

- **Model Details**

Il classificatore Music x Mental Health è in grado di classificare i generi musicali preferiti dei partecipanti e predire in base a determinate feature lo stato depressivo di ogni utente. Inoltre può essere addestrato anche per predire altri disagi come ad esempio ansia, insonnia e disturbo ossessivo compulsivo (ocd).

- **Intended Use**

Il modello è stato sviluppato per fini puramente didattici e per divertimento.

- **Factors**

I fattori evidenziati in questo modello mettono in relazione i vari generi musicali, con disturbi come ansia, depressione, insonnia e ocd. Altri fattori rilevanti sono sicuramente l'età dei partecipanti. Inoltre bisogna tener conto che le valutazioni riguardo i disturbi mentali sopra citati non sono state effettuate da persone competenti in quel determinato ambito, bensì i valori riportati corrispondono ad una autodescrizione del soggetto, che spesso potrebbe essere influenzata da stati d'animo passeggeri.

- **Evaluation Data**

Il dataset si presenta fortemente squilibrato. La maggioranza delle persone che hanno partecipato al sondaggio hanno un'età compresa tra i 15 ed i 25 anni. Non a caso tra le operazioni di Pre Processing è stata adoperata una banale tecnica di oversampling per provare a bilanciare i dati senza perdere informazioni o alterare esageratamente i dati. Inoltre la varianza dei dati è molto alta rapportata al numero di record totali iniziali (circa 650 record). Anche per questo motivo, ad esempio, l'analisi condotta sulla metrica per riconoscere la depressione è stata effettuata a valle di una modifica di questi dati. In particolare nel dataset originario, la colonna Depression era di tipo float e poteva contenere qualsiasi valore compreso tra 0 e 10. Per effettuare l'addestramento, ho convertito questa colonna in una di tipo intero, che

contenesse solo i valori 0 ed 1 (0 → non depresso, 1 → depresso).

Inoltre sono presente molte feature quasi inconsistenti, ad esempio i BPM. Potrebbe influenzare il risultato in quanto si potrebbe pensare che generi con BPM più lenti siano ascoltati da persone più predisposte alla depressione, quando in realtà non è così. I valori di BPM infatti si aggirano quasi tutti tra i 100 e i 150 (la maggior parte dei generi musicali ricade in questo intervallo !).

- **Metrics**

Sono state utilizzate metriche di confusione, come ad esempio la matrice di confusione, nel caso di modello di addestramento binario. Inoltre sono state confrontate le misure di accuracy tra due modelli: deep learning e decision tree.

Il modello decision tree si presentava più performante rispetto a quello di deep learning, per questo le metriche di accuracy sono state calcolate considerando i risultati del classificatore decision tree.

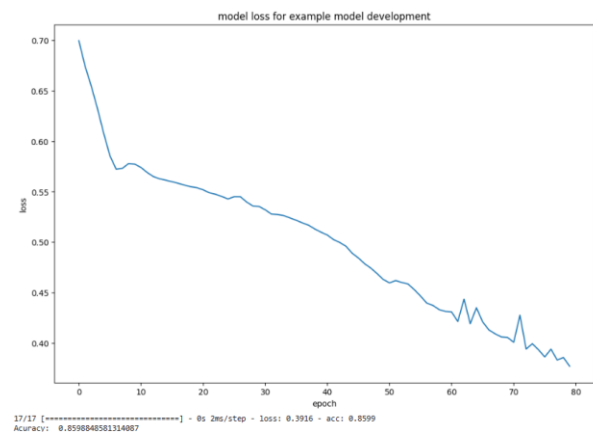


Figura 1- Model loss deep learning

```
0.8733205374280231
Predicted 0.0 1.0 All
Actual
0.0      357  28  385
1.0      38  98  136
All      395 126 521
```

Figura 2- accuracy decision tree

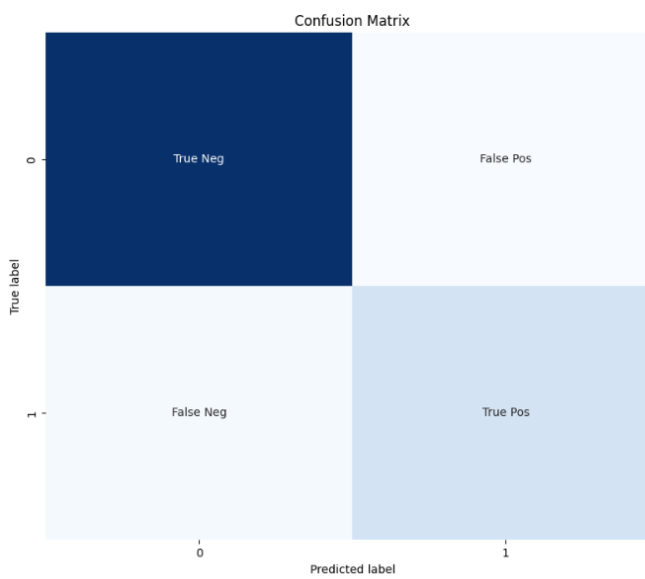


Figura 3- Confusion Matrix (Depression Y or N)

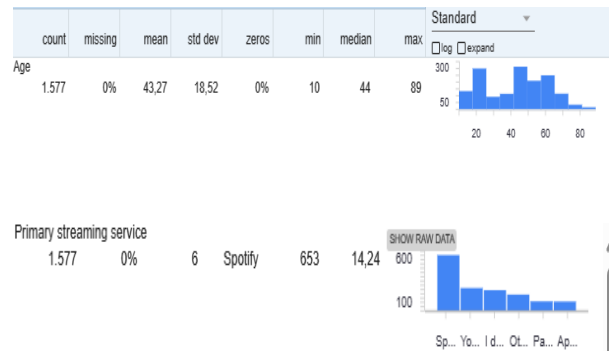


Figura 4- Some Data Bias