
Predizione sull'incidenza di Patologie Cardiache utilizzando una Rete Neurale

Andrea Saggio • 757HHHINGINFOR

Predizione sull'incidenza di Patologie Cardiache utilizzando una Rete Neurale

Argomento:

- Machine Learning Applicato a Predizioni di Patologie Cardiache
- Intelligenza Artificiale
- Rete Neurale

Scopo:

- Rilevare con Precisione la presenza di Patologie Cardiache in un Paziente utilizzando Tecniche di Apprendimento Automatico.



Dataset

heart_research.csv

- Il Dataset è un file csv (comma-separated values) composto da 14 colonne e 303 righe e contiene una serie di informazioni mediche sui pazienti

Informazioni

AGE	SEX
CP	TRESTBPS
CHOL	FBS
RESTECG	THALACH
EXANG	OLDPEAK
SLOPE	CA
THAL	TARGET



Strumenti Utilizzati

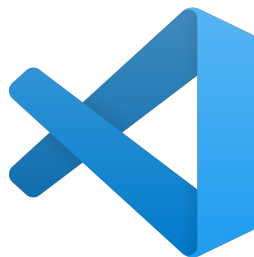
Linguaggio

- Linguaggio Python versione 3.10.6.



IDE (Integrated Development Environment)

- Visual Studio code.



Librerie

Pandas	Matplotlib
Seaborn	Scikit-Learn
Keras	FBS

I vantaggi di Python

Rapidità

- Python è un linguaggio molto veloce grazie alla sua natura interpretata e all'uso di librerie altamente ottimizzate per la computazione numerica e scientifica.

Efficienza

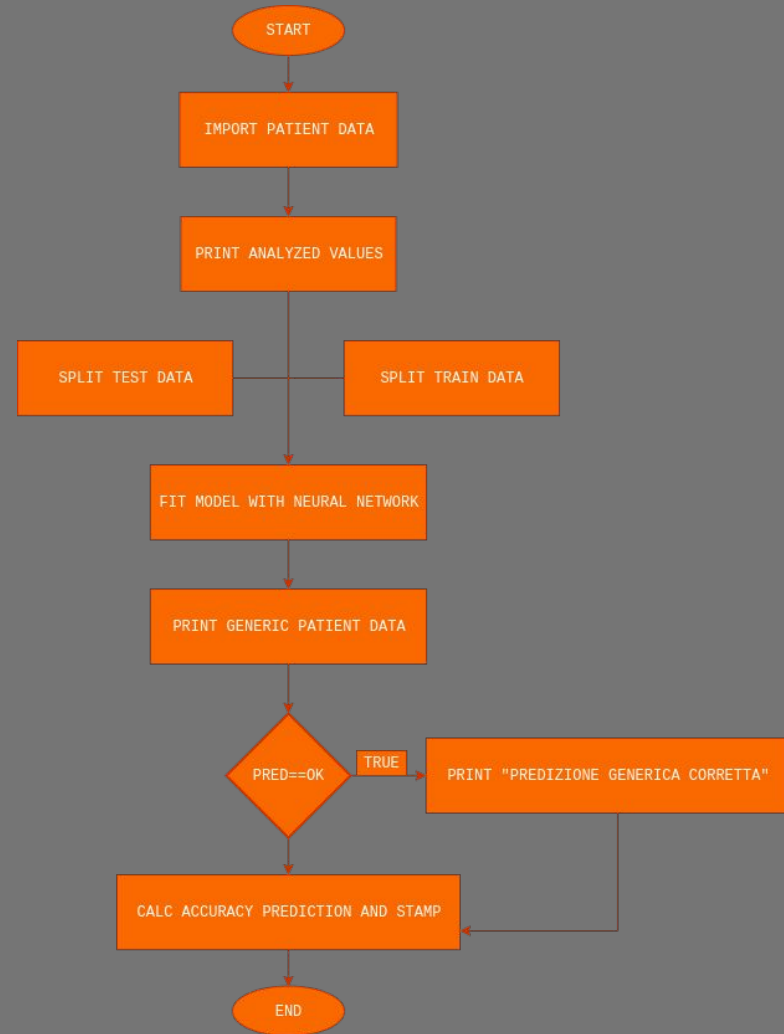
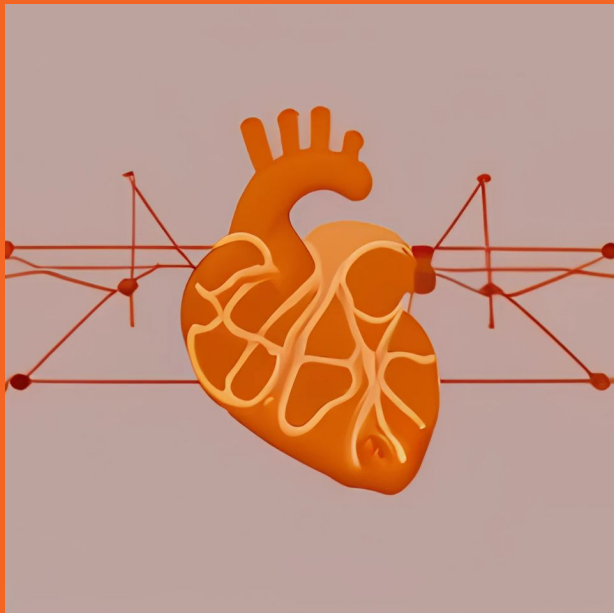
- Python è in grado di gestire grandi quantità di dati in modo efficiente grazie all'uso di strutture dati native come le liste, i dizionari e le tuple, che consentono di elaborare e manipolare grandi quantità di dati in modo rapido ed efficiente.

