenti tra l'elenco delle feature selezionate. A titolo di esempio si riportano alcuni dei boxplot generati tra le metriche individuate.

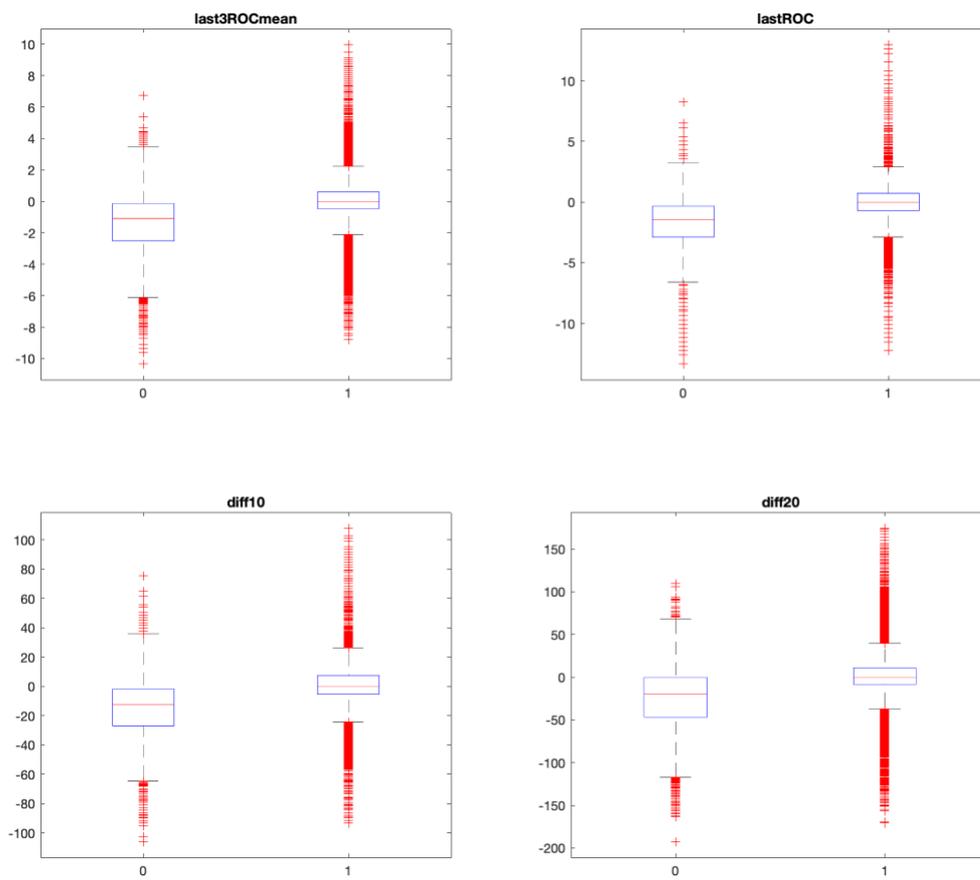


Figura 4.3: alcuni dei boxplot aventi la mediana esterna al box dell'altra classe.

CAPITOLO 5

METODOLOGIE PER LA PREDIZIONE DELLE IPOGLICEMIE

Nel seguente capitolo verranno descritte sinteticamente le metodologie scelte per la predizione delle ipoglicemie post-bariatriche. Come anticipato nella presentazione degli obiettivi della tesi, si è lavorato nell'ambito della classificazione binaria tramite apprendimento supervisionato (supervised learning). Formalmente, dato un insieme di vettori \mathbf{x} appartenente al Domain Set X e un insieme di valori y appartenenti al Target Set Y , l'apprendimento supervisionato sfrutta la conoscenza a priori delle coppie $(\mathbf{x}_i, y_i) \in X \times Y$ contenute in un Training Set $S = ((\mathbf{x}_1, y_1) \dots (\mathbf{x}_m, y_m))$ per stimare una funzione (predittore) $h : X \rightarrow Y$ in grado di associare a nuovi punti di X una predizione del valore target di Y . Nella classificazione binaria Y è un insieme di due elementi, tipicamente $\{0,1\}$ o $\{-1,+1\}$, che identifica l'appartenenza ad una classe positiva o negativa. In particolare, le metodologie scelte per la predizione delle ipoglicemie post-bariatriche sono: logistic regression, support vector machine (SVM), decision tree, random forest, logistic model tree. Ognuno di questi modelli avrà come input l'insieme delle 20 metriche scelte e avrà come output $Y = \{0,1\}$, con 0 ad indicare nessuna ipoglicemia imminente e 1 viceversa.

5.1 LOGISTIC REGRESSION

La regressione logistica (logistic regression) è una tecnica di classificazione lineare utilizzata per modellare la probabilità di esistenza di un evento o di appartenenza ad una certa classe. È chiamata regressione, nonostante sia impiegata nell'ambito della classificazione, perché la funzione stimata $h: X \rightarrow Y$ mappa i vettori del domain space nell'intervallo $[0,1]$, interpretabile come la probabilità di appartenenza alla classe positiva. Invece di servirsi direttamente dei valori di output che, come si è detto, sono rappresentati da variabili dicotomiche, la regressione logistica modella il log odds(Y), definito come il logaritmo del rapporto tra la probabilità p per cui $y=1$ e la probabilità $(1-p)$ per cui $y = -1$. Formalmente il modello è ricavato tramite la risoluzione del problema:

$$\ln\left(\frac{p}{1-p}\right) = b + \sum_{j=1}^n w_j x_j \quad (5.1)$$

dove b e w_j sono i coefficienti del modello e x_j le feature del domain space. Matematicamente la (5.1) può essere riscritta in funzione di p come:

$$p = \frac{1}{1 + e^{-(b + \sum_{j=1}^n w_j x_j)}} = \frac{1}{1 + e^{(-logit)}} \quad (5.2)$$

dove si evince che la probabilità è modellata come una funzione sigmoidale dipendente dalla combinazione lineare delle feature e dei coefficienti del modello, chiamata funzione logistica. La classe di ipotesi associata al modello logistico è quindi la composizione di una funzione sigmoidale $\phi_{sig}: \mathbb{R} \rightarrow [0,1]$ sulla classe di funzioni lineari affini

$L_d = \{h_{\mathbf{w},b} \mapsto \langle \mathbf{w}, \mathbf{x} \rangle + b : \mathbf{w} \in \mathbb{R}^d, b \in \mathbb{R}\}$, matematicamente definita come:

$$H_{sig} = \phi_{sig} \circ L_d = \{\mathbf{x} \mapsto \phi_{sig}(\langle \mathbf{w}, \mathbf{x} \rangle + b) : \mathbf{w} \in \mathbb{R}^d\} \quad (5.3)$$

Osservando la figura (5.1) si può notare che la funzione logistica appena citata gode delle seguenti proprietà:

- Ha per codominio l'intervallo $(0,1)$, come i valori assunti da una probabilità;
- Tende a 1 quando $\langle \mathbf{w}, \mathbf{x} \rangle$ è molto grande;
- Tende a 0 quando $\langle \mathbf{w}, \mathbf{x} \rangle$ è molto piccolo;
- È pari a 0.5 quando $\langle \mathbf{w}, \mathbf{x} \rangle + b$ è nullo.

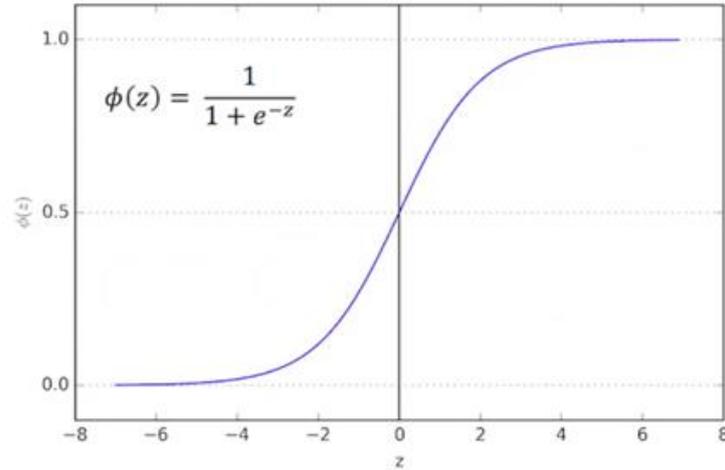


Figura 5.1: Grafico funzione logistica.

Intuitivamente, maggiore è il $|\langle \mathbf{w}, \mathbf{x} \rangle|$ maggiore è la probabilità di appartenenza ad una specifica classe. Per cui l'assegnazione ad una delle due si basa sul seguente criterio:

$$\begin{cases} y = 1 & \text{se } P(y = 1|\mathbf{x}) \geq \alpha \\ y = 0 & \text{se } P(y = 0|\mathbf{x}) < \alpha \end{cases} \quad (5.4)$$

Dove α è una soglia di probabilità da fissare, ma generalmente posta a 0.5. La scelta del parametro α è forse uno degli aspetti cruciali per il miglioramento delle performance del modello, in quanto, molto spesso, il valore di default 0.5 non si adatta alle reali distribuzioni delle classi, specialmente quando c'è uno forte sbilanciamento tra le due. In questi casi può essere utile adottare una strategia di correzione della soglia di classificazione. Tra le più comuni si citano quelle di analisi della *receiver operating characteristic curve* (curva ROC) per la massimizzazione della G-mean o della Youden s J statistic, o la strategia di massimizzazione dell'F1-score, di cui si parlerà nel prossimo capitolo.

Per stimare i parametri del modello è necessario definire una funzione costo che assegni una penalità tanto più grande quanto minore è la probabilità assegnata ad ogni osservazione del training set di appartenere alla classe vera. Nella regressione logistica questa funzione è definita come *logistic loss function*, definita come:

$$l(h_{\mathbf{w},b}(\mathbf{x}, y)) = \log(1 + e^{-y(\langle \mathbf{w}, \mathbf{x} \rangle + b)}) \quad (5.5)$$

Per cui, dato l'insieme di coppie $(x_i, y_i) \in X \times Y$ contenuto nel training set, il problema della stima dei coefficienti del modello si riduce alla minimizzazione della funzione:

$$\underset{\mathbf{w}, b}{\operatorname{argmin}} \frac{1}{m} \sum_{i=1}^m \log(1 + e^{-y_i(\langle \mathbf{w}, x_i \rangle + b)}) \quad (5.6)$$

Il vantaggio legato all'uso della *logistic loss function* è che la funzione è convessa rispetto a \mathbf{w} . Per cui il problema, sebbene non abbia una soluzione in forma chiusa, può essere risolto efficientemente attraverso l'uso di specifici algoritmi iterativi per la minimizzazione di funzioni convesse.

5.2 SUPPORT VECTOR MACHINES

Le support vector machines (SVM) sono dei modelli di apprendimento supervisionato largamente usati nell'ambito della classificazione. Il loro obiettivo è quello di individuare un iperpiano nel feature space in grado di separare la classe positiva da quella negativa e contemporaneamente massimizzare il margine di separazione, definito come la minima distanza tra uno dei punti del training set e l'iperpiano stesso. Attraverso le SVM, si cerca quindi di dividere lo spazio delle feature in due semispazi, ciascuno contenente le osservazioni di una delle due classi, ponendo la condizione di trovare la soluzione più robusta possibile. Prendendo come esempio la figura (5.2) a pagina seguente, è facile intuire come, tra tutti i diversi iperpiani (in questo caso rette) in grado di separare la classe dei pallini blu da quella dei quadrati rossi, quello individuato sia quello più robusto. Una perturbazione dei dati, infatti, non inficerebbe sulla loro corretta classificazione. Al contrario, iperpiani posizionati vicino ai margini definiti dalle linee tratteggiate sono più soggetti a cambiamenti della classificazione dei dati del training set a causa di una loro perturbazione. I punti più vicini all'iperpiano sono chiamati support vector e lo definiscono in maniera univoca. Algebricamente l'iperpiano ottimo definito dai parametri (\mathbf{w}, b) si ottiene risolvendo il problema:

$$\underset{(\mathbf{w}, b)}{\operatorname{argmin}} \|\mathbf{w}\|^2 \quad \text{s.t.} \quad \forall i, y_i(\langle \mathbf{w}, x_i \rangle + b) \geq 1 \quad (5.7)$$

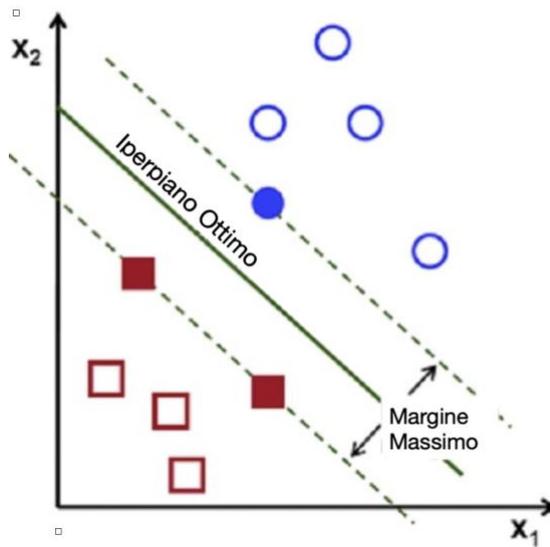


Figura 5.2: Esempio di Hard-SVM.

Nella formula appena riportata viene posto come costrizione che la soluzione trovata debba classificare correttamente tutte le osservazioni del training set. In questi casi si parla di Hard-SVM. Nella realtà, tuttavia, questa condizione è difficilmente realizzabile, soprattutto in presenza di problemi complessi come quello della predizione delle ipoglicemie, dove i dati a disposizione non sono linearmente separabili e presentano parziale sovrapposizione nel feature space. Per gestire questo problema è necessario modificare le condizioni di ottimizzazione in (5.7) e adottare un approccio più “morbido”, chiamato Soft-SVM. La differenza principale tra l’hard-SVM e il Soft-SVM è che in quest’ultimo la condizione $y_i(\langle \mathbf{w}, \mathbf{x}_i \rangle + b) \geq 1, \forall i$ viene rilassata consentendo ad alcune osservazioni del training set di violare tale restrizione, come visibile nella figura successiva.

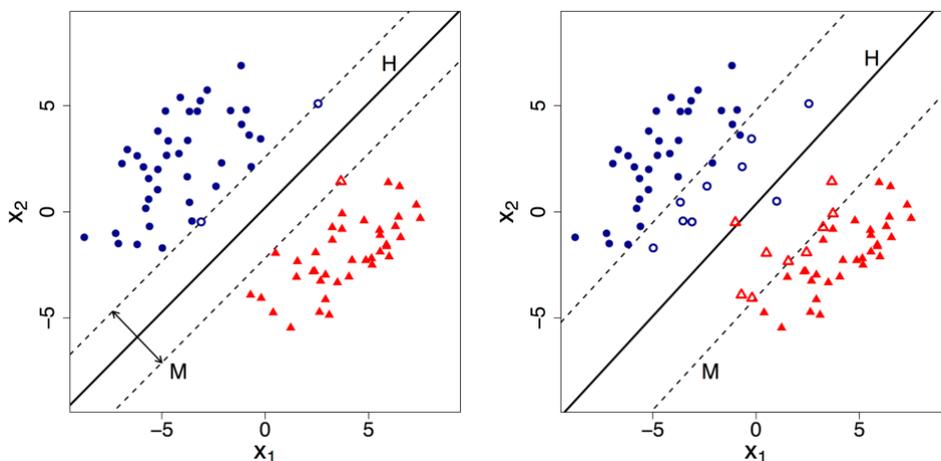


Figura 5.3: Hard-SVM (a sinistra) vs Soft-SVM (a destra).

