UNIVERSITÀ DI PISA

DIPARTIMENTO DI INGEGNERIA DELL'INFORMAZIONE

Laurea Triennale in Ingegneria Informatica

eXplainable Artificial Intelligence: Sviluppo di una web app per la ricostruzione del battito cardiaco partendo dall'EEG tramite una Rete Neurale Convoluzionale

12 Dicembre 2024

Relatore:

Prof: Antonio Luca Alfeo

Prof: Mario G.C.A. Cimino

Candidato:

Lorenzo Monaci

ANNO ACCADEMICO 2023/2024

Indice

Abstract

In medicina sorge spesso la necessità di sottoporre i pazienti a diversi esami per riuscire a stabilire una diagnosi, i quali possono essere dispendiosi in termini economici, di energia e tempo. Quindi ci chiediamo: *è possibile estrarre più informazioni da un singolo esame?* e qualora lo fosse, *come si può svolgere l'interazione con l'utente finale?*

Il nostro progetto di tesi ha l'obiettivo di rispondere (positivamente) a entrambe le domande. Si propone dunque un metodo per l'estrazione del segnale del battito cardiaco HR (Heart Rate) a partire dall'analisi dell' Elettroencefalogramma (EEG) utilizzando una Rete Neurale Convoluzionale (CNN), e lo sviluppo di un'applicazione che possa rendere fruibile questa tecnica agli utenti finali.

Si è scelto di utilizzare questo tipo di rete neurale perché diversi studi ne hanno dimostrato l'efficacia nell'analisi di dati correlati temporalmente. Per quanto riguarda lo sviluppo del'applicazione si fa uso di Google Colab, un servizio di Notebook Jupyter che offre libero accesso a risorse di calcolo assieme a Streamlit, un framework per lo sviluppo di applicazioni web interattive che fa uso del linguaggio di programmazione Python e infine Ngrok, un servizio di tunneling gratuito che ci permette di creare un link pubblico a cui gli utenti possono collegarsi.

La rete neurale è in grado di ricostruire il battito cardiaco dei singoli soggetti con elevata precisione e di soggetti multipli con una modesta accuratezza partendo da EEG filtrati dall'interferenza di segnali fisiologici maggiori (come il movimento degli arti) ma comunque affetti da rumore generato da piccoli movimenti (come il movimento involontario degli occhi o la respirazione). L'applicazione si interfaccia con l'utente in modo da ricevere come input dei parametri per l'analisi e file in formato csv (comma separated values) gli EEG di un numero variabile di pazienti, così da addestrare la rete in modo da generare un modello che sia il più possibile generico. Si propone inoltre all'utente i risultati dell'elaborazione della CNN e le prestazioni di quest'ultima, in forma di grafici

Capitolo 1

Introduzione

Il nostro progetto punta a sviluppare un'applicazione web per ottenere il battito cardiaco dei pazienti di cui analizziamo l'elettroencefalogramma utilizzando una **Rete Neurale Convoluzionale**.

1.1 Reti Neurali Convoluzionali, un approccio vantaggioso

Questa tecnica è molto efficace per via delle elevate prestazioni di questo tipo di reti per l'analisi di segnali temporali. Un software del genere permetterebbe di accorciare notevolmente le tempistiche e i costi delle visite mediche nonché fornire informazioni che normalmente vengono estratte da esami separati, si pensi quindi all'estensione di questa tecnica per la ricostruzione di altri segnali fisiologici, come la pressione sanguigna o la fotopletismografia [1].

1.2 Impieghi nella realtà, l'XAI a supporto della scienza

É ovvio pensare che un oggetto come questo susciterebbe grande interesse all'interno di vari campi della medicina *(ad es. cardiologia, neurologia, . . .)*, ma si può pensare anche a soggetti interessati allo studio di *Machine Learning* e in particolare all'implementazione di quest'ultimo in ambito medico.

Prendiamo come esempio per il primo caso i medici di base, il loro lavoro sarebbe sicuramente semplificato da questo strumento in quanto potrebbero concentrarsi su un singolo esame a cui sottoporre vari pazienti e ricevere come risposta a questi sia HR che EEG.

Invece immaginiamo degli studenti di machine learning, potrebbero prendere spunto per studi futuri o basare ricerche sul nostro caso di studio.

