Klasifikacija zdravlja fetusa na temelju kardiotokograma

-dokumentacija koda-

Valentina Križ, Jelena Kurilić, Lucija Valentić

Prirodoslovno-matematički fakultet

Sveučilište u Zagrebu

Prije pokretanja jupyter bilježnica potrebno je instalirati XGBoost i imbalanced-learn pomoću priloženih uputa za instalaciju.

Cijeli projekt podijeljen je 3 bilježnice. U 'Eksploratorna analiza podataka.ipynb' nalazi se eksploratorna analiza odabranog skupa podataka. U 'treniranje.ipynb' odvija se podjela skupa podataka u train i test skupove i treniranje odabranih modela te spremanje najboljih modela u pickle datoteke. Pretpostavlja se da se sve datoteke i bilježnice nalaze u istom direktoriju. Usporedba dobivenih modela i grafički prikaz rezultata