A livello implementativo questo si traduce nell'introduzione di alcune variabili non negative, chiamate anche *slack variables*. Le slack variables ($\xi_1, \xi_2, \dots, \xi_m$), forniscono una misura di quanto può essere violata la restrizione posta nell'Hard-SVM. Il problema di ottimizzazione definito in (5.7) diventa quindi:

$$\underset{\mathbf{w}, b, \xi}{\operatorname{argmin}} \left(\lambda \|\mathbf{w}\|^2 + \frac{1}{m} \sum_{i=1}^m \xi_i \right) \quad \text{s.t.} \quad \forall i, y_i(\langle \mathbf{w}, \mathbf{x}_i \rangle + b) \geq 1 - \xi_i, \quad \xi_i \geq 0 \quad (5.8)$$

Dove λ è un iperparametro non negativo, inserito per gestire il trade-off tra l'incremento dei margini e il soddisfacimento della corretta classificazione dei dati del training set. Per fornire un ulteriore chiarimento è sufficiente dire che per valori di λ molto piccoli il costo delle violazioni diventa predominante e il comportamento del Soft-SVM sarà simile a quello dell'hard-SVM (margini stretti e poche violazioni), mentre per λ molto grandi avverrà esattamente l'opposto (margini ampi e molte violazioni).

Si è detto che per la classificazione tramite Support Vector Machines è necessario identificare un iperpiano che separi quanto meglio possibile la classe positiva da quella negativa. Tuttavia, non sempre questo è possibile. Si veda per esempio il caso riportato in figura (5.4). I dati rappresentati nello spazio bidimensionale originale non sono in alcun modo linearmente separabili e la ricerca dell'iperpiano (in questo caso retta) porterebbe a risultati del tutto insoddisfacenti. Tuttavia, è possibile ovviare a questo problema mappando i dati in uno spazio a dimensionalità superiore, dove essi risultano linearmente separabili. Questa procedura da un lato risolve il problema della corretta classificazione dei dati, ma dall'altro ne introduce di nuovi prettamente di carattere computazionale. Mappare i dati in uno spazio a dimensionalità maggiore non solo incrementa la richiesta di memoria, ma anche il carico computazionale necessario per l'identificazione dell'iperpiano.

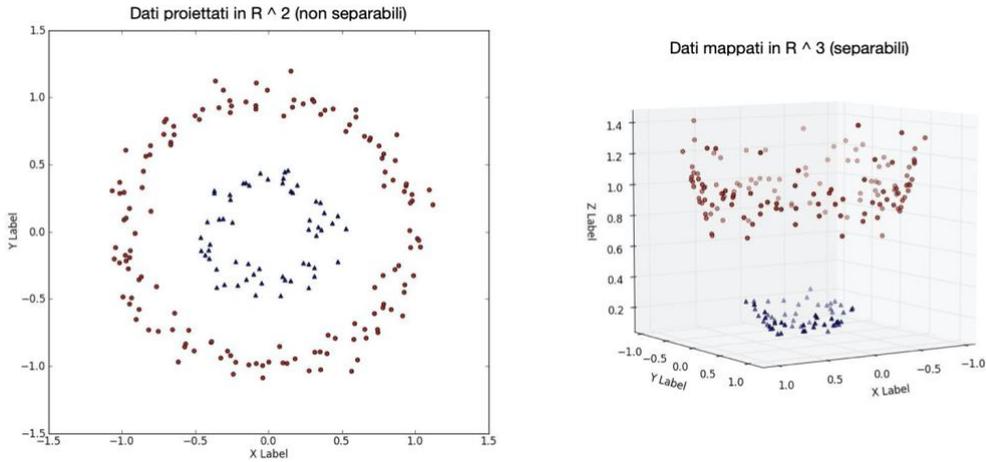


Figura 5.4: Dati nel feature space originale (a sinistra) mappati nello spazio tridimensionale secondo la funzione $\psi(\mathbf{x}) = (x_1, x_2, \sqrt{x_1^2 + x_2^2})$. Si noti che in tale spazio i dati sono linearmente separabili.

Per ovviare a questo problema si ricorre ad una procedura particolare denominata kernel trick. Per capire di cosa si tratta è necessario descrivere prima una delle interessanti proprietà delle support vector machines. Finora è stato detto che per identificare l'iperpiano ottimo è necessario risolvere il problema di ottimizzazione (5.8). Tuttavia, è possibile percorrere un'altra strada avvalendosi del principio di dualità e affrontando il problema di ottimizzazione duale:

$$\min_w \max_{\alpha \in R^m: \alpha \geq 0} \left(\frac{1}{2} |w|^2 + \sum_{i=1}^m \alpha_i (1 - y_i \langle w, x_i \rangle) \right) \quad (5.9)$$

Omettendo la dimostrazione, comunque consultabile nel testo “understanding machine learning: from theory to algorithms” [30], è possibile verificare che tra i due problemi sussiste una dualità forte. In altre parole, è possibile trovare la soluzione della (5.8) risolvendo la (5.9). Senza inoltrarsi in lunghe procedure matematiche si può dimostrare che la (5.9) può essere riscritta nella forma:

$$\max_{\alpha \in R^m: \alpha \geq 0} \left(\sum_{i=1}^m \alpha_i - \frac{1}{2} \sum_{i=1}^m \sum_{j=1}^m \alpha_i \alpha_j y_i y_j \langle x_i, x_j \rangle \right) \quad (5.10)$$

Una volta individuato i valori di α , i coefficienti dell'iperpiano si estrapolano dalla semplice relazione:

$$w = \sum_{i=1}^m \alpha_i y_i x_i \tag{5.11}$$

L'equazione (5.10) ha l'interessante proprietà di dipendere esclusivamente dal prodotto scalare dei campioni del training set. Questo aspetto può essere efficacemente sfruttato per risolvere i problemi legati alla maggiore complessità computazionale introdotta dalla mappatura delle feature in un nuovo spazio a dimensionalità maggiore. Invece di mappare esplicitamente ogni campione del training set per risolvere il problema primario, è sufficiente definire una funzione per il calcolo del prodotto scalare tra due osservazioni mappate, così da risolvere il problema duale. Tale funzione è chiamata Kernel function e si definisce come $K(\mathbf{x}_i, \mathbf{x}_j) = \langle \psi(\mathbf{x}_i), \psi(\mathbf{x}_j) \rangle$. Per poter essere applicato il kernel deve soddisfare il teorema di Mercer, che afferma che la matrice di Gram G , con $G_{i,j} = K(\mathbf{x}_i, \mathbf{x}_j)$, deve essere positiva semidefinita. Tra i kernel più comuni che soddisfano tale teorema si citano:

- Kernel lineare: $K(\mathbf{x}_i, \mathbf{x}_j) = \langle \mathbf{x}_i, \mathbf{x}_j \rangle$;
- Kernel polinomiale di grado d : $K(\mathbf{x}_i, \mathbf{x}_j) = (1 + \gamma \langle \mathbf{x}_i, \mathbf{x}_j \rangle)^d$;
- Radial Basis Function (RBF) Kernel: $K(\mathbf{x}_i, \mathbf{x}_j) = e^{-\gamma \|\mathbf{x}_i - \mathbf{x}_j\|^2}$.

Dove d e γ sono degli iperparametri che determinano rispettivamente il grado del polinomio e il peso di ogni singola istanza.

5.3 DECISION TREE

Gli alberi decisionali (decision tree) sono dei modelli di apprendimento supervisionato utilizzati sia nell'ambito della classificazione che in quello della regressione. Si tratta di modelli molto intuitivi, basati sulla costruzione di grafi decisionali che mirano a predire la classe o il valore y del target space viaggiando da un nodo radice fino ad un nodo foglia contenente il valore da assegnare. Nell'ambito della classificazione binaria, il tipo di classificazione scelto per gli scopi della tesi, ogni nodo interno dell'albero divide ricorsivamente il domain space in due parti sulla

base di predefinite regole di divisione. Una di queste, sicuramente la più popolare, prevede di scegliere quale dei due percorsi del nodo percorrere tramite sogliatura di una singola feature.

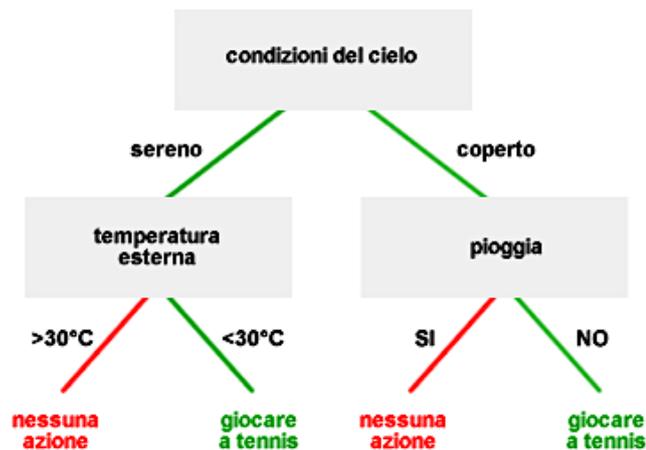


Figura 5.5: Esempio di albero di decisione binario.

Per l'addestramento degli alberi decisionali sono stati proposti numerosi algoritmi euristici, basati sulla loro graduale costruzione tramite ottimizzazioni locali eseguite durante l'identificazione dei nodi. Affrontarli tutti, o anche solo i più comuni, sarebbe inutile ed esterno ai fini della tesi. Pertanto, ci si limita a descrivere l'algoritmo CART per classificazione e regressione tramite strutture ad albero, utilizzato dal software Matlab® per l'identificazione di tali modelli. Gli step dell'algoritmo sono i seguenti:

1. Comincia con tutti i dati ed esamina tutte le possibili divisioni binarie su ogni feature;
2. Seleziona e impone la divisione migliore scelta in base ad un criterio di ottimizzazione (Indice di diversità di Gini, devianza statistica);
3. Ripeti ricorsivamente lo step precedente per ognuno dei due nodi figli.

La divisione ricorsiva termina quando viene soddisfatta una delle seguenti condizioni:

- Il nodo è puro, cioè osservazioni di una sola classe;
- Il nodo contiene un numero di osservazioni minore di un valore predefinito (in Matlab è indicato dalla variabile `MinParentSize`);
- Il nodo produce figli aventi un numero di osservazioni minore di un valore predefinito (in Matlab è indicato dalla variabile `MinLeafSize`);
- È stato raggiunto il massimo numero di nodi (in Matlab è indicato dalla variabile `MaxNumSplits`).

Un problema tipico di questi algoritmi è legato alla loro tendenza di restituire alberi decisionali molto profondi, che spesso soffrono di overfitting. Per limitare questo problema si è soliti operare su due fronti. Il primo prevede di intervenire durante la fase di training ottimizzando i parametri di terminazione degli split, mentre il secondo prevede di attuare un'operazione di *pruning* (potatura) una volta terminato l'algoritmo, in modo da ridurre la profondità dell'albero e garantire un giusto compromesso tra complessità del modello e bontà del fit.

5.4 RANDOM FOREST

Un problema comune per i decision tree è la tendenza all'overfitting. Per quanto si possa limitare il problema con l'ottimizzazione dei parametri e il successivo pruning, eliminarlo è molto difficile. Una strategia per ridurre il rischio di overfitting è quella di applicare i metodi di ensemble. I metodi di ensemble sono delle tecniche di classificazione o regressione che mirano a combinare i risultati di molti modelli deboli per ottenere una predizione più affidabile e, di conseguenza, prestazioni migliori. Tra i metodi di ensemble più conosciuti si descrive quello della Random Forest, introdotto da Breiman nel 2001 [31]. La Random Forest consiste in una collezione di alberi decisionali, ognuno dei quali addestrato su un training set diverso, ricavato da quello originale. Nel software Matlab® le random forest vengono costruite tramite metodo bootstrap aggregating, soprannominato Bagging. Nel bagging, come riportato nello schema in figura (5.6), ogni albero viene addestrato su un training set ottenuto tramite campionamento uniforme e con rimpiazzamento del training set originale. Le predizioni di un generico dato $\mathbf{x} \in X$ avvengono per votazione, cioè associando la classe predetta maggiormente dai diversi alberi.

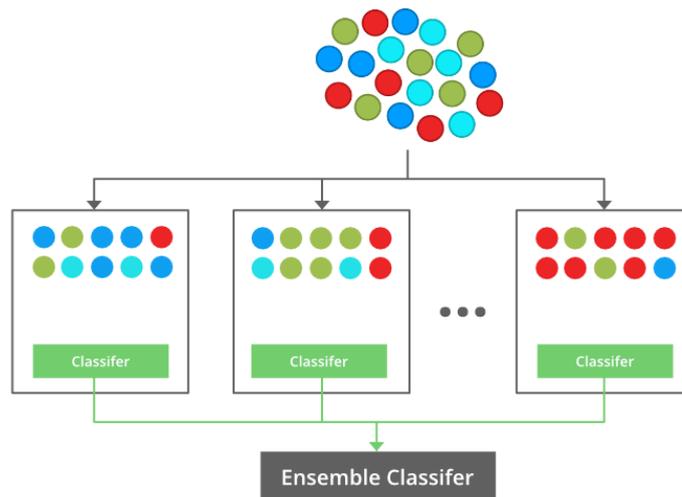


Figura 5.6: Schema del bagging per il Random Forest.

5.5 LOGISTIC MODEL TREES

L'ultima classe di modelli scelta per la predizione delle ipoglicemie è quella dei Logistic Model Tree. Si tratta di alberi decisionali per la classificazione aventi come foglie, al posto della classe predetta, dei modelli logistici. Nella trattazione del modello di Landwehr [32] vengono proposti due metodi diversi per la costruzione dei logistic model tree. Il primo, più semplice e immediato, prevede prima di costruire l'albero decisionale e poi, in un secondo momento, i vari modelli logistici da inserire nelle foglie, costruiti utilizzando solo le osservazioni del training set che cadono nelle foglie stesse. L'altro, più complesso, prevede l'applicazione dell'algoritmo di logit boost.

CAPITOLO 6

PROCEDURE E RISULTATI DELLA CLASSIFICAZIONE

In questo capitolo verranno presentati e commentati i dati raccolti durante la valutazione delle prestazioni delle metodologie di classificazione analizzate. Come per la sezione relativa all'analisi di correlazione, prima di riportare tutti i risultati ottenuti verranno illustrate le decisioni prese in merito alle modalità di preparazione del dataset (processing e divisione training/test) e i criteri di valutazione delle prestazioni dei modelli. Per l'addestramento dei modelli, sono stati utilizzati il statistic and machine learning toolbox e il parallel computing toolbox del software Matlab®.

6.1 ASSEGNAZIONE DELLE CLASSI

Prima di illustrare le procedure di elaborazione dei dati pre-processati è necessario descrivere le modalità di assegnazione delle classi. Contrariamente alle tipiche procedure di predizione delle ipoglicemie, l'obiettivo degli algoritmi proposti non è quello di predire l'esatto valore di glucosio al tempo definito dall'orizzonte di predizione, ma di individuare, tramite classificazione binaria, le osservazioni che si trovano vicine ad un evento di ipoglicemia. Questo aspetto è molto importante, perché garantisce una maggiore libertà riguardo alla scelta dei criteri di selezione delle classi positive e negative, che non deve per forza essere vincolata alla situazione del tracciato al tempo esatto $t_{windowend} + PH$. Per chiarire questo concetto, si ponga come esempio un'osservazione temporalmente localizzata 5 minuti prima di una ipoglicemia della durata di 15 minuti. Se la sua classificazione fosse basata solo sulla situazione all'orizzonte di predizione, che si può supporre essere pari a 30 minuti, la classe a cui apparterebbe tale osservazione sarebbe quella negativa, perché in quell'istante il soggetto non sarà in ipoglicemia, ma ne sarà appena uscito. Tuttavia, è logico pensare che in questa situazione sia meglio assegnare la classe positiva, visto che all'interno di quei 30 minuti è presente un evento di ipoglicemia completo. Tenendo conto di quanto appena detto sono stati utilizzati due metodi di assegnazione delle classi, chiamati per semplicità "a PH esatto" e "a PH misto", di cui si spiegheranno criteri e utilizzi nelle prossime sottosezioni.