[1] Un fotopletismogramma (FPG) è un pletismogramma ottenuto per via ottica che può essere usato per rilevare variazioni del volume sanguigno all'interno di tessuti microvascolari. [Fonte:Wikipedia]

Capitolo 2

Related Works

L'utilizzo delle CNN è fondamentale per una predizione accurata nel caso di serie temporali, dato che queste sono 'comprese' in modo molto preciso grazie appunto alla convoluzione che viene applicata, che nel caso di segnali correlati temporalmente permette di fornire una descrizione molto precisa di questi ultimi.

Dato che l'EEG è appunto un segnale temporale, si sceglie di analizzarlo utilizzando queste reti per, nel nostro caso estrarre il battito cardiaco, ma come già detto in precedenza potremmo concentrarci su qualsiasi altro segnale fisiologico.

2.1 L'analisi delle serie temporali

Come dimostrato in uno studio pubblicato sul Journal of Big Data: 'Time-series analysis with smoothed Convolutional Neural Network' *(Aji Prasetya Wibawa, Agung Bella Putra Utama)*[4] l'impiego delle CNN non è sempre pensato per analizzare serie temporali, infatti nasce come approccio all'analisi delle immagini.

Infatti le CNN necessitano di dati filtrati da rumori troppo evidenti per avere delle predizioni consistenti, quindi nasce la necessità di rendere i dati *smooth*, cioè adatti a essere dati in input alla rete.

2.2 Applicazione a segnali fisiologici

I segnali elettrici che il cervello emette e che vengono quindi campionati dai vari sensori dell'apparecchio sono spesso misti a vari tipi di rumore, da quello più grande in ampiezza e facile da riconoscere e quindi rimuovere *(movimenti articolari, rumore termico, ...)* a quello lieve di movimenti impercettibili e involontari del nostro corpo *(Movimento delle palpebre, respiro, organi interni, ...)*.

La maggior parte di queste interferenze possono essere facilmente eliminabili tramite tecniche di *filtraggio*, ma come analizzato nello studio pubblicato sul Journal of Computer Science and Information Technology: 'Convolutional Neural Network Application in Biomedical Signals' *(Haya, A.)*[1] ci sono alcuni artefatti che non possono essere rimossi (rumore termico) o individuati facilmente (micro-movimenti), quindi è necessario

far si che la rete neurale 'si abitui' alla presenza di questi disturbi e li etichetti come tali.

2.3 Il nostro approccio

Cerchiamo di combinare le due cose e rendere in primis i dati privi di ampi artefatti di rumore grazie a operazioni di *filtraggio in frequenza*, successivamente si applica una divisione tra dati di testing e training, utilizzando *un solo dataset* per la validazione e il resto come dataset di addestramento. Ciò permette di dotare la rete di una visione più 'ampia' del problema e generare così un modello che riesca a adattarsi a qualunque soggetto. In seguito si applica una *normalizzazione* ai dati, facendo si che abbiano *media* nulla e *varianza* unitaria in modo da ridurre la variabilità dei dati.

2.3.1 Rappresentazione dell'output

'Keep it simple, keep it stupid'[2]

In base a questo principio alla base di molte applicazioni informatiche e ingegneristiche scegliamo una tecnica di rappresentazione dell'output basata sui colori, in modo che possa essere colta al volo la differenza tra una predizione giusta e una sbagliata e dove si trova il picco cardiaco se presente.

Capitolo 3

Design e Implementazione

3.1 Interfaccia utente

Per interagire con l'utente si fa uso dell'API di **Streamlit**, la quale offre gli strumenti per ricevere in input i parametri e i dati necessari.