Espressività

- Python è un linguaggio molto espressivo grazie alla sua sintassi leggibile e comprensibile, che consente di scrivere codice in modo chiaro e conciso.



ALGORITMO



Importazione delle librerie necessarie per il codice, tra cui Pandas, Matplotlib, Seaborn, Sklearn e Keras.

Calcolo della matrice di correlazione tra le feature del dataset e visualizzazione dei risultati tramite una heatmap.

Addestramento del modello e calcolo dell'accuracy score per la valutazione delle prestazioni del modello.



Caricamento del dataset in formato CSV attraverso la funzione "pd.read_csv".

Divisione dei dati in training set e test set.

OUTPUT

Epoche

- Vi è riportato il risultato di un addestramento di una rete neurale utilizzando il linguaggio Python e la libreria Keras.
- L'addestramento è stato effettuato per un totale di 2000 epoche, e il risultato di ogni epoca viene riportato in termini di perdita (loss) e accuratezza (accuracy).

```
Epoch 1964/2000
8/8 [=====] - 0s 1ms/step - loss: 0.3474 - accuracy: 0.8388
Epoch 1965/2000
8/8 [=====] - 0s 2ms/step - loss: 0.3347 - accuracy: 0.8554
Epoch 1966/2000
8/8 [=====] - 0s 1ms/step - loss: 0.3300 - accuracy: 0.8512
Epoch 1967/2000
8/8 [=====] - 0s 1ms/step - loss: 0.3284 - accuracy: 0.8554
Epoch 1968/2000
8/8 [=====] - 0s 1ms/step - loss: 0.3330 - accuracy: 0.8595
Epoch 1969/2000
8/8 [=====] - 0s 1ms/step - loss: 0.3415 - accuracy: 0.8471
Epoch 1970/2000
8/8 [=====] - 0s 1ms/step - loss: 0.3390 - accuracy: 0.8512
```



RISULTATI

Predizione Paziente Generico

```
--- IL PAZIENTE CON QUESTI VALORI :  
age          64.0  
sex          1.0  
cp          2.0  
trestbps    125.0  
chol        309.0  
fbs         0.0  
restecg     1.0  
thalach     131.0  
exang       1.0  
oldpeak     1.8  
slope       1.0  
ca          0.0  
thal        3.0  
Name: 225, dtype: float64  
--- HA DATO QUESTO RISULTATO (0 Assente - 1 Presente) : [0.21311358]  
--- IL VALORE REALE DEL PAZIENTE E' (0 Assente - 1 Presente) : 0  
--- LA PREDIZIONE E' STATA : CORRETTA
```

Percentuale di Precisione Ottenuta:

```
--- La PRECISIONE DELLA PREDIZIONE GENERALE E' STATA DEL : 80.33 %  
o asaggio@asaggio-Vostro-3500:~/progetto_uni$
```



Conclusioni

- Il modello creato ha dimostrato di essere abbastanza preciso nella predizione della presenza o meno di malattie cardiache, con un'accuratezza del 80.33%. Ciò suggerisce che i dati utilizzati per addestrare il modello erano sufficienti per identificare i pattern che indicano la presenza di malattie cardiache.



Sviluppo Futuro

1. Usare Architetture di Rete Neurale Convolute.
 2. Utilizzo di Tecniche di Regolarizzazione e di Ottimizzazione dei parametri del modello.
 3. l'esplorazione di nuove fonti di dati per migliorare la capacità di generalizzazione del modello
-