6.1.1 Classi PH esatto

L'assegnazione delle classi secondo PH esatto è molto intuitiva e di facile implementazione. Si tratta di assegnare all'osservazione considerata la classe positiva nel caso in cui l'acquisizione CGM al tempo $t_{windowend} + PH$ sia all'interno di un evento di ipoglicemia, viceversa se non lo è. Un modo intuitivo per capire quale classe assegnare è quello di rispondere alla domanda "fra ESATTAMENTE PH minuti il soggetto sarà in ipoglicemia?". In realtà, come si riporterà in seguito, il sottocampionamento a cui sono stati sottoposti i dati durante la fase di processing permette di risolvere il problema precedentemente esposto, eliminando tutte le classi con associazioni dubbie. Si precisa, inoltre, che questo metodo è stato applicato solo in prima analisi nei modelli logistici con il preciso intento di individuare un set di parametri da utilizzare per gli sviluppi successivi. Bisogna tenere conto, infatti, che per valutare correttamente tutti i modelli proposti è necessario delineare un set di parametri tra gli orizzonti di predizione, le grandezze delle finestre di estrazione e i metodi di bilanciamento delle classi. Sicuramente i risultati dell'analisi di correlazione forniscono un buon punto di partenza per la loro scelta, ma non sono sufficienti. Per questo si è deciso di raccogliere ulteriori dati prima di procedere con le reali valutazioni degli algoritmi di predizione.

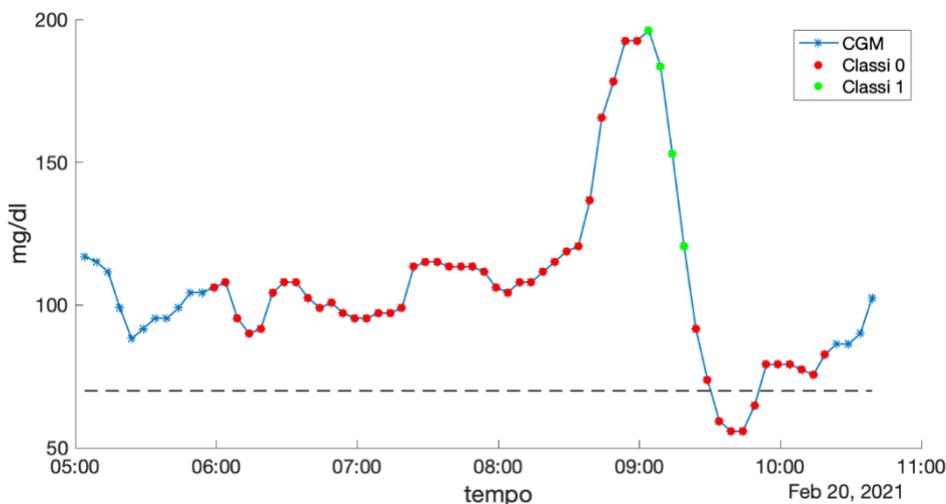


Figura 6.1: Assegnazione classi secondo PH esatto. In figura si è usato $PH=30$ [min]. Si ricorda che per assegnare la classe ad un campione è sufficiente valutare lo stato della glicemia PH minuti dopo.

6.1.2 Classi PH misto

L'assegnazione delle classi secondo PH misto è un metodo meno intuitivo del precedente ma sicuramente più adatto per lo sviluppo di algoritmi di predizione delle ipoglicemie post-ariatriche. Si tratta di assegnare all'osservazione considerata la classe positiva se almeno una tra le successive acquisizioni CGM fino a quella al tempo $t_{windowend} + PH$ è all'interno di un evento di ipoglicemia, altrimenti negativa. Anche in questo caso un modo intuitivo per capire quale classe assegnare è quello di rispondere alla domanda: "ENTRO PH minuti il soggetto sarà in ipoglicemia?". Questo criterio non presenta i problemi legati all'assegnazione dubbia di alcune classi e pertanto è stato utilizzato nelle fasi di implementazione delle metodologie predittive da valutare. Inoltre, presenta delle similarità con il metodo di valutazione delle prestazioni degli algoritmi descritto nel paragrafo (6.5), rendendolo ancora più idoneo all'applicazione rispetto a quello a PH esatto.

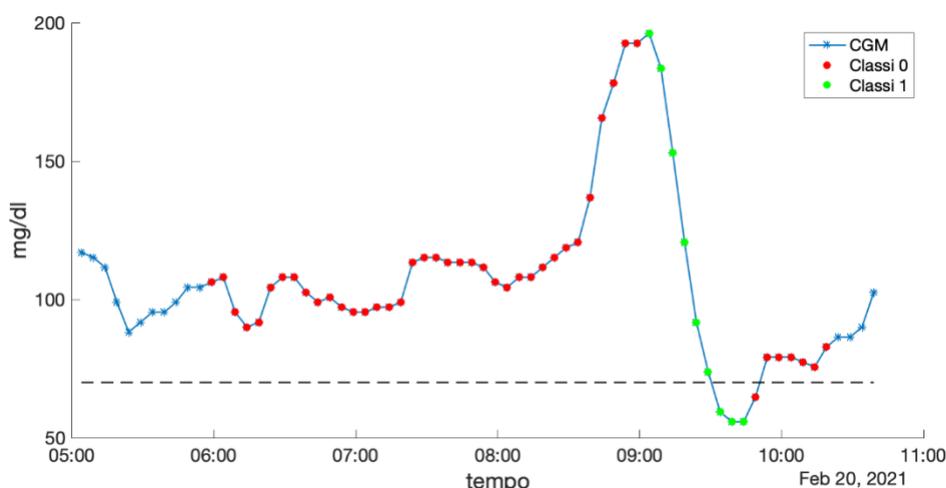


Figura 6.2: Assegnazione classi secondo PH misto. In figura si è usato $PH=30$ [min]. Si noti come rispetto al metodo precedente tutte le osservazioni che portano all'ipoglicemia siano etichettate come positive.

6.2 PROCESSING DEI DATI PER LA CLASSIFICAZIONE

Anche per lo sviluppo degli algoritmi di predizione delle ipoglicemie si è reso necessario sottoporre i dati pre-processati ad una ulteriore fase di elaborazione. Lo scopo di questo passaggio è quello di creare un insieme di coppie $(\mathbf{x}_i, y_i) \in X \times Y$, con \mathbf{x}_i vettore di metriche e y_i classe corrispondente, con cui addestrare e valutare i vari modelli predittivi.

Sebbene questa fase sia simile a quella descritta nel paragrafo (4.1) per l'analisi di correlazione, si preferisce riportare in maniera lineare tutti le varie fasi del processing, così da non creare ambiguità e risultare quanto più chiari possibile. La prima fase del processing consiste nella creazione dei vettori di osservazioni su cui estrarre le metriche, di cui si riportano i vari passaggi eseguiti sequenzialmente:

1. Fissare la grandezza della finestra di estrazione, l'orizzonte di predizione (PH) e il criterio di assegnazione delle classi;
2. Scegliere una delle giornate ottenute dopo la fase di pre-processing;
3. Salvare i primi N_w campioni di una giornata nella prima riga di una matrice X ;
4. Salvare la classe del campione posizionato alla fine della finestra al primo indice di un vettore Y ;
5. Salvare il numero o l'identificativo del soggetto da cui sono stati estratti i dati al primo posto di un vettore ID ;
6. Traslare la finestra un campione in avanti ed estrarre nuovamente i rispettivi N_w campioni contenuti al suo interno, la relativa classe e l'ID del paziente. Salvarli agli indici successivi delle variabili corrispondenti;
7. Ripetere lo step cinque fino al raggiungimento dell'ultimo campione su cui è possibile associare una classe;
8. Ripetere i passaggi dal tre al sette per ognuna delle restanti giornate presenti nel dataset.

Si noti come in questa fase sia richiesto di conservare il numero o l'identificativo del soggetto dal quale proviene ogni osservazione estratta. Tale informazione sarà indispensabile, infatti, per la successiva partizione dell'insieme in Training e Test set.

In seguito all'estrazione dei vettori di acquisizioni CGM, si è svolto un sottocampionamento preliminare per eliminare tutte le osservazioni non necessarie e ridurre parzialmente l'enorme sbilanciamento fra le classi. Il sottocampionamento prevede l'esecuzione sequenziale dei seguenti passaggi:

1. Eliminare tutte le osservazioni in cui il campione estratto alla fine della finestra è localizzato all'interno di un evento di ipoglicemia;
2. Se il criterio di assegnazione delle classi scelto è "a PH esatto", conservare solo le osservazioni positive in cui la fine della finestra di estrazione è posizionata PH minuti prima dell'inizio di un evento di ipoglicemia ed eliminare tutte quelle successive ottenute tramite traslazione della finestra fino all'inizio dell'evento stesso. Se il criterio di assegnazione

delle classi è a PH misto, conservare semplicemente tutte le classi positive rimaste dopo il punto uno;

3. Conservare le quattro (o quelle disponibili) classi negative precedenti a quella positiva ottenuta posizionando la finestra di estrazione PH minuti prima dell'inizio di un evento di ipoglicemia;
4. Conservare, delle classi negative rimaste, una ogni quattro procedendo in ordine, per cui senza alcuna estrazione casuale;
5. Eliminare tutti i doppioni tra le classi negative.

A questo punto dalle osservazioni rimaste vengono estratte le metriche selezionate dall'analisi di correlazione.

6.3 SUDDIVISIONE IN TRAINING E TEST SET

Come in tutte le applicazioni di machine learning, per poter addestrare e valutare le prestazioni dei modelli descritti nel capitolo cinque, è necessario dividere l'insieme delle osservazioni in due sottoinsiemi disgiunti e indipendenti chiamati training e test set. Un approccio usuale per la rapida esecuzione della procedura è basato sulla semplice partizione casuale delle osservazioni contenute nel dataset. Applicare una metodologia di questo tipo, tuttavia, non è possibile, poiché la presenza in entrambi gli insiemi di osservazioni provenienti da uno stesso soggetto potrebbe introdurre dei bias sulla stima dei parametri legati alle prestazioni dei modelli. Per eliminare qualsiasi forma di correlazione tra il training e test set si è quindi scelto di dividere il dataset allocando tutte le osservazioni di un soggetto in uno solo dei due insiemi. Per l'individuazione dei soggetti da collocare in una delle due partizioni, sono state applicate le seguenti condizioni:

- Il Test set deve contenere al massimo sette soggetti (18 % del numero totale);
- I pazienti che hanno tempi di monitoraggio molto lunghi, un numero di ipoglicemie troppo elevato, o nessuna ipoglicemia, devono sempre essere inseriti nel training set;
- La partizione deve garantire che, nell'insieme di osservazioni ottenuto dopo sottocampionamento, vengano soddisfatti i seguenti requisiti di prevalenza e rapporto tra le classi:

$$\frac{C1_{train}}{C0_{train}} - \frac{C1_{test}}{C0_{test}} < 0.001 \quad (6.1)$$

$$0.195 < \frac{C0_{test} + C1_{test}}{C0_{train} + C1_{train}} < 0.205 \quad (6.2)$$

Una volta stabilita la partizione dei soggetti, i due insiemi vengono creati ponendo:

- nel Training set tutte le osservazioni dei soggetti scelti provenienti dall'insieme sottocampionato;
- nel Test set tutte le osservazioni ottenute dai soggetti scelti ripetendo il processing dei dati senza eseguire alcun sottocampionamento.

Il motivo per cui il test set viene così costruito è legato alla metodologia scelta per la valutazione delle prestazioni degli algoritmi, che prevede di simulare una reale applicazione dell'algoritmo sulle giornate acquisite. Per questo motivo, è indispensabile disporre di ogni singola osservazione ricavata dai soggetti del test set.

Le condizioni poste per la partizione del dataset permettono di identificare un insieme di differenti suddivisioni in training e test set. Questo deriva dalla necessità di stimare la variabilità del potere predittivo dei modelli in relazione ad un cambiamento delle osservazioni del training set su cui è costruito. Stima ottenibile tramite l'applicazione di una simulazione di Monte Carlo. In una simulazione di Monte Carlo il processo di addestramento e di valutazione dell'algoritmo viene ripetuto per ogni partizione individuata, riportando poi una sintesi della distribuzione dei risultati ottenuti tramite il calcolo di alcuni indici statistici di posizione e di dispersione. In questo modo viene eliminata e quantificata l'influenza legata alla scelta dei soggetti da inserire nel test set e viene fornita una stima più realistica della capacità predittiva del modello.

6.4 BILANCIAMENTO DELLE CLASSI

Nonostante il sottocampionamento preliminare eseguito durante la fase di processing dei dati elimini una buona parte delle classi negative, lo sbilanciamento delle classi rimane comunque molto alto. Tale aspetto può essere una fonte di problemi per l'addestramento dei classificatori poiché maggiore è lo sbilanciamento delle classi, maggiore è la tendenza di quest'ultimi ad

ignorare la classe minoritaria per garantire una maggiore accuratezza. Per risolvere il problema dello sbilanciamento è possibile avvalersi di diverse strategie, applicabili sia singolarmente che in maniera combinata. Tra le più comuni si citano:

- Il sovradimensionamento della classe minoritaria tramite sintesi di nuove osservazioni, per esempio tramite applicazione dell'algoritmo SMOTE;
- Il sottodimensionamento della classe maggioritaria tramite eliminazione di osservazioni poco rilevanti.

Quest'ultimo approccio è stato scelto per gestire il problema dello sbilanciamento nel dataset in esame. Sebbene la sintesi di nuove classi positive sia sicuramente un approccio efficace per la risoluzione dello sbilanciamento in molti problemi di classificazione, si è preferito evitarlo per non mescolare dati sintetici a quelli ottenuti dal trial clinico. Per bilanciare le classi tramite sottodimensionamento della classe maggioritaria sono stati valutati due diversi approcci, uno semplice e uno più complesso:

- **Bilanciamento casuale:** le classi negative vengono ridotte al numero di quelle positive tramite eliminazione casuale;
- **One Sided Selection (OSS):** è un metodo di sottocampionamento della classe maggioritaria proposto da Kubat e Matwin [33]. L'algoritmo prevede l'esecuzione dei seguenti step, riportati e tradotti dal paper originale:
 1. Si indichi con S il training set originale;
 2. Inizialmente, C, l'insieme sottocampionato, contiene tutte le osservazioni positive di S e una osservazione negativa casualmente selezionata;
 3. Si effettui la classificazione 1-NN di S usando le osservazioni di C e si comparino le etichette assegnate con quelle originali. Si spostino poi in C tutte le osservazioni di S classificate erroneamente;
 4. Si rimuovano da C tutte le osservazioni negative partecipanti nel Tomei Link. Si precisa che due osservazioni A e B partecipano nel Tomei Link quando sono di classi differenti e ognuna è il nearest neighbor dell'altra.

L'insieme sottocampionato tramite OSS è consistente con il training set originale e privo di osservazioni borderline o rumorose.

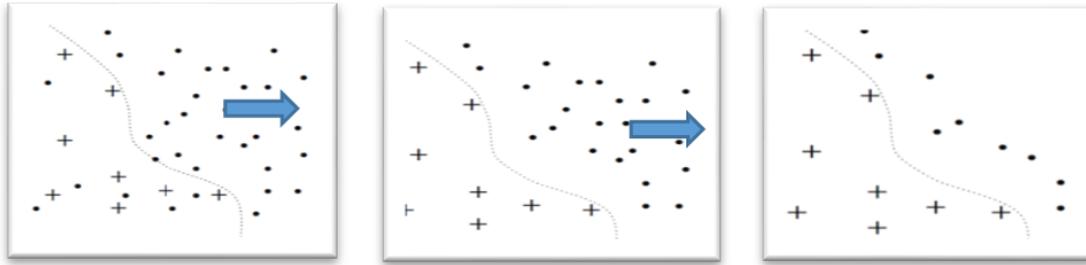


Figura 6.3: One-Sided Selection (OSS) applicato su un dataset di esempio.