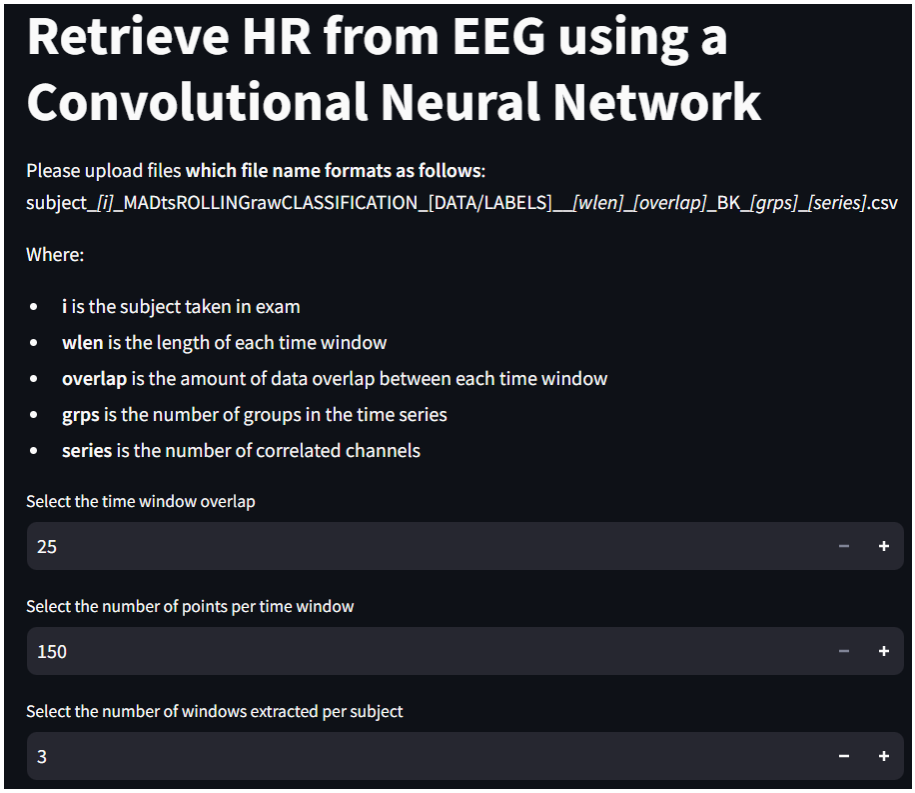

Figura 3.1: User interface

3.1.1 Formato dei file di input e parametri

Si informa innanzitutto l'utente del formato che devono avere i file e di come vanno rinominati in modo da far funzionare l'applicazione, dopodiché vengono richiesti la *finestra di overlap* e il *numero di campioni per finestra temporale*, i quali partono rispettivamente da un minimo di 25

e 150 per arrivare a un massimo di 150 e 250, incrementando di 5 e 50 punti.

3.1.2 Modalità di visualizzazione

Dato che dopo l'analisi di ogni soggetto verrà esposto all'utente un insieme di finestre temporali che rappresentano la performance della CNN, viene richiesto quante finestre temporali si vogliono visualizzare, fino a un massimo di 10 per soggetto.

3.1.3 File Upload

Seguono due *caricatori di file*[1], completi di descrizione ed etichetta di aiuto per l'upload dei **dataset** rappresentanti l'EEG dei soggetti e le **label** degli stessi, dato che questo tipo di rete viene addestrata con un approccio *supervisionato*.