6.5 VALUTAZIONE DELLE PRESTAZIONI DEGLI ALGORITMI

Riguardo alla valutazione della capacità degli algoritmi di predire gli eventi di ipoglicemia, si è scelto di seguire parzialmente un approccio già proposto per il medesimo scopo e applicato su dati provenienti dalla popolazione diabetica [34]. In tale approccio le osservazioni di ogni giornata del Test Set vengono classificate in maniera sequenziale, considerando come un allarme ogni predizione positiva. Inoltre, per evitare il lancio di più allarmi consecutivi, viene applicato un silenziamento di 20 minuti (4 osservazioni) successivamente ad ogni predizione positiva. La valutazione della correttezza degli allarmi è stata fatta considerando una detection window (DW) di 30 minuti e assegnando le seguenti etichette:

- Vero Positivo (TP): l'allarme viene lanciato dai 5 ai 30 (DW) minuti prima dell'inizio di un evento di ipoglicemia;
- Falso Positivo (FP): l'allarme viene lanciato oltre 30 (DW) minuti prima dell'inizio di un evento di ipoglicemia;
- Falso Negativo (FN): nessun allarme viene lanciato tra i 5 e i 30 (DW) minuti prima dell'inizio di un evento di ipoglicemia;
- Non Etichettabile (NC): l'allarme viene lanciato all'interno di un evento di ipoglicemia e quindi in ritardo. Si precisa che, nella popolazione post-bariatrica possono manifestarsi fenomeni di roller-coasting (sali-scendi del glucosio) con conseguenti ipoglicemie anche molto ravvicinate. Nel caso in cui un allarme sia stato lanciato all'interno di un evento di ipoglicemia ma contemporaneamente si trovi tra i 5 e i 30 minuti prima dell'inizio dell'evento successivo, esso verrà comunque considerato NC. Inoltre, considerando che il tempo di silenziamento è minore della detection window, in presenza di due allarmi Veri Positivi riferiti alla stessa ipoglicemia si è considerato come TP quello lanciato prima e NC l'altro.

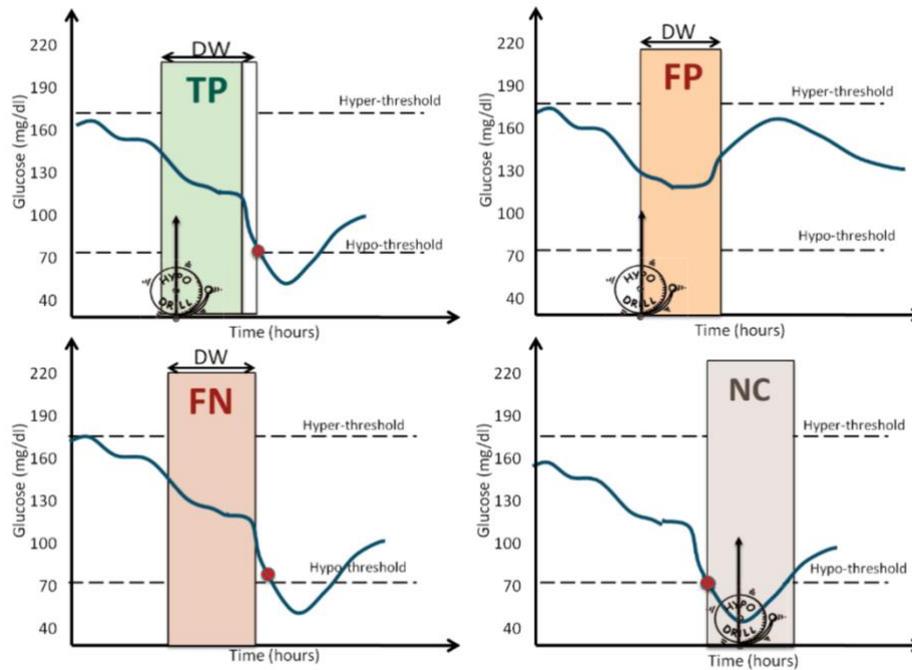


Figura 6.3: Esempi di Vero Positivo (in alto a sinistra), Falso Positivo (in alto a destra), Falso Negativo (in basso a sinistra) e Non Etichettabile (in basso a destra). (Tratto da [34])

Una volta etichettati tutti gli allarmi secondo i criteri appena definiti, sono state calcolate le seguenti metriche:

$$\text{precisione} = \frac{TP}{TP + FP} \quad (6.3)$$

$$\text{sensibilità} = \frac{TP}{TP + FN} \quad (6.4)$$

$$\text{F1-score} = 2 \cdot \frac{\text{precisione} \cdot \text{sensibilità}}{\text{precisione} + \text{sensibilità}} \quad (6.5)$$

Si noti che non è prevista l'identificazione dei Veri Negativi né tantomeno il calcolo di metriche che li coinvolgono, come l'accuratezza o la specificità. Questo per via del forte sbilanciamento fra le classi che rende pressoché inutile la loro considerazione. Oltre alle metriche appena descritte si è calcolato anche i valori di media e mediana del Time Gain (TG), definito come la differenza temporale tra l'inizio dell'evento di ipoglicemia e l'allarme Vero Positivo lanciato.

6.6 RISULTATI

Definite tutte le scelte implementative attuate per lo sviluppo e la valutazione degli algoritmi, si può finalmente passare all'analisi dei risultati dei vari modelli testati. Per ogni metodologia esaminata sono state eseguite numerose simulazioni Monte Carlo variando, variando i seguenti parametri:

- Grandezza della Finestra di estrazione;
- Orizzonte di Predizione;
- Soglia di classificazione;
- Metodo di bilanciamento.

Vista l'enorme mole di dati raccolti si è preferito riportare solo i risultati ottenuti con soglia di ipoglicemia di 70 mg/dl, che si è verificato essere la soglia migliore. I risultati ottenuti scegliendo altre soglie (65, 60, 54 mg/dl) sono consultabili nella sezione A.2 dell'appendice. Si precisa inoltre che le tabelle riportano i valori medi delle metriche e tra parentesi la loro relativa deviazione standard, ad eccezione per la mediana del Time Gain, che è stata riportata insieme alle mediane dei range interquartili di ogni simulazione. I valori di PH e grandezza della finestra di estrazione riportati si riferiscono in minuti.

6.6.1 Logistic regression

Come anticipato nel paragrafo dedicato ai criteri di assegnazione delle classi, il modello logistico è stato scelto non solo come metodologia di valutazione, ma anche come modello su cui identificare il set di parametri da utilizzare sulle altre metodologie. Per questo motivo la trattazione dei risultati sarà suddivisa in due parti. Nella prima verranno riportate e commentate le performance ottenute successivamente all'uso del criterio di assegnazione delle classi a PH esatto, così da giustificare la scelta dei parametri utilizzati per le successive analisi. Nella seconda si riporteranno i risultati ottenuti con i parametri scelti e criterio di assegnazione delle classi a PH misto. Si precisa che per ogni modello identificato è stata individuata la soglia di classificazione migliore secondo il criterio massimizzazione dell'F1-score.

Logistic regression: PH esatto

In questa prima parte si è fissato il criterio di assegnazione delle classi a PH esatto e si è scelto di valutare un range molto ampio degli altri parametri, così da individuare:

- Un limite massimo di orizzonte di predizione da usare;
- Un insieme di grandezze di finestre da utilizzare;
- Il metodo di bilanciamento migliore.

Per facilitare la consultazione si separeranno i risultati ottenuti in base al metodo di bilanciamento, evitando di riportare i valori mediani del Time Gain.

PRESTAZIONI MODELLO LOGISTICO: PH ESATTO, NESSUN BILANCIAMENTO

Precisione				
PH \ finestra	5	10	15	20
30	0.524 (0.024)	0.482 (0.022)	0.391 (0.011)	0.327 (0.018)
45	0.527 (0.012)	0.468 (0.028)	0.421 (0.024)	0.315 (0.017)
60	0.528 (0.015)	0.464 (0.013)	0.397 (0.023)	0.312 (0.026)
90	0.461 (0.013)	0.470 (0.016)	0.380 (0.015)	0.331 (0.016)
120	0.453 (0.026)	0.472 (0.024)	0.392 (0.016)	0.306 (0.012)

Sensibilità				
PH \ finestra	5	10	15	20
30	0.676 (0.034)	0.650 (0.048)	0.625 (0.023)	0.495 (0.032)
45	0.707 (0.013)	0.663 (0.052)	0.527 (0.024)	0.505 (0.024)
60	0.667 (0.016)	0.655 (0.019)	0.539 (0.033)	0.476 (0.033)
90	0.654 (0.014)	0.614 (0.025)	0.547 (0.030)	0.453 (0.021)
120	0.623 (0.032)	0.556 (0.024)	0.548 (0.025)	0.500 (0.029)

F1-score				
PH \ finestra	5	10	15	20
30	0.589 (0.014)	0.552 (0.011)	0.481 (0.012)	0.393 (0.018)
45	0.604 (0.012)	0.546 (0.011)	0.467 (0.011)	0.387 (0.012)
60	0.590 (0.011)	0.543 (0.009)	0.456 (0.007)	0.376 (0.012)
90	0.541 (0.011)	0.532 (0.011)	0.448 (0.008)	0.382 (0.012)
120	0.523 (0.014)	0.510 (0.014)	0.456 (0.010)	0.379 (0.010)

Media Time Gain				
PH \ finestra	5	10	15	20
30	7.260 (0.422)	9.583 (0.326)	12.753 (0.423)	14.769 (0.394)
45	7.419 (0.272)	9.963 (0.340)	12.767 (0.506)	15.536 (0.511)
60	7.540 (0.308)	10.193 (0.288)	13.242 (0.462)	15.528 (0.646)
90	8.307 (0.279)	10.663 (0.278)	13.782 (0.485)	15.901 (0.686)
120	8.803 (0.421)	10.935 (0.506)	13.456 (0.057)	15.886 (0.523)

Tabelle 6.1: Prestazioni modello logistico con PH esatto e nessun bilanciamento.

Osservando attentamente i dati riportati nelle tabelle, si può notare come un aumento del PH porti ad una diminuzione delle performance del modello. Questo è un risultato atteso in quanto il sottocampionamento a cui è sottoposto il dataset prevede, per il PH esatto, la conservazione delle sole osservazioni positive associate all'inizio di un evento di ipoglicemia. Di conseguenza, ad un aumento di quest'ultimo corrisponde una perdita di correlazione tra la classe positiva e l'evento di ipoglicemia ad esso associato, con un conseguente calo del potere predittivo del modello. Anche la perdita di performance indotta dall'aumento della finestra di estrazione è da considerarsi un risultato atteso, se si considera che l'uso di finestre ampie può ridurre la diversificazione tra le classi dei valori di alcune delle feature estratte. L'aumento della media del Time Gain dovuto all'incremento dell'orizzonte di predizione, invece, non è da considerarsi come un fattore positivo, in quanto deve essere valutato insieme al calo della precisione, che evidenzia come per valori di PH alti il modello lanci allarmi senza discernimento.

PRESTAZIONI MODELLO LOGISTICO: PH ESATTO, BILANCIAMENTO CASUALE

Precisione				
PH \ finestra	5	10	15	20
30	0.496 (0.018)	0.445 (0.024)	0.364 (0.028)	0.332 (0.036)
45	0.490 (0.015)	0.447 (0.024)	0.394 (0.050)	0.351 (0.044)
60	0.481 (0.022)	0.448 (0.035)	0.394 (0.042)	0.337 (0.050)
90	0.482 (0.027)	0.475 (0.039)	0.408 (0.053)	0.357 (0.062)
120	0.524 (0.032)	0.490 (0.046)	0.432 (0.068)	0.390 (0.080)

Recall				
PH \ finestra	5	10	15	20
30	0.720 (0.027)	0.679 (0.037)	0.616 (0.037)	0.540 (0.039)
45	0.732 (0.025)	0.679 (0.032)	0.571 (0.043)	0.491 (0.038)
60	0.713 (0.040)	0.659 (0.034)	0.556 (0.046)	0.508 (0.036)
90	0.699 (0.055)	0.650 (0.063)	0.558 (0.057)	0.480 (0.047)
120	0.685 (0.047)	0.681 (0.062)	0.578 (0.065)	0.503 (0.046)

F1-score				
finestra \ PH	5	10	15	20
30	0.587 (0.008)	0.536 (0.012)	0.456 (0.019)	0.408 (0.024)
45	0.587 (0.008)	0.537 (0.016)	0.464 (0.043)	0.407 (0.032)
60	0.573 (0.010)	0.532 (0.023)	0.459 (0.034)	0.402 (0.032)
90	0.569 (0.022)	0.547 (0.040)	0.470 (0.049)	0.406 (0.046)
120	0.592 (0.434)	0.568 (0.044)	0.491 (0.059)	0.435 (0.058)

Media Time Gain				
finestra \ PH	5	10	15	20
30	7.737 (0.181)	9.821 (0.427)	12.626 (0.019)	13.814 (1.203)
45	7.975 (0.200)	10.086 (0.525)	12.321 (1.375)	13.798 (1.427)
60	8.132 (0.272)	9.960 (0.766)	12.284 (1.246)	13.836 (1.556)
90	8.230 (0.298)	9.546 (1.091)	11.929 (1.688)	13.952 (1.981)
120	7.974 (0.434)	9.147 (1.140)	11.278 (2.100)	12.733 (2.490)

Tabelle 6.2: Prestazioni modello logistico con PH esatto e metodo di Bilanciamento Casuale.

Il metodo di bilanciamento casuale ha la particolarità di creare modelli caratterizzati da una forte aggressività indotta dalla drastica riduzione del numero di osservazioni negative fornite in fase di training. Questo aspetto, riscontrabile tra l'altro nella riduzione della precisione e nell'aumento della sensibilità, viene mitigato dal processo di correzione della soglia di classificazione che, identificando valori di α molto grandi anche superiori allo 0.9, limita la generazione di un elevato numero di allarmi. Anche il silenziamento previsto dal metodo di valutazione delle performance contribuisce alla riduzione dell'alto numero di falsi positivi generati dal modello, impedendo di lanciare numerosi allarmi consecutivi. Tutto questo va a beneficio dell'F1-score, il parametro decisivo per il confronto delle varie metodologie, che non sembra subire cambiamenti significativi rispetto ai valori ottenuti dai modelli costruiti senza bilanciamento. In conclusione, non sembra esserci un reale vantaggio nell'uso di un criterio di bilanciamento simile, e per questo si è preferito scartarlo. Una riduzione così elevata del numero di classi negative in fase di training implica infatti la perdita di buona parte dell'informazione associata alla classe maggioritaria, senza contare che la casualità con cui vengono scelte quelle da conservare influisce negativamente sulla varianza dei modelli.

Come ulteriore conferma di quanto detto si riporta un esempio delle performance di modelli logistici ottenuti senza correzione della soglia di classificazione

Bilanciamento	precisione	Sensibilità	F1-score
Nessuno	0.559 (0.016)	0.502 (0.016)	0.529 (0.012)
Casuale	0.211 (0.005)	0.969 (0.003)	0.346 (0.006)

Tabella 6.3: Esempio prestazioni modelli logistici (PH= 5 min, Finestra= 30 min, soglia di classificazione fissa a 0.5). Si noti come l'assenza della correzione della soglia di classificazione induca un brusco calo delle performance per il bilanciamento casuale.

PRESTAZIONI MODELLO LOGISTICO: PH ESATTO, BILANCIAMENTO OSS

Precisione				
PH \ finestra	5	10	15	20
30	0.512 (0.025)	0.480 (0.017)	0.416 (0.016)	0.298 (0.026)
45	0.527 (0.014)	0.484 (0.016)	0.435 (0.027)	0.298 (0.024)
60	0.525 (0.014)	0.482 (0.014)	0.412 (0.015)	0.298 (0.025)
90	0.466 (0.025)	0.481 (0.018)	0.404 (0.020)	0.320 (0.024)
120	0.506 (0.025)	0.455 (0.027)	0.403 (0.025)	0.301 (0.019)

Sensibilità				
PH \ finestra	5	10	15	20
30	0.689 (0.039)	0.687 (0.032)	0.594 (0.035)	0.494 (0.045)
45	0.703 (0.018)	0.686 (0.015)	0.539 (0.027)	0.494 (0.033)
60	0.676 (0.016)	0.664 (0.021)	0.553 (0.021)	0.477 (0.035)
90	0.678 (0.039)	0.622 (0.030)	0.520 (0.029)	0.451 (0.028)
120	0.630 (0.027)	0.616 (0.031)	0.549 (0.035)	0.509 (0.028)

F1-score				
PH \ finestra	5	10	15	20
30	0.586 (0.015)	0.565 (0.014)	0.488 (0.013)	0.370 (0.021)
45	0.603 (0.013)	0.567 (0.012)	0.480 (0.014)	0.371 (0.017)
60	0.591 (0.012)	0.558 (0.011)	0.472 (0.009)	0.366 (0.018)
90	0.551 (0.011)	0.542 (0.012)	0.453 (0.009)	0.373 (0.018)
120	0.560 (0.014)	0.522 (0.014)	0.463 (0.012)	0.377 (0.016)

Media Time Gain				
PH \ finestra	5	10	15	20
30	7.311 (0.481)	9.852 (0.354)	12.956 (0.451)	15.963 (0.666)
45	7.309 (0.288)	9.762 (0.270)	13.149 (0.443)	16.258 (0.546)
60	7.552 (0.340)	9.857 (0.295)	13.293 (0.462)	16.387 (0.550)
90	8.279 (0.340)	10.415 (0.318)	13.443 (0.575)	16.335 (0.649)
120	8.081 (0.357)	10.842 (0.395)	13.337 (0.845)	16.376 (0.614)

Tabella 6.4: Prestazioni modello logistico con PH esatto e metodo di Bilanciamento OSS.

Simili risultati si ottengono con il metodo di bilanciamento One-Sided Selection, che non sembra portare vistosi cambiamenti rispetto all'assenza di bilanciamento. Questo fatto è legato ancora una volta alla capacità della correzione della soglia di classificazione e del metodo di valutazione delle performance di mascherare pregi e difetti delle varie tecniche di bilanciamento usate, rendendo difficile stabilire quale dei tre sia il migliore. Tuttavia, va considerato che la capacità dell'algoritmo One-Sided Selection di creare training set consistenti a dimensionalità minore favorisce la riduzione dei tempi di addestramento dei modelli più complessi e la possibilità di applicare strategie di ottimizzazione degli iper-parametri più efficaci. Inoltre, osservando la tabella 6.5, si può notare come senza la correzione della soglia di classificazione le prestazioni dei modelli siano migliori. Anche questo è un aspetto molto importante se si considera che alcune delle altre metodologie scelte non dispongono di soglie di classificazione da correggere.

Bilanciamento	precisione	Sensibilità	F1-score
Nessuno	0.559 (0.016)	0.502 (0.016)	0.529 (0.012)
OSS	0.547 (0.016)	0.584 (0.018)	0.565 (0.014)

Tabella 6.5: Esempio prestazioni modelli logistici
(PH= 5 min, Finestra= 30 min, soglia di classificazione fissa a 0.5).

Una volta analizzato attentamente i risultati appena riportati, si è quindi scelto di adoperare i seguenti parametri per l'implementazione dei vari modelli:

- Limite massimo di PH posto a 20 minuti;
- Grandezza Finestra di estrazione di 30, 45 e 60 minuti;
- Metodo di bilanciamento OSS.

Logistic Regression: PH misto

Stabiliti i parametri da utilizzare, si riportano quindi i risultati delle performance del modello logistico ottenuto tramite l'uso del criterio di assegnazione delle classi a PH misto, che d'ora in avanti sarà usato anche sugli altri modelli.

PRESTAZIONI MODELLO LOGISTICO

Precisione				Sensibilità			
PH finestra	5-10	5-15	5-20	PH finestra	5-10	5-15	5-20
30	0.542 (0.015)	0.478 (0.011)	0.467 (0.016)	30	0.675 (0.028)	0.697 (0.039)	0.654 (0.030)
45	0.513 (0.019)	0.478 (0.015)	0.465 (0.017)	45	0.694 (0.027)	0.685 (0.026)	0.641 (0.028)
60	0.503 (0.021)	0.488 (0.019)	0.454 (0.024)	60	0.701 (0.026)	0.660 (0.026)	0.612 (0.035)

F1-Score				Media Time Gain			
PH finestra	5-10	5-15	5-20	PH finestra	5-10	5-15	5-20
30	0.601 (0.012)	0.566 (0.012)	0.545 (0.012)	30	8.412 (0.327)	9.660 (0.350)	10.080 (0.196)
45	0.590 (0.010)	0.562 (0.009)	0.538 (0.010)	45	8.576 (0.328)	9.938 (0.269)	10.885 (0.247)
60	0.585 (0.011)	0.560 (0.010)	0.520 (0.008)	60	8.845 (0.366)	9.869 (0.324)	11.255 (0.376)

Mediana Time Gain			
PH finestra	5-10	5-15	5-20
30	5 [5 - 10]	10 [5 - 10]	10 [5 - 10]
45	5 [5 - 10]	10 [5 - 10]	10 [5 - 15]
60	5 [5 - 10]	10 [5 - 10]	10 [5 - 15]

Tabelle 6.6: Prestazioni modello logistico con PH misto e metodo di Bilanciamento OSS.

Anche in questo caso la variazione del criterio di assegnazione delle classi non sembra aver portato alcun giovamento nelle performance del modello. Ancora una volta la causa di ciò è da associarsi alla capacità della correzione della soglia di classificazione e del metodo di valutazione delle performance di mascherare i pregi e difetti dei criteri utilizzati. Osservando la tabella 6.7, dove sono riportate le metriche ottenute senza applicazione della correzione della soglia di classificazione, si può notare come l'uso del criterio di assegnazione delle classi a PH misto porti all'identificazione di un modello più aggressivo, che tuttavia non stravolge le performance del modello, come invece era accaduto per bilanciamento casuale. L'F1-score che ne risulta è leggermente più alto, anche se non in maniera statisticamente significativa.

Criterio PH	precisione	Sensibilità	F1-score
PH esatto	0.547 (0.016)	0.584 (0.018)	0.565 (0.014)
PH misto	0.462 (0.013)	0.765 (0.020)	0.576 (0.017)

Tabella 6.7: Confronto tra le prestazioni dei modelli logistici costruiti con diverso criterio di assegnazione delle classi (PH esatto = 5 min, PH misto = 5-10 min, grandezza finestra = 30 min, bilanciamento OSS).

6.6.2 Support Vector Machine

Per l'addestramento delle SVM si è utilizzato la funzione `fitcsvm` di Matlab[®]. Come in molte altre funzioni del pacchetto di machine learning, il software offre la possibilità di applicare l'ottimizzazione baesiana per stimare alcuni tra i principali iper-parametri del modello quali il tipo kernel, il costo delle violazioni e il coefficiente di scaling della matrice di Gram. Purtroppo, a differenza di quanto fatto per le successive metodologie, non è stato possibile eseguire tale procedura a causa dell'alta sovrapposizione delle classi, già riscontrabile consultando le performance dei modelli logistici, che ha reso eccessivamente lunghi i tempi di addestramento dei singoli modelli. Per ovviare a questo problema sono state valutate su alcune partizioni di riferimento diverse combinazioni dei tre parametri da individuare. Ciò che ne è risultato è che i parametri migliori per l'addestramento dei modelli SVM sono:

- Kernel lineare;
- Costo violazioni (parametro C della funzione `fitcsvm`) uguale a 1;
- Kernel scale (parametro omonimo della funzione `fitcsvm`) uguale a 1.

PRESTAZIONI SUPPORT VECTOR MACHINES

Precisione				Sensibilità			
finestra \ PH	5-10	5-15	5-20	finestra \ PH	5-10	5-15	5-20
	30	0.536 (0.013)	0.506 (0.014)		0.498 (0.017)	30	0.705 (0.034)
45	0.528 (0.020)	0.509 (0.027)	0.511 (0.032)	45	0.692 (0.042)	0.705 (0.042)	0.655 (0.060)
60	0.520 (0.019)	0.520 (0.022)	0.524 (0.017)	60	0.724 (0.049)	0.707 (0.045)	0.636 (0.020)

F1-Score				Media Time Gain			
finestra \ PH	5-10	5-15	5-20	finestra \ PH	5-10	5-15	5-20
	30	0.608 (0.013)	0.597 (0.013)		0.589 (0.013)	30	8.312 (0.249)
45	0.597 (0.010)	0.590 (0.009)	0.571 (0.011)	45	8.274 (0.338)	8.734 (0.478)	9.010 (0.637)
60	0.604 (0.011)	0.598 (0.010)	0.574 (0.007)	60	8.286 (0.253)	8.434 (0.352)	9.034 (0.240)

Mediana Time Gain			
finestra \ PH	5-10	5-15	5-20
	30	5 [5 - 10]	10 [5 - 10]
45	5 [5 - 10]	5 [5 - 10]	5 [5 - 10]
60	5 [5 - 10]	5 [5 - 10]	5 [5 - 10]

Tabelle 6.8: Prestazioni SVM (PH misto, metodo di Bilanciamento OSS, Kernel Lineare).

Dalle tabelle riportate non si riscontra un miglioramento della capacità predittiva del modello. Tuttavia, sembra che l'aumento del PH non penalizzi molto l'F1-score, consentendo l'identificazione di modelli con un migliore bilanciamento tra capacità predittiva e aggressività nella generazione degli allarmi. Si veda per esempio il modello ottenuto con finestra di 30 min e PH= [5-15].

6.6.3 Decision Tree

Per l'addestramento degli alberi è stata utilizzata la funzione `fitctree` di Matlab[®]. A differenza delle Support Vector Machine, questa volta è stato possibile eseguire l'ottimizzazione bayesiana dei parametri già citati nel paragrafo 5.3 durante la descrizione del CART implementato dal software.

PRESTAZIONI DECISION TREE

Precisione			
PH \ finestra	5-10	5-15	5-20
30	0.433 (0.017)	0.352 (0.017)	0.285 (0.136)
45	0.430 (0.024)	0.352 (0.012)	0.309 (0.014)
60	0.417 (0.019)	0.366 (0.020)	0.296 (0.015)

Sensibilità			
PH \ finestra	5-10	5-15	5-20
30	0.858 (0.027)	0.912 (0.021)	0.941 (0.014)
45	0.860 (0.024)	0.888 (0.020)	0.908 (0.017)
60	0.856 (0.021)	0.885 (0.015)	0.920 (0.019)

F1-Score			
PH \ finestra	5-10	5-15	5-20
30	0.575 (0.014)	0.507 (0.017)	0.437 (0.015)
45	0.573 (0.016)	0.504 (0.012)	0.461 (0.016)
60	0.561 (0.017)	0.517 (0.019)	0.448 (0.017)

Media Time Gain			
PH \ finestra	5-10	5-15	5-20
30	9.397 (0.417)	11.866 (0.512)	13.561 (0.430)
45	9.390 (0.459)	11.518 (0.396)	13.060 (0.367)
60	9.532 (0.443)	11.526 (0.527)	13.235 (0.502)

Mediana Time Gain			
PH \ finestra	5-10	5-15	5-20
30	10 [5 - 10]	10 [5 - 15]	15 [10 - 20]
45	10 [5 - 10]	10 [5 - 15]	10 [10 - 15]
60	10 [5 - 10]	10 [5 - 15]	10 [10 - 20]

Tabelle 6.9: Prestazioni Decision Trees (PH misto, Bilanciamento OSS, Ottimizzazione bayesiana).

Come atteso, i modelli ottenuti sono affetti da overfitting, che si traduce in una disparità tra gli alti valori di sensibilità e quelli bassi di precisione. L’F1-score risultante è minore rispetto alle metodologie precedentemente esaminate, segno che l’incremento della sensibilità non è in grado di controbilanciare il calo della precisione. Il Time Gain medio (e mediano) è più alto, ma è direttamente correlato alla forte aggressività del modello.

6.6.4 Random Forest

Per l’addestramento delle Random Forest, Matlab® offre la possibilità di usare o la funzione `fitcensemble`, dotata della funzionalità di identificazione degli iper-parametri tramite ottimizzazione bayesiana, o la funzione `Treebagger`, che non esegue in maniera autonoma alcuna ottimizzazione ma accetta molti dei parametri utilizzabili nella funzione `fitctree`. Tra le due opzioni, la funzione `Treebagger` si è rivelata quella in grado di fornire mediamente dei modelli migliori. L’addestramento delle Random Forest è stato eseguito individuando i parametri ottimi del singolo albero tramite l’ottimizzazione di `fitctree`, che sono stati forniti in ingresso alla funzione `Treebagger`, utilizzata per l’identificazione della collezione di classificatori.

PRESTAZIONI RANDOM FOREST

Precisione				Sensibilità			
finestra \ PH	5-10	5-15	5-20	finestra \ PH	5-10	5-15	5-20
	30	0.443 (0.012)	0.371 (0.010)		0.308 (0.008)	30	0.857 (0.025)
45	0.455 (0.013)	0.381 (0.011)	0.330 (0.010)	45	0.818 (0.026)	0.878 (0.013)	0.906 (0.012)
60	0.466 (0.016)	0.400 (0.013)	0.339 (0.009)	60	0.812 (0.022)	0.878 (0.029)	0.917 (0.010)
90	0.476 (0.017)	-	-	90	0.799 (0.024)	-	-
120	0.482 (0.026)	-	-	120	0.816 (0.029)	-	-

F1-Score				Media Time Gain			
PH finestra	5-10	5-15	5-20	PH finestra	5-10	5-15	5-20
30	0.584 (0.013)	0.528 (0.011)	0.464 (0.010)	30	9.240 (0.291)	11.024 (0.300)	12.935 (0.380)
45	0.585 (0.014)	0.532 (0.012)	0.483 (0.012)	45	9.159 (0.305)	10.942 (0.348)	12.564 (0.262)
60	0.592 (0.015)	0.550 (0.011)	0.495 (0.010)	60	9.184 (0.277)	10.691 (0.337)	12.128 (0.339)
90	0.596 (0.011)	-	-	90	9.041 (0.280)	-	-
120	0.605 (0.015)	-	-	120	8.859 (0.298)	-	-

Mediana Time Gain			
PH finestra	5-10	5-15	5-20
30	10 [5 - 10]	10 [5 - 15]	10 [10 - 15]
45	10 [5 - 10]	10 [5 - 15]	10 [10 - 15]
60	10 [5 - 10]	10 [5 - 15]	10 [5 - 15]
90	10 [5 - 10]	-	-
120	10 [5 - 10]	-	-

Tabella 6.10: Prestazioni Random Forest (PH misto, Bilanciamento OSS, Ottimizzazione bayesiana).

L'uso delle Random Forest ha permesso di mitigare l'overfitting dei singoli alberi decisionali. Ne consegue un leggero abbassamento della sensibilità, adeguatamente controbilanciato dall'aumento della precisione, che ha permesso il raggiungimento dei valori di F1-score, almeno per il PH= 5-10 min, analoghi a quelli del modello logistico e delle SVM. A differenza di quest'ultimi, sembra che l'aumento della finestra di estrazione non causi un peggioramento delle prestazioni del modello, ma bensì un miglioramento, seppur leggero. Per questo motivo sono state valutate, per il solo PH=5-10 min (il migliore), i modelli ottenuti tramite l'uso di una finestra di estrazione di 90 e 120 min. Il risultato è che il modello ottenuto con finestra di 120

min e PH= [5-10] minuti è finora il migliore in termini di bilanciamento fra F1-score e Time Gain medio-mediano.