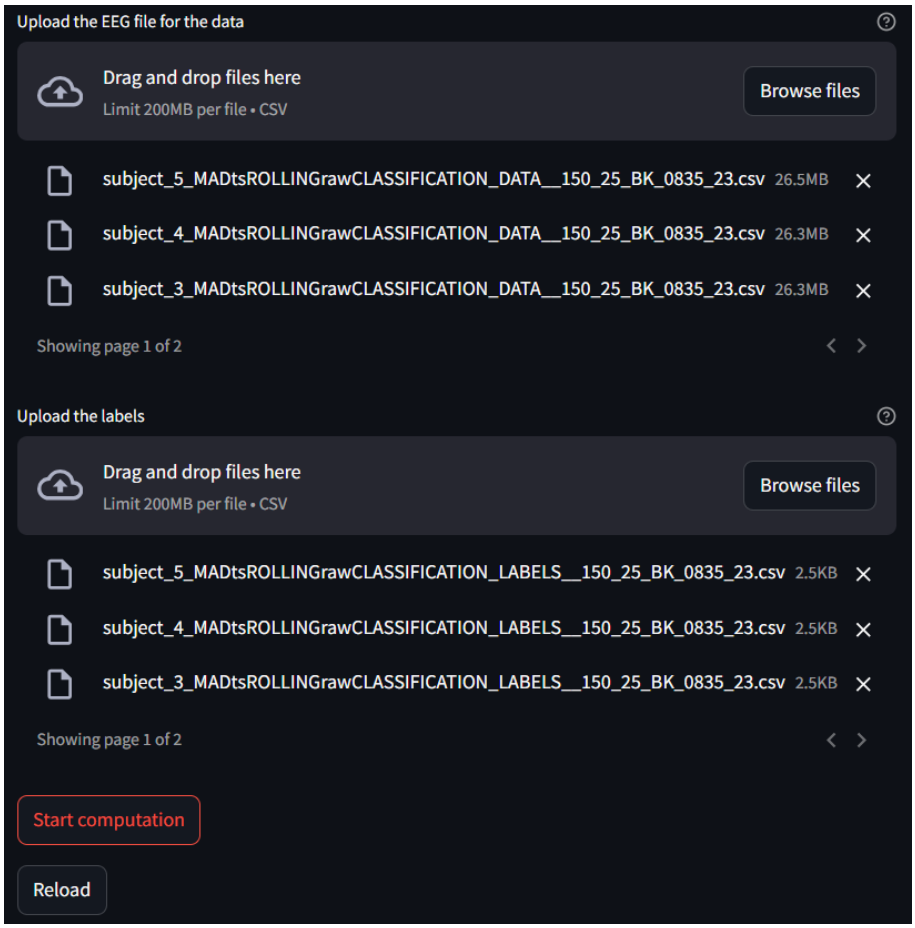

Figura 3.2: User interface

Infine un apposito pulsante *'Start computation'* fa partire l'applicazione, premuto il quale l'utente è informato che l'elaborazione è in esecuzione tramite un apposito *widget animato*

[1] Max 200MB per file

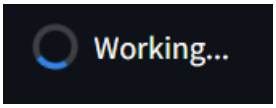

Figura 3.3: Widget di caricamento

3.2 Costruzione del modello di rete neurale

Si tratta di un modello sequenziale composto da 5 *layer*:

Conv1D: un layer convoluzionale caratterizzato da 64 neuroni e una finestra di 5 campioni per volta impegnati nella convoluzione

MaxPooling1D: un layer che riduce il numero di campioni prendendo quello con valore massimo in un pool di valore specificato, nel nostro caso 2

Flatten: layer che si occupa di appiattire la dimensione del tensore in esame, rendendolo monodimensionale

Dense: due layer con filtri densamente connessi[2] con funzioni di attivazione *rettificatrice* e *sigmoidea* riepsttivamente con un numero di filtri pari a 128 e 1.

Il modello è poi compilato con un `batch size` pari a 64 e impostato per andare avanti fino a 100 epoche.

Questo design è efficiente per il nostro caso di studio in quanto rimane un modello relativamente semplice e con un numero limitato di epoche per evitare l'overfitting (stiamo analizzando dataset relativamente piccoli, neanche vicini ai **Big Data**[3]), e un batch di dimensioni intermedie per analizzare una modesta quantità di dati per volta[4].

3.3 Data Preprocessing

I dati inseriti dall'utente in formato `csv` vengono trattati in due modi differenti a seconda del numero di soggetti che stiamo analizzando, per garantire l'analisi di un particolare soggetto oppure per cercare di generare un modello generico.

3.3.1 Soggetto singolo

In questo caso si adotta uno split del dataset stesso seguendo il Principio di Pareto[3]. Così facendo addestriamo la rete neurale su un soggetto particolare, quindi il modello sarà creato 'ad-hoc' per quello specifico soggetto.

[2] Ogni filtro è connesso a tutti qulli del layer precedente e successivo

[3] I quali si considerano tali per grandezze del centinaio di TB

[4] Si ricorda la regola di Pareto[3], per la quale l'80% dei risultati proviene dal 20% dei dati

3.3.2 Soggetti multipli

Analizzando più soggetti alla volta consideriamo il fatto di addestrare la rete su tutti i soggetti meno uno per ogni soggetto, utilizzando quest'ultimo come soggetto di test.

3.3.3 Normalizzazione

Successivamente allo split viene applicata una normalizzazione ai dataset tramite un ridimensionamento standard:

$$z = \frac{x_i - \mu}{\sigma}$$

3.4 CNN Performance & Metrics

3.4.1 Introduzione alle metriche utilizzate

La compilazione del modello avviene tenendo conto di 4 metriche:

1. **Accuracy**: metrica standard, la quale misura l'accuratezza del modello

2. **F1 score**[5]: metrica implementata all'interno del codice che misura anch'essa l'accurateza del modello ma è particolarmente rilevante in casi come il nostro in cui abbiamo dataset *sbilanciati* ed è espressa con la formula
$$2 \cdot \frac{Precision \cdot Recall}{Precision + Recall}$$

3. **Precision**[6]: misura la proporzione di campioni classificati come positivi che sono effettivamente positivi, rispondendo alla domanda *'Quanto possiamo fidarci della classificazione positiva del modello'?*
$$\frac{TP}{TP + TN}$$

4. **Recall**[7]: misura la proporzione di campioni effettivamente positivi che sono stati classificati come tali, rispondendo alla domanda *'Quanto è abile il modello a catturare campioni effettivamente positivi'?*
$$\frac{TP}{TP + FN}$$

3.4.2 Visualizzazione delle Performance

Durante l'addestramento della nostra rete neurale è possibile visualizzare il suo comportamento attraverso le epoche (che ricordiamo essere 100) e durante la validazione.