6.6.5 Logistic Model Tree

Attualmente Matlab® non fornisce funzioni per l'identificazione dei Logistic Model Tree. Seguendo quindi la procedura di addestramento semplificata descritta nel paragrafo 5.5 si è individuato i modelli addestrando prima l'albero decisionale ottimizzato tramite comando fitctree e successivamente i vari modelli logistici da inserire alle foglie. Per ogni modello logistico è stata poi svolta una correzione della soglia di classificazione seguendo i seguenti step:

1. Selezionare il primo modello logistico e identificare la soglia di classificazione che garantisce l'F1-score più alto;
2. Fissare la soglia per tale modello;
3. Ripetere i primi due step per ognuno degli altri modelli logistici, procedendo in sequenza;
4. Ripetere i primi tre step un'altra volta.

Per questa metodologia sono state valutate solo alcune delle combinazioni (finestra,PH), scegliendo tra quelle che si sono rivelate migliori con i modelli precedenti.

Finestra	PH	Precisione	Sensibilità	F1-score	Media Time Gain	Mediana Time Gain
30	5-10	0.550 (0.023)	0.763 (0.045)	0.638 (0.016)	8.339 (0.390)	5 [5 - 10]
30	5-20	0.539 (0.028)	0.755 (0.041)	0.628 (0.016)	9.110 (0.516)	5 [5 - 10]
60	5-10	0.542 (0.029)	0.748 (0.036)	0.628 (0.447)	8.514 (0.447)	5 [5 - 10]
60	5-20	0.541 (0.021)	0.763 (0.036)	0.632 (0.011)	9.288 (0.464)	5 [5 - 10]
120	5-20	0.537 (0.022)	0.797 (0.024)	0.641 (0.013)	8.786 (0.429)	5 [5 - 10]

Tabella 6.11: Prestazioni Random Forest (PH misto, Bilanciamento OSS, Ottimizzazione bayesiana).

Rispetto agli altri modelli sembra esserci un miglioramento significativo dell’F1-score a discapito di un calo del Time Gain mediano. Tuttavia, il Time Gain Medio sembra essere in linea con i modelli migliori ottenuti dalle altre metodologie. Il logistic Model Tree si è quindi rivelato il candidato migliore per la predizione delle ipoglicemie post-bariatriche tra le metodologie esaminate.

6.6.6 Sintesi finale

In quest’ultima sezione si riportano i modelli migliori individuati tra le varie metodologie esaminate. Per ogni metodologia sono stati riportati i due modelli migliori selezionando:

- Quello con l’F1-score più alto;
- Quello con l’F1-score più alto e Time Gain mediano almeno di 10 minuti. Se non disponibile è stato riportato il secondo modello con l’F1-score più alto.

Tipo	Finestra	PH	Precisione	Sensibilità	F1-score	Media Time Gain	Mediana Time Gain
Logistico	30	5-10	0.542 (0.015)	0.675 (0.028)	0.601 (0.012)	8.412 (0.327)	5 [5 - 10]
Logistico	45	10-10	0.484 (0.016)	0.686 (0.015)	0.567 (0.012)	9.762 (0.270)	10 [5 - 10]
Decision Tree	30	5-10	0.433 (0.017)	0.858 (0.027)	0.575 (0.014)	9.397 (0.417)	10 [5 - 10]
Decision Tree	45	5-10	0.430 (0.024)	0.860 (0.024)	0.573 (0.016)	9.390 (0.459)	10 [5 - 10]
Random Forest	120	5-10	0.482 (0.026)	0.816 (0.029)	0.605 (0.015)	8.859 (0.298)	10 [5 - 10]
Random Forest	90	5-10	0.476 (0.017)	0.799 (0.024)	0.596 (0.011)	9.041 (0.280)	10 [5 - 10]
SVM	30	5-10	0.536 (0.013)	0.705 (0.034)	0.608 (0.013)	8.312 (0.249)	5 [5 - 10]
SVM	30	5-15	0.506 (0.014)	0.731 (0.034)	0.597 (0.013)	8.965 (0.341)	10 [5 - 10]
Logistic Model Tree	120	5-20	0.537 (0.022)	0.797 (0.024)	0.641 (0.013)	8.786 (0.429)	5 [5 - 10]
Logistic Model Tree	30	5-10	0.550 (0.023)	0.763 (0.045)	0.638 (0.016)	8.339 (0.390)	5 [5 - 10]

Tabella 6.12: Prestazioni dei modelli migliori fra quelli esaminati.

Osservando la tabella qui riportata si nota come tra le varie metodologie esaminate non ci sia una differenza significativa di F1-score differenzi molto, ad eccezione del Logistic Model Tree. I Time Gain mediani sono in generale molto bassi, e questo può essere un problema se si considera che tra l'assunzione di un hypo treatment e la sua effettiva azione passano dei minuti. Prima di concludere l'analisi delle varie metodologie si è infine provato a ottimizzare il peso tra le classi per i modelli selezionati in 6.12, ad eccezione del Logistic Model Tree.

Tipo	Finestra	PH	Precisione	Sensibilità	F1-score	Media Time Gain	Mediana Time Gain
Logistico	30	5-10	0.530 (0.011)	0.713 (0.027)	0.608 (0.013)	8.319 (0.283)	5 [5 - 10]
Logistico	45	10-10	0.483 (0.014)	0.688 (0.017)	0.567 (0.011)	9.684 (0.276)	10 [5 - 10]
Decision Tree	30	5-10	0.535 (0.026)	0.712 (0.038)	0.610 (0.017)	7.830 (0.368)	5 [5 - 10]
Decision Tree	45	5-10	0.548 (0.023)	0.700 (0.031)	0.614 (0.015)	8.468 (0.427)	5 [5 - 10]
Random Forest	120	5-10	0.533 (0.013)	0.756 (0.029)	0.625 (0.013)	8.497 (0.317)	5 [5 - 10]
Random Forest	90	5-10	0.537 (0.020)	0.710 (0.036)	0.611 (0.014)	8.564 (0.370)	5 [5 - 10]
SVM	30	5-10	0.528 (0.013)	0.761 (0.023)	0.624 (0.013)	8.151 (0.197)	5 [5 - 10]
SVM	30	5-15	0.514 (0.015)	0.743 (0.028)	0.608 (0.012)	8.403 (0.290)	5 [5 - 10]
Logistic Model Tree	120	5-20	0.537 (0.022)	0.797 (0.024)	0.641 (0.013)	8.786 (0.429)	5 [5 - 10]
Logistic Model Tree	30	5-10	0.550 (0.023)	0.763 (0.045)	0.638 (0.016)	8.339 (0.390)	5 [5 - 10]

Tabella 6.13: Prestazioni dei modelli migliori fra quelli esaminati con ottimizzazione dei pesi (logistic model tree escluso).

L'ottimizzazione dei pesi ha portato ad un miglioramento dell'F1-score dei modelli a discapito di una riduzione dei parametri di Time Gain. Il logistic Model Tree resta comunque il modello migliore, segno che l'azione esercitata dai modelli logistici sulla mitigazione dell'overfitting degli alberi decisionali si è rivelata la soluzione più vincente tra quelle proposte per la predizione delle ipoglicemie post-bariatriche.

CAPITOLO 7

CONCLUSIONI

Gli algoritmi sviluppati in questa tesi per la predizione delle ipoglicemie post-bariatriche hanno dimostrato di avere prestazioni superiori a quelle dell'algoritmo proposto da Laguna Sanz et al., come evidenziato in tabella 7.1 dove è riportato un riassunto delle metriche della migliore configurazione (logistic model tree) confrontate proprio contro quelle dell'algoritmo stato dell'arte. Tuttavia, i risultati evidenziano come ci sia ancora grande margine di miglioramento per quel che riguarda le prestazioni ottenute, in particolare relativamente alla precisione e al time gain. Tra le varie cause che non hanno permesso il raggiungimento di risultati migliori possiamo evidenziare: la presenza di un'elevata componente rumorosa sui dati CGM; il relativamente basso numero di soggetti acquisiti presenti nel dataset; la scarsa capacità dei modelli adottati di generalizzare il problema. Sicuramente l'utilizzo dei soli dati CGM ha reso lo sviluppo degli algoritmi di predizione delle ipoglicemie post-bariatriche un compito molto più arduo. Si pensi ad esempio all'importanza che l'informazione legata ai pasti potrebbe avere nel miglioramento delle prestazioni raggiunte.

L'abbandono delle tecniche di machine learning per l'investigazione di algoritmi predittivi basati su modello (AR, ARMA, ARIMA, etc.) non sembra essere una strategia adeguata a questo tipo di dati. Le dinamiche sono troppo rapide e variano nel tempo, rendendo il processo fortemente non-stazionario. Pertanto, i modelli costruiti non sarebbero utili per la stima dei valori di concentrazione di glucosio futuri, poiché l'identificazione dei loro parametri si basa sull'ipotesi di stazionarietà del processo di uscita, che in questo caso non è soddisfatta (si rimanda all'appendice dei risultati per verificare quanto detto).

Modello	precisione	Sensibilità	F1-score	Media Time Gain	Mediana Time Gain
Laguna Sanz et al.	0.247	0.812	0.379	14.317	15 [10-20]
Logistic Model Tree	0.537 (0.022)	0.797 (0.024)	0.641 (0.013)	8.786 (0.429)	5 [5 - 10]

Tabella 7.1: Confronto tra il modello migliore e una implementazione dell'algoritmo di Laguna et al. sulla popolazione acquisita, ottenute considerando i tracciati dei soggetti come un'unica serie temporale.

7.1 ULTERIORI SVILUPPI

Il lavoro qui presentato può essere considerato come un'analisi preliminare al problema della predizione delle ipoglicemie nella popolazione post-bariatrica, che ancora non è stato ampiamente affrontato. In prosecuzione del lavoro svolto si potrebbe indagare l'uso di metodologie più complesse, come le varie architetture delle reti neurali, che si sono rivelate particolarmente efficaci nella predizione delle ipoglicemie nella popolazione diabetica [35]. O ancora arricchire l'analisi già svolta utilizzando dataset più ricchi e contenenti informazioni esterne ai dati CGM, come il tempo di assunzione dei pasti.

Un interessante aspetto da approfondire è quello legato all'enorme importanza giocata dal rischio dinamico nella classificazione tramite alberi decisionali (e quindi random forest e logistic model tree), come visibile nel grafico riportato in figura 7.1. Essendo il calcolo di tale feature dipendente dai parametri della funzione di trasformazione e dal fattore di amplificazione associato alla derivata, si potrebbe valutare come un'ottimizzazione di quest'ultimi possa beneficiare sulle prestazioni del modello.

Inoltre, con le opportune considerazioni del caso, sarebbe interessante estendere l'analisi anche alla popolazione diabetica, dove algoritmi simili sono già stati testati con buoni risultati [36].

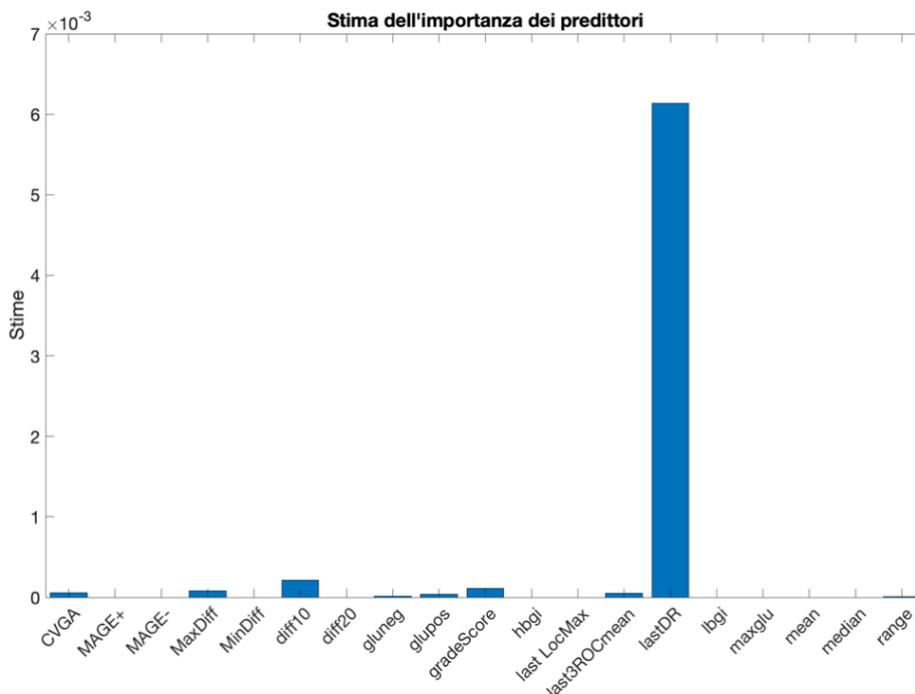


Figura 7.1: Grado di importanza dei predittori negli alberi decisionali.

In conclusione, ci si riserva di dire che il problema della predizione delle ipoglicemie nella popolazione post-bariatrica è ancora poco investigata e merita una particolare attenzione, soprattutto in considerazione del crescente numero di interventi che vengono eseguiti ogni anno e dei benefici che la risoluzione di tale problema potrebbe portare nel progresso degli obiettivi della terapia, nella prevenzione contro l'introduzione di terapie farmacologiche e, più in generale, nel miglioramento delle condizioni di vita della persona che ne è affetta.

APPENDICE A

A.1 Valutazione preliminare di algoritmi di predizione basati su modello

Nella prima sezione dell'appendice si riportano i risultati dell'analisi preliminare delle prestazioni degli algoritmi di predizione basati su modello lineare Black-Box. Prima di scegliere l'adozione di metodologie predittive basate su un approccio alla Machine Learning, infatti, sono state valutate le performance dell'unico algoritmo proposto in letteratura congiuntamente a quelle di alcuni semplici modelli lineari, come il Running Line e l'AR (RLS). Poiché l'applicazione di tali modelli pone come assunto di base la stazionarietà della serie temporale, che non può essere soddisfatta in presenza di andamenti così rapidi come quelli rilevati nella popolazione post-bariatrica, le aspettative sulle loro capacità predittive non erano buone. Tuttavia, si è preferito testare lo stesso alcune metodologie, per essere pienamente sicuri di quanto ipotizzato.

A.1.1 PRESTAZIONI ALGORITMO DI PREDIZIONE DI LAGUNA SAINZ ET AL.

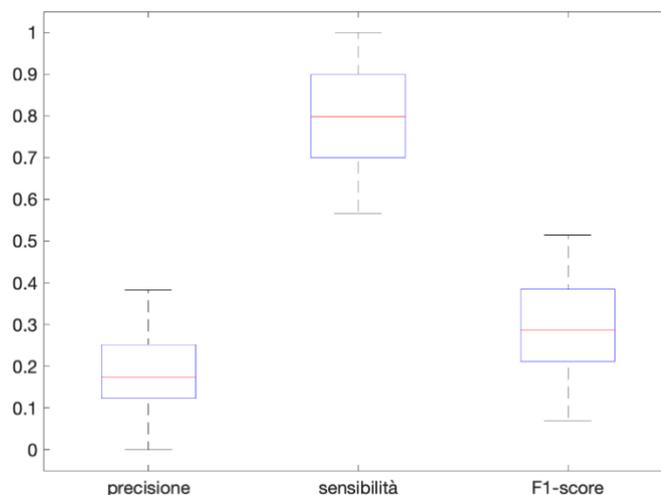


Figura A.1: Boxplot delle performance dell'implementazione dell'algoritmo di Laguna Sainz et al. sui soggetti acquisiti.

A.1.2 PRESTAZIONI MODELLO RUNNING LINE CON FORGETTING FACTOR

Running Line PH=30 min				Running Line PH=20 min			
μ	precisione	sensibilità	F1-score	μ	precisione	sensibilità	F1-score
0.9	0.078	0.081	0.079	0.9	0.100	0.079	0.088
0.8	0.161	0.377	0.226	0.8	0.155	0.272	0.197
0.7	0.218	0.700	0.332	0.7	0.246	0.625	0.353
0.5	0.211	0.852	0.339	0.5	0.256	0.845	0.393

Tabelle A.1: Prestazioni modello running line al variare del PH e del forgetting factor μ .

A.1.3 PRESTAZIONI MODELLO AR (RLS) CON FORGETTING FACTOR

AR (RLS) PH=30 min				AR (RLS) PH=20 min			
μ	precisione	sensibilità	F1-score	μ	precisione	sensibilità	F1-score
0.99	0.366	0.375	0.370	0.99	0.359	0.325	0.342
0.9	0.317	0.804	0.455	0.9	0.377	0.733	0.498
0.8	0.270	0.885	0.414	0.8	0.333	0.849	0.478
0.7	0.245	0.892	0.385	0.7	0.309	0.885	0.458

Tabelle A.2: Prestazioni modello AR (RLS) al variare del PH e del forgetting factor μ .

A.2 Performance delle metodologie per differenti soglie di ipoglicemia

Di seguito si riportano le performance degli algoritmi predittivi per le varie metodologie esaminate al variare della soglia di ipoglicemia. Nello specifico sono stati testati i valori di soglia di 65, 60 e 54 mg/dl. Come anticipato nell'introduzione del capitolo 6, si è riscontrato un calo costante delle performance dei modelli al diminuire del valore di soglia utilizzato. Questo può derivare sia dalla diminuzione del numero di classi positive (ridurre la soglia significa cercare eventi più severi, che sono anche più rari) che dalla perdita di correlazione di alcune metriche con il valore di glicemia futuro.

A.2.1 PRESTAZIONI METODOLOGIE: SOGLIA 65 mg/dl

Modello Logistico: soglia 65 mg/dl

Precisione			
PH finestra	5-10	5-15	5-20
30	0.461 (0.019)	0.453 (0.026)	0.443 (0.027)
45	0.451 (0.022)	0.415 (0.015)	0.394 (0.022)
60	0.473 (0.018)	0.428 (0.018)	0.420 (0.029)

Sensibilità			
PH finestra	5-10	5-15	5-20
30	0.765 (0.024)	0.720 (0.033)	0.632 (0.040)
45	0.764 (0.025)	0.719 (0.020)	0.681 (0.030)
60	0.700 (0.028)	0.669 (0.040)	0.628 (0.043)

F1-Score			
PH finestra	5-10	5-15	5-20
30	0.572 (0.015)	0.555 (0.014)	0.520 (0.019)
45	0.567 (0.015)	0.526 (0.014)	0.498 (0.013)
60	0.564 (0.015)	0.521 (0.012)	0.502 (0.014)

Media Time Gain			
PH finestra	5-10	5-15	5-20
30	8.512 (0.300)	9.415 (0.224)	9.773 (0.476)
45	8.833 (0.340)	10.102 (0.361)	11.069 (0.256)
60	8.602 (0.282)	9.701 (0.364)	10.367 (0.438)

Mediana Time Gain			
PH finestra	5-10	5-15	5-20
30	5 [5 - 10]	10 [5 - 10]	10 [5 - 10]
45	5 [5 - 10]	10 [5 - 15]	10 [5 - 15]
60	5 [5 - 10]	10 [5 - 10]	10 [5 - 15]

Tabelle A.3: Prestazioni modello logistico (soglia di ipoglicemia: 65 mg/dl).

Support Vector Machine: soglia 65 mg/dl

Precisione				Sensibilità			
finestra \ PH	5-10	5-15	5-20	finestra \ PH	5-10	5-15	5-20
	30	0.476 (0.023)	0.468 (0.024)		0.468 (0.046)	30	0.756 (0.029)
45	0.473 (0.017)	0.476 (0.033)	0.473 (0.037)	45	0.744 (0.017)	0.695 (0.048)	0.698 (0.057)
60	0.468 (0.023)	0.497 (0.024)	0.474 (0.018)	60	0.749 (0.040)	0.643 (0.032)	0.703 (0.035)

F1-Score				Media Time Gain			
finestra \ PH	5-10	5-15	5-20	finestra \ PH	5-10	5-15	5-20
	30	0.583 (0.014)	0.571 (0.015)		0.558 (0.019)	30	8.420 (0.419)
45	0.578 (0.014)	0.563 (0.014)	0.561 (0.012)	45	8.454 (0.290)	8.879 (0.599)	8.923 (0.708)
60	0.575 (0.016)	0.560 (0.013)	0.566 (0.013)	60	8.718 (0.405)	8.546 (0.411)	8.988 (0.366)

Mediana Time Gain			
finestra \ PH	5-10	5-15	5-20
	30	5 [5 - 10]	10 [5 - 10]
45	5 [5 - 10]	10 [5 - 10]	10 [5 - 10]
60	5 [5 - 10]	5 [5 - 10]	10 [5 - 10]

Tabelle A.4: Prestazioni SVM (soglia di ipoglicemia: 65 mg/dl).

Decision Tree: soglia 65 mg/dl

Precisione			
PH finestra	5-10	5-15	5-20
30	0.394 (0.023)	0.325 (0.019)	0.276 (0.012)
45	0.386 (0.020)	0.330 (0.015)	0.289 (0.019)
60	0.386 (0.023)	0.335 (0.016)	0.298 (0.020)

Sensibilità			
PH finestra	5-10	5-15	5-20
30	0.865 (0.023)	0.917 (0.013)	0.930 (0.016)
45	0.861 (0.019)	0.903 (0.013)	0.913 (0.014)
60	0.852 (0.018)	0.893 (0.020)	0.901 (0.022)

F1-Score			
PH finestra	5-10	5-15	5-20
30	0.541 (0.020)	0.480 (0.546)	0.426 (0.013)
45	0.533 (0.018)	0.483 (0.016)	0.438 (0.021)
60	0.530 (0.018)	0.487 (0.017)	0.448 (0.022)

Media Time Gain			
PH finestra	5-10	5-15	5-20
30	8.956 (0.344)	10.824 (0.546)	12.419 (0.447)
45	9.226 (0.307)	11.007 (0.488)	12.419 (0.637)
60	9.187 (0.382)	10.725 (0.450)	12.246 (0.620)

Mediana Time Gain			
PH finestra	5-10	5-15	5-20
30	10 [5 - 10]	10 [5 - 15]	10 [5 - 15]
45	10 [5 - 10]	10 [5 - 15]	10 [5 - 15]
60	10 [5 - 10]	10 [5 - 15]	10 [5 - 15]

Tabelle A.5: Prestazioni Decision Tree (soglia di ipoglicemia: 65 mg/dl).

Random Forest: soglia 65 mg/dl

Precisione				Sensibilità			
finestra \ PH	5-10	5-15	5-20	finestra \ PH	5-10	5-15	5-20
	30	0.420 (0.019)	0.360 (0.016)		0.314 (0.012)	30	0.857 (0.023)
45	0.440 (0.016)	0.372 (0.019)	0.327 (0.013)	45	0.824 (0.036)	0.885 (0.032)	0.909 (0.010)
60	0.437 (0.023)	0.370 (0.012)	0.328 (0.012)	60	0.805 (0.038)	0.895 (0.023)	0.906 (0.012)

F1-Score				Media Time Gain			
finestra \ PH	5-10	5-15	5-20	finestra \ PH	5-10	5-15	5-20
	30	0.563 (0.016)	0.514 (0.014)		0.468 (0.014)	30	8.814 (0.218)
45	0.573 (0.015)	0.523 (0.016)	0.481 (0.014)	45	8.928 (0.278)	10.558 (0.278)	11.319 (0.317)
60	0.566 (0.016)	0.523 (0.010)	0.482 (0.014)	60	9.028 (0.335)	10.438 (0.285)	11.295 (0.379)

Mediana Time Gain			
finestra \ PH	5-10	5-15	5-20
	30	10 [5 - 10]	10 [5 - 15]
45	10 [5 - 10]	10 [5 - 15]	10 [5 - 15]
60	10 [5 - 10]	10 [5 - 15]	10 [5 - 15]

Tabelle A.6: Prestazioni Random Forest (soglia di ipoglicemia: 65 mg/dl).

A.2.2 PRESTAZIONI METODOLOGIE: SOGLIA 60 mg/dl

Modello Logistico: soglia 60 mg/dl

Precisione			
PH finestra	5-10	5-15	5-20
30	0.473 (0.035)	0.445 (0.021)	0.429 (0.026)
45	0.456 (0.034)	0.468 (0.021)	0.453 (0.018)
60	0.474 (0.031)	0.436 (0.036)	0.434 (0.030)

Sensibilità			
PH finestra	5-10	5-15	5-20
30	0.613 (0.060)	0.625 (0.052)	0.601 (0.041)
45	0.624 (0.060)	0.586 (0.041)	0.561 (0.036)
60	0.586 (0.058)	0.578 (0.036)	0.560 (0.028)

F1-Score			
PH finestra	5-10	5-15	5-20
30	0.531 (0.022)	0.518 (0.017)	0.499 (0.016)
45	0.524 (0.021)	0.520 (0.020)	0.500 (0.017)
60	0.522 (0.020)	0.495 (0.020)	0.488 (0.020)

Media Time Gain			
PH finestra	5-10	5-15	5-20
30	7.428 (0.441)	8.207 (0.594)	8.767 (0.549)
45	7.794 (0.551)	7.826 (0.511)	8.040 (0.381)
60	7.248 (0.606)	8.016 (0.733)	8.303 (0.548)

Mediana Time Gain			
PH finestra	5-10	5-15	5-20
30	5 [5 - 10]	5 [5 - 10]	5 [5 - 10]
45	5 [5 - 10]	5 [5 - 10]	5 [5 - 10]
60	5 [5 - 10]	5 [5 - 10]	5 [5 - 10]

Tabelle A.7: Prestazioni modello logistico (soglia di ipoglicemia: 60 mg/dl).

Support Vector Machine: soglia 60 mg/dl

Precisione				Sensibilità			
finestra \ PH	5-10	5-15	5-20	finestra \ PH	5-10	5-15	5-20
	30	0.426 (0.024)	0.415 (0.020)		0.412 (0.019)	30	0.654 (0.039)
45	0.409 (0.027)	0.398 (0.016)	0.400 (0.017)	45	0.659 (0.072)	0.671 (0.027)	0.657 (0.027)
60	0.412 (0.030)	0.400 (0.017)	0.403 (0.018)	60	0.642 (0.057)	0.665 (0.032)	0.654 (0.037)

F1-Score				Media Time Gain			
finestra \ PH	5-10	5-15	5-20	finestra \ PH	5-10	5-15	5-20
	30	0.515 (0.022)	0.511 (0.019)		0.508 (0.017)	30	7.658 (0.440)
45	0.502 (0.017)	0.499 (0.015)	0.497 (0.015)	45	8.065 (0.771)	8.016 (0.345)	7.926 (0.412)
60	0.499 (0.018)	0.499 (0.015)	0.498 (0.016)	60	7.983 (0.739)	7.983 (0.373)	7.948 (0.447)

Mediana Time Gain			
finestra \ PH	5-10	5-15	5-20
	30	5 [5 - 10]	5 [5 - 10]
45	5 [5 - 10]	5 [5 - 10]	5 [5 - 10]
60	5 [5 - 10]	5 [5 - 10]	5 [5 - 10]

Tabelle A.8: Prestazioni SVM (soglia di ipoglicemia: 60 mg/dl).

Decision Tree: soglia 60 mg/dl

Precisione			
PH finestra	5-10	5-15	5-20
30	0.351 (0.031)	0.283 (0.019)	0.244 (0.015)
45	0.327 (0.022)	0.280 (0.019)	0.245 (0.016)
60	0.345 (0.030)	0.283 (0.028)	0.247 (0.020)

Sensibilità			
PH finestra	5-10	5-15	5-20
30	0.833 (0.041)	0.890 (0.025)	0.912 (0.019)
45	0.828 (0.033)	0.886 (0.031)	0.869 (0.036)
60	0.801 (0.045)	0.864 (0.041)	0.890 (0.030)

F1-Score			
PH finestra	5-10	5-15	5-20
30	0.492 (0.028)	0.429 (0.023)	0.385 (0.018)
45	0.468 (0.022)	0.425 (0.022)	0.382 (0.019)
60	0.481 (0.025)	0.426 (0.029)	0.386 (0.024)

Media Time Gain			
PH finestra	5-10	5-15	5-20
30	8.652 (0.594)	10.357 (0.553)	11.519 (0.527)
45	8.648 (0.583)	10.177 (0.527)	11.172 (0.658)
60	8.500 (0.725)	9.909 (0.839)	10.984 (0.598)

Mediana Time Gain			
PH finestra	5-10	5-15	5-20
30	5 [5 - 10]	10 [5 - 15]	10 [5 - 15]
45	5 [5 - 10]	10 [5 - 15]	10 [5 - 15]
60	5 [5 - 10]	10 [5 - 15]	10 [5 - 15]

Tabelle A.9: Prestazioni Decision Tree (soglia di ipoglicemia: 60 mg/dl).

Random Forest: soglia 60 mg/dl

Precisione				Sensibilità			
finestra \ PH	5-10	5-15	5-20	finestra \ PH	5-10	5-15	5-20
	30	0.368 (0.020)	0.304 (0.013)		0.267 (0.013)	30	0.814 (0.032)
45	0.369 (0.025)	0.311 (0.016)	0.276 (0.012)	45	0.777 (0.037)	0.851 (0.029)	0.867 (0.019)
60	0.396 (0.033)	0.319 (0.017)	0.282 (0.011)	60	0.729 (0.083)	0.842 (0.022)	0.882 (0.016)

F1-Score				Media Time Gain			
finestra \ PH	5-10	5-15	5-20	finestra \ PH	5-10	5-15	5-20
	30	0.506 (0.021)	0.452 (0.013)		0.506 (0.021)	30	8.205 (0.379)
45	0.500 (0.021)	0.455 (0.016)	0.500 (0.021)	45	8.031 (0.461)	9.651 (0.492)	10.662 (0.467)
60	0.510 (0.022)	0.462 (0.019)	0.510 (0.022)	60	7.758 (0.424)	9.187 (0.488)	10.234 (0.486)

Mediana Time Gain			
finestra \ PH	5-10	5-15	5-20
	30	5 [5 - 10]	10 [5 - 15]
45	5 [5 - 10]	10 [5 - 15]	10 [5 - 15]
60	5 [5 - 10]	5 [5 - 10]	10 [5 - 15]

Tabelle A.10: Prestazioni Random Forest (soglia di ipoglicemia: 60 mg/dl).