Per fare ciò vengono mostrati a schermo 3 grafici, i quali riportano l'andamento delle metriche descritte all'inizio di questa sezione durante

[5] F1 Score on Wikipedia
[6] Precision on Wikipedia
[7] Recall on Wikipedia

l'addestramento e la validazione, con l'aggiunta di una quinta metrica (standard): **loss**. Essa indica quanto bene (o male) il modello sta predicendo, basandosi appunto sulle label fornite dall'utente.

E' oltretutto la funzione che il modello cerca di minimizzare per mezzo di metodi basati per esempio sulla *discesa del gradiente*, noi utilizziamo la funzione di ottimizzazione **Adam**[8]

Nel nostro caso utilizziamo la funzione di loss **Binary Cross-Entropy**, la quale è specifica per problemi di classificazione binaria e definita per mezzo dell'espressione:

$$Loss = -\frac{1}{N}\sum_{i=1}^{N} y_i \log(\hat{y}_i) + (1-y)\log(1-\hat{y}_i)$$

Dove y_i è una label che vale 0 o 1 e $\hat{y}_i$ è la probabilità predetta dal modello.

I nostri grafici sono così composti: i primi due indicano rispettivamente l'andamento di accuracy, f1, precision e recall durante l'addestramento, il secondo invece durante la validazione. Il terzo grafico riporta invece l'andamento del valore della loss sia durante l'addestramento che durante la validazione, in modo da avere un confronto diretto tra i due.

Si è scelto di rappresentare la loss in un grafico separato per la grande differenza tra i valori di quest'ultima e le altre metriche, che nei casi più estremi differiscono infatti di *almeno un paio di ordini di grandezza*.

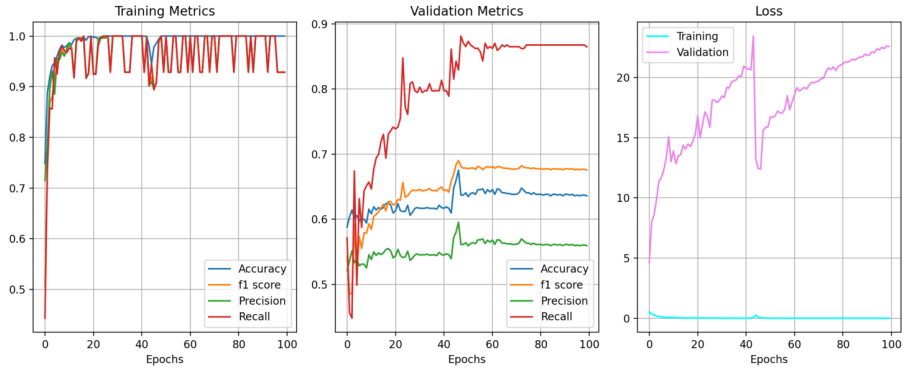

Figura 3.4: Rappresentazione delle metriche

3.4.3 Plotting

La visualizzazione a schermo dei grafici discussi precedentemente avviene per mezzo del pacchetto `matplotlib.pyplot` (per creare le figure) assieme al metodo `pyplot()` del pacchetto `streamlit`

Viene scelta una colorazione apposita per rappresentare a schermo le linee dei vari grafici, la quale è riproposta successivamente in fase di stampa dei risultati.

[8] Adaptive Moment Estimation (Keras)

3.5 Output e Download dei risultati

Durante l'operazione di testing vengono memorizzate le metriche di validazione, che vengono poi mostrate numericamente all'utente utilizzando gli stessi colori dei grafici descritti nella sottosezione 3.4.3.

Nel caso in cui vi siano più soggetti da analizzare l'applicazione memorizza la migliore e peggiore accuracy tra quelle misurate nel corso dell'elaborazione, ricordandosi anche a quale soggetto appartengono.