A.2.3 PRESTAZIONI METODOLOGIE: SOGLIA 54 mg/dl

Modello Logistico: soglia 54 mg/dl

Precisione			
PH finestra	5-10	5-15	5-20
30	0.416 (0.030)	0.402 (0.017)	0.377 (0.025)
45	0.458 (0.040)	0.443 (0.033)	0.428 (0.015)
60	0.446 (0.036)	0.448 (0.020)	0.447 (0.031)

Sensibilità			
PH finestra	5-10	5-15	5-20
30	0.672 (0.073)	0.666 (0.040)	0.645 (0.045)
45	0.601 (0.016)	0.612 (0.049)	0.599 (0.023)
60	0.660 (0.072)	0.628 (0.024)	0.599 (0.038)

F1-Score			
PH finestra	5-10	5-15	5-20
30	0.511 (0.014)	0.501 (0.013)	0.474 (0.016)
45	0.518 (0.016)	0.512 (0.013)	0.499 (0.009)
60	0.528 (0.015)	0.522 (0.013)	0.510 (0.014)

Media Time Gain			
PH finestra	5-10	5-15	5-20
30	7.514 (0.722)	7.931 (0.622)	7.923 (0.657)
45	6.911 (0.481)	7.248 (0.453)	7.428 (0.414)
60	7.024 (0.349)	7.162 (0.423)	6.979 (0.398)

Mediana Time Gain			
PH finestra	5-10	5-15	5-20
30	5 [5 - 10]	5 [5 - 10]	5 [5 - 10]
45	5 [5 - 5]	5 [5 - 10]	5 [5 - 10]
60	5 [5 - 5]	5 [5 - 10]	5 [5 - 5]

Tabella A.11: Prestazioni modello logistico (soglia di ipoglicemia: 54 mg/dl).

Support Vector Machine: soglia 54 mg/dl

Precisione				Sensibilità			
finestra \ PH	5-10	5-15	5-20	finestra \ PH	5-10	5-15	5-20
	30	0.419 (0.014)	0.419 (0.015)		0.418 (0.010)	30	0.568 (0.027)
45	0.409 (0.015)	0.407 (0.014)	0.381 (0.015)	45	0.564 (0.029)	0.568 (0.028)	0.604 (0.040)
60	0.411 (0.019)	0.411 (0.013)	0.400 (0.016)	60	0.571 (0.044)	0.564 (0.029)	0.580 (0.029)

F1-Score				Media Time Gain			
finestra \ PH	5-10	5-15	5-20	finestra \ PH	5-10	5-15	5-20
	30	0.482 (0.011)	0.481 (0.012)		0.481 (0.010)	30	6.581 (0.494)
45	0.473 (0.013)	0.473 (0.013)	0.466 (0.011)	45	6.482 (0.570)	6.617 (0.507)	7.167 (0.380)
60	0.476 (0.014)	0.475 (0.012)	0.473 (0.012)	60	6.479 (0.595)	6.413 (0.499)	6.814 (0.439)

Mediana Time Gain			
finestra \ PH	5-10	5-15	5-20
	30	5 [5 - 5]	5 [5 - 5]
45	5 [5 - 5]	5 [5 - 5]	5 [5 - 10]
60	5 [5 - 5]	5 [5 - 5]	5 [5 - 10]

Tabelle A.12: Prestazioni SVM (soglia di ipoglicemia: 54 mg/dl).

Decision Tree: soglia 54 mg/dl

Precisione			
PH finestra	5-10	5-15	5-20
30	0.299 (0.022)	0.253 (0.029)	0.230 (0.019)
45	0.309 (0.020)	0.261 (0.028)	0.233 (0.024)
60	0.314 (0.025)	0.267 (0.022)	0.257 (0.023)

Sensibilità			
PH finestra	5-10	5-15	5-20
30	0.796 (0.045)	0.855 (0.051)	0.837 (0.044)
45	0.775 (0.056)	0.836 (0.053)	0.852 (0.039)
60	0.757 (0.071)	0.816 (0.049)	0.813 (0.057)

F1-Score			
PH finestra	5-10	5-15	5-20
30	0.434 (0.025)	0.389 (0.033)	0.352 (0.024)
45	0.441 (0.022)	0.396 (0.030)	0.365 (0.028)
60	0.443 (0.029)	0.402 (0.028)	0.390 (0.026)

Media Time Gain			
PH finestra	5-10	5-15	5-20
30	8.185 (0.685)	9.074 (0.829)	10.145 (0.816)
45	8.006 (0.530)	8.626 (0.601)	9.089 (0.764)
60	8.070 (0.580)	9.158 (0.558)	9.327 (0.611)

Mediana Time Gain			
PH finestra	5-10	5-15	5-20
30	5 [5 - 10]	5 [5 - 10]	10 [5 - 15]
45	5 [5 - 10]	5 [5 - 10]	5 [5 - 10]
60	5 [5 - 10]	10 [5 - 10]	10 [5 - 10]

Tabelle A.13: Prestazioni Decision Tree (soglia di ipoglicemia: 54 mg/dl).

Random Forest: soglia 54 mg/dl

Precisione				Sensibilità			
finestra \ PH	5-10	5-15	5-20	finestra \ PH	5-10	5-15	5-20
	30	0.329 (0.017)	0.289 (0.012)		0.275 (0.008)	30	0.723 (0.059)
45	0.355 (0.021)	0.313 (0.026)	0.290 (0.020)	45	0.664 (0.067)	0.804 (0.077)	0.837 (0.038)
60	0.353 (0.023)	0.332 (0.031)	0.313 (0.027)	60	0.618 (0.070)	0.675 (0.090)	0.742 (0.018)

F1-Score				Media Time Gain			
finestra \ PH	5-10	5-15	5-20	finestra \ PH	5-10	5-15	5-20
	30	0.452 (0.020)	0.427 (0.013)		0.414 (0.012)	30	7.874 (0.412)
45	0.460 (0.017)	0.448 (0.022)	0.430 (0.021)	45	7.616 (0.393)	8.215 (0.465)	8.805 (0.524)
60	0.447 (0.022)	0.441 (0.024)	0.437 (0.018)	60	7.542 (0.502)	8.228 (0.521)	8.636 (0.629)

Mediana Time Gain			
finestra \ PH	5-10	5-15	5-20
	30	5 [5 - 10]	5 [5 - 10]
45	5 [5 - 10]	5 [5 - 10]	5 [5 - 10]
60	5 [5 - 10]	5 [5 - 10]	5 [5 - 10]

Tabelle A.14: Prestazioni Random Forest (soglia di ipoglicemia: 54 mg/dl).

Bibliografia

1. Wolfe BM, Kvach E, Eckel RH. Treatment of obesity: weight loss and bariatric surgery. *Circ res.* 2016 May 27. 118(11):1844-1855.
2. Mechanik JI, Apovian C, Brethauer S, Garvey WT, Joffe AM, Kim J, et al. Clinical practice guidelines for the perioperative nutrition, metabolic, and nonsurgical support of patients undergoing bariatric procedures – 2019 update: cosponsored by American Association of Clinical Endocrinologists/ American College of Endocrinology, The Obesity Society, American Society for Metabolic & Bariatric Surgery, Obesity Medicine Association, and American Society of Anesthesiologists. *Surg Obes Relat Dis.* 2020. 16:175-247.
3. Courcoulas AP, Kiing WC, Belle SH, Berk P, Flum DR, Barcia L. Seven-year weight trajectories and health outcomes in the longitudinal assessment of bariatric surgery (LABS) study. *JAMA Surg.* 2018 May 1. 153(5):427-434.
4. O'Brien PE, Hindle A, Brennan L, Skinner S, Burton P, Smith A. Long-term outcomes after bariatric surgery: a systematic review and meta-analysis of weight loss at 10 or more years for all bariatric procedures and a single-centre review of 20-year outcomes after adjustable gastric banding. *Obes Surg.* 2019 Jan. 29(1):3-14.
5. Mingrone G, Panunzi S, De Gaetano A, Guidone C, Iaconelli A, Capristo E, et al. Metabolic surgery versus conventional medical therapy in patients with type 2 diabetes: 10-years follow-up of an open-label, single-centre, randomized controlled trial. *Lancet.* 2021 Jan 23. 397(10271):293-304.
6. Affinati AH, Esfandiari NH, Oral EA, Kraftson AT. Bariatric surgery in the treatment of type 2 diabetes. *Curr Diab Rep.* 2020 Sep 29. 19(12):156.
7. Kassem MAM, Durda MA, Stoicea N, Cavus O, Sahin L, Rogers B. The impact of bariatric surgery on type 2 diabetes mellitus and the management of hypoglycemic events. *Front Endocrinol.* 2017 Mar 01. 8:37.
8. Kassir R, Debs T, Blanc P, Gugenheim J, Ben Amor I, Boutet C, et al. Complications of bariatric surgery: presentation and emergency management. *Int J Surg.* 2016 Mar. 27:77-81.
9. Brix JM, Kopp HP, Höllerl F, Scherthaner GH, Ludvik B, Scherthaner G, Frequency of hypoglycaemia after different bariatric surgical procedures. *Obes Fact.* 2019. 12(4):397-506.

10. Goldfine AB, Patti ME. How common is hypoglycemia after gastric bypass?. *Obesity*. 2016. 24(6):1210-1211.
11. Danne T, Nimri R, Battelino T, Bergenstal RM, Close KL, DeVries JH, et al. International consensus on use of continuous glucose monitoring. *Diabetes Care*. 2017. 40:1631-1640.
12. Salehi M, Vella A, McLaughlin T, Patti ME. Hypoglycemia after gastric bypass surgery: current concept and controversies. *J Endocrinol Metab*. 2018 Aug 1. 103(8):2815-2826.
13. Craig CM, Liu L, Deacon CF, Holst JJ, McLaughling TL. Critical role for GLP-1 in symptomatic post-bariatric hypoglycemia. *Diabetologia*. 2017 Mar. 60(3):531-540.
14. Suhl E, Anderson-Haynes S, Mulla C, Patti ME. Medical nutrition therapy for post-bariatric hypoglycemia: practical insights. *Surg Obes Relat Dis*. 2017 May. 13(5):888-896.
15. Oviedo S, Vehí J, Calm R, Armengol J. A review of personalized blood glucose prediction strategies for T1DM patients. *Int J Numer Method Biomed Eng*. 2017 Jun. 33(6).
16. Laguna Sanz AJ, Mulla CM, Fowler KM, Cloutier E, Goldfine AB, Newswanger B, et al. Design and clinical evaluation of a novel low-glucose prediction algorithm with mini-dose stable glucagon delivery in post bariatric hypoglycemia. *Diabetes Technol Ther*. 2018 Feb. 20(2):127-139.
17. Harvey RA, Dassau E, Zisser H, Seborg DE, Jovanović L, Doyle F. Design of the health monitoring system for the artificial pancreas: low glucose prediction module.
18. Mulla MC, Zavitsanou S, Laguna Sanz AJ, Pober D, Richardson L, Walcott P, et al. A randomized, placebo-controlled double-blind trial of a closed-loop glucagon system for postbariatric hypoglycemia. *J Clin Endocrinol Metab*. 2020 Apr 1. 105(4):e1260-e1271.
19. Wójcicki JM. "J"-index. A new proposition of the assessment of current glucose control in diabetic patients. *Horm Metab Res*. 1995 Jan. 27(1):41-32.
20. Longato E. Classificazione dello stato di avanzamento del diabete tramite indici di variabilità glicemica. Tesi di laurea, università di Padova. 2016. Padua@thesis.
21. Marling CR, Shubrook JH, Vernier SJ, Wiley MT, Schwartz FL. Characterizing blood glucose variability using new metrics with continuous glucose monitoring data. *J Diabetes Sci Technol*. 2011 Jul 1. 5(4):871-8.
22. Magni L, Raimondo DM, Dalla Man C, Breton M, Patek S, De Nicolao G, et al. Evaluating the efficacy of closed-loop glucose regulation via control-variability grid analysis. *J Diabetes Sci Technol*. 2008 Jul. 2(4):630-635.
23. Hill NR, Hindmarsh PC, Stevens RJ, Stratton IM, Levy JC, Matthews DR. A method for assessing quality of control from glucose profiles. *Diabet Med*. 2007 Jul. 24(7):753-758.

24. Schlichtkrull J, Munck O, Jersild M. The M-value, an index of blood sugar control in diabetics. *Acta Med Scand*. 1965 Jan. 177:95-102.
25. Wójcicki JM. Mathematical descriptions of the glucose control in diabetes therapy. Analysis of the Schlichtkrull “M”-value. *Horm Metab Res*. 1995 Jan. 27(1):1-5.
26. Kovatchev BP, Cox DJ, Gonder-Frederick LA, Clarke W. Symmetrization of the blood glucose measurement scale and its applications. *Diabetes Care*. 1997 Nov. 20(11):1655-1658.
27. Guerra S, Sparacino G, Facchinetti A, Schiavon M, Dalla Man C, Cobelli C. A dynamic risk measure from continuous glucose monitoring data. *Diabetes Technol Ther*. 2011 Aug. 13(8):843-852.
28. Camerlingo N, Vettoretti M, Del Favero S, Cappon G, Sparacino G, Facchinetti A. A real-time continuous glucose monitoring-based algorithm to trigger hypotreatments to prevent/mitigate hypoglycemic events. *Diabetes Technol Ther*. 2019 Nov. 21(11):644-655.
29. Evans JD. *Straightforward statistics for the behavioral sciences*. Thomson Brooks/Cole Publishing Co. 1996.
30. Shalev-Schwartz S, Ben-David S. *Understanding machine learning: from theory to algorithms*. Cambridge University Press. 1st ed. 2014.
31. Breiman L. Random forests. *Machine Learning*. 2001. 45:5-32.
32. Landwehr N, Hall M, Frank E. Logistic Model Trees. *Mach Learn*. 2005. 59:161-205.
33. Kubat M, Matwin S. Addressing the curse of imbalanced training sets: one-sided selection. *Proceedings of the 14th international conference of Machine Learning*. 1997. 179-186.
34. Prendin F, Del Favero S, Vettoretti M, Sparacino G, Facchinetti A. Forecasting of glucose levels and hypoglycemic events: head-to-head comparison of linear and nonlinear data-driven algorithms based on continuous glucose monitoring data only. *Sensors*. 2021 Feb 27. 21(5):1647.
35. Li K, Liu C, Zhu T, Herrero P, Georgiou P. GluNet: a deep learning framework for accurate glucose forecasting. *IEEE J Biomed Health Inform*. 2020 Feb. 24(2):414-423.
36. Dave D, DeSalvo DJ, Haridas B, McKay S, Shenoy A, Joh CJ, et al. Feature-based machine learning model for real-time hypoglycemia prediction. *J Diabetes Sci Technol*. 2021 Jul. 15(4):842-845.

SITI CONSULTATI

- MathWorks help center – www.mathworks.com/help
- PubMed® - www.pubmed.ncbi.nlm.nih.gov.
- ScienceDirect – www.sciencedirect.com .
- SpringerLink – www.springerlink.com.
- Google scholar – scholar.google.com.
- Wikipedia - www.wikipedia.org.
- Machine Learning Mastery – machinelearningmastery.com.
- American Society for Metabolic and Bariatric Surgery – www.asmb.org.
- World Health Organization – www.who.int.
- Clinical Trials Database – www.clinicaltrials.gov.

FONTI IMMAGINI

- Figura 1.1: <https://www.unavitasumisura.it/i-tipi-di-intervento/>.
- Figura 1.2: <https://dottmarcelloavallone.it/roux-en-y-gastric-bypass-laparoscopico/>.
- Figura 1.4: <https://www.dexcom.com/it-IT>.
- Figura 5.2: https://miro.medium.com/max/575/1*p5p7AfXw1Zo3krGZp_Z1MA.png.
- Figura 5.3: https://editor.analyticsvidhya.com/uploads/24817choose_the_hyperplane.png.
- Figura 5.4: https://external-preview.redd.it/wUFvdXRTFanY_2dA-bqhJ7-_mYY3vYoGUHB7RmVIY1A.jpg?auto=webp&s=f89922f1227cbf67cab53abdc5b4668279ce1e23.
- Figura 5.5: <https://www.okpedia.it/data/okpedia/albero-decisionale-esempio.gif>.
- Figura 5.6: <https://media.geeksforgeeks.org/wp-content/uploads/20210707140912/Bagging.png>.
- Figura 6.3: <https://www.semanticscholar.org/paper/COMPARISON-OF-SAMPLING-TECHNIQUES-FOR-IMBALANCED-Durahim/>.