Questo può aiutarci a capire quali soggetti evidenzino un'atipicità rispetto agli altri, riconoscendo possibili anomalie nel tracciato dell'EEG o casi particolari che richiedono particolare cura nel trattamento.

Oltre a questo viene memorizzato anche il tempo di training per ogni soggetto, per poi calcolare e mostrare all'utente il tempo medio di addestramento alla fine di tutto il processo.

3.5.1 Serie temporale originale e Finestre temporali

Per avere un'idea di come si è comportata la rete neurale e soprattutto per cercare una *spiegazione* al fenomeno analizzato si plotta l'intera serie temporale del soggetto in questione (Figura 3.5) e in seguito un numero scelto dall'utente (come spiegato nella sottosezione 3.1.2) di finestre temporali da visualizzare.

Queste ultime sono delimitate da una riga verticale tratteggiata per rendere evidente la separazione tra esse, viene inoltre delimitato l'overlap con una riga verticale più sottile, a indicare quanti punti temporali fanno parte di una finestra e della precedente/successiva.

Le finestre temporali che vengono mostrate vengono scelte in modo randomico, estraendo un numero estratto con una distribuzione uniforme discreta di probabilità[9] a cui poi viene sommato il numero di finestre scelto dall'utente secondo la formula

$$start = n\_points \cdot r$$

$$stop = n\_points \cdot (r + n\_finestre\_utente)$$

Dove r è il numero estratto, $n\_finestre\_utente$ è il numero di finestre temporali che l'utente desidera visualizzare e n_points il numero di punti temporali per finestra.

La linea continua rappresentata nel grafico può assumere il colore **rosso** nel caso in cui la rete abbia generato una previsione errata, nel caso contrario sarà colorata in **verde**.

Se nella finestra temporale è inoltre presente una label positiva (che sta a indicare che è presente un picco cardiaco) viene indicata con un quadratino **blu** in corrispondenza del valore del picco più alto presente all'interno della finestra.

I grafici sopra descritti sono accompagnati da una legenda che spiega all'utente i colori associati alle varie linee nel grafico.

[9] Fonte: numpy.org

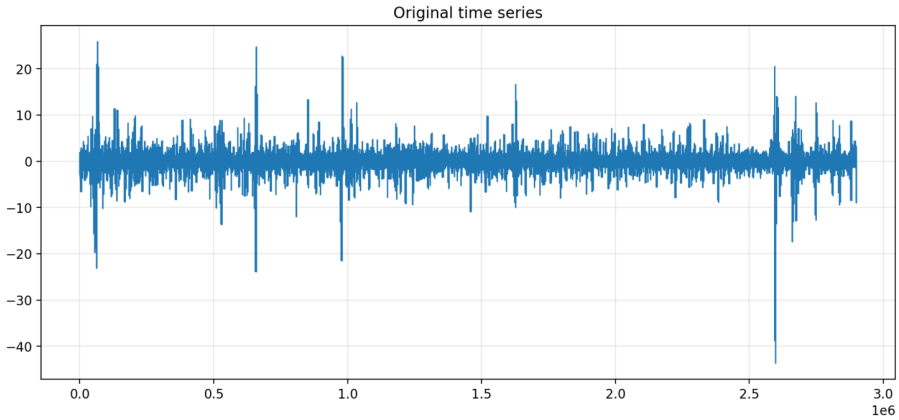

Figura 3.5: Rappresentazione delle finestre temporali

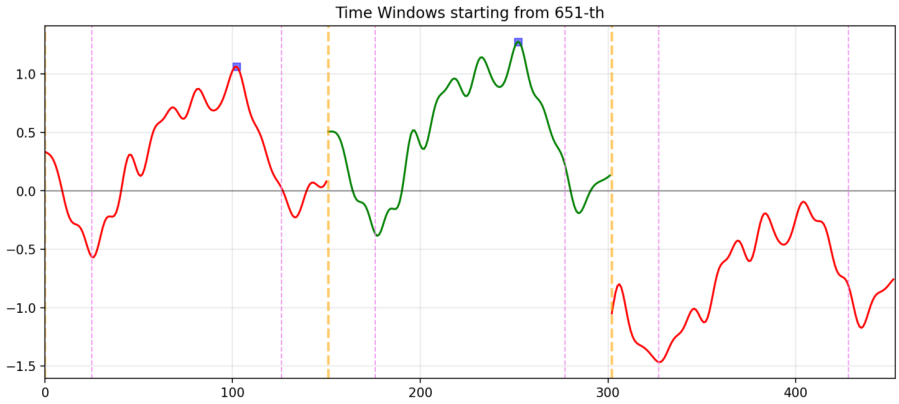

Figura 3.6: Rappresentazione delle finestre temporali

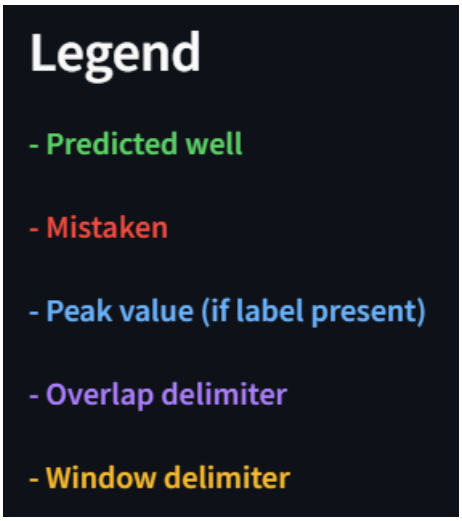

Figura 3.7: Legenda

3.5.2 Performance over all the subjects

Se abbiamo analizzato dati relativi a soggetti multipli, vengono plottate le varie metriche discusse nella sottosezione 3.4.1 (inclusa la *loss*) in due grafici separati, in funzione di ogni soggetto analizzato (Figura 3.8).

3.5.3 Download di risultati e immagini in formato .zip

Nel caso l'utente voglia salvare nel proprio filesystem i risultati dell'analisi appena conclusa può farlo tramite due appositi pulsanti (Figura 3.8).

File di testo

Un pulsante è dedicato al salvataggio dei risultati numerici visualizzati a schermo, viene creato un archivio `.zip` contenente un file per ogni soggetto contenente le performance della CNN per quel soggetto.

File immagini

Un altro pulsante invece si occupa del salvataggio delle immagini contenenti i grafici discussi nella sottosezione 3.5.1, nello stesso modo descritto precedentemente.

Figura 3.8: Risultati, metriche per ogni soggetto e pulsanti di download

3.6 Google Colab

3.7 Tunneling tramite Ngrok

Appendice A

Questa è un'appendice

Fusce mauris. Vestibulum luctus nibh at lectus. Sed bibendum, nulla a faucibus semper, leo velit ultricies tellus, ac venenatis arcu wisi vel nisl. Vestibulum diam. Aliquam pellentesque, augue quis sagittis posuere, turpis lacus congue quam, in hendrerit risus eros eget felis. Maecenas eget erat in sapien mattis porttitor. Vestibulum porttitor. Nulla facilisi. Sed a turpis eu lacus commodo facilisis. Morbi fringilla, wisi in dignissim interdum, justo lectus sagittis dui, et vehicula libero dui cursus dui. Mauris tempor ligula sed lacus. Duis cursus enim ut augue. Cras ac magna. Cras nulla. Nulla egestas. Curabitur a leo. Quisque egestas wisi eget nunc. Nam feugiat lacus vel est. Curabitur consectetuer.

Appendice B

Questa è un'altra appendice

Suspendisse vel felis. Ut lorem lorem, interdum eu, tincidunt sit amet, laoreet vitae, arcu. Aenean faucibus pede eu ante. Praesent enim elit, rutrum at, molestie non, nonummy vel, nisl. Ut lectus eros, malesuada sit amet, fermentum eu, sodales cursus, magna. Donec eu purus. Quisque vehicula, urna sed ultricies auctor, pede lorem egestas dui, et convallis elit erat sed nulla. Donec luctus. Curabitur et nunc. Aliquam dolor odio, commodo pretium, ultricies non, pharetra in, velit. Integer arcu est, nonummy in, fermentum faucibus, egestas vel, odio.

Bibliografia

[1] A. Haya. Convolutional Neural Network Application in Biomedical Signals. *ResearchGate*, 1:16, Gennaio 2018.

[2] U.S. Navy, 1960.

[3] Vilfredo Pareto, 1848-1923.

[4] Aji Prasetya Wibawa and Agung Bella Putra Utama. Time-series analysis with smoothed Convolutional Neural Network. *Journal of Big Data*, 1:18, Aprile 2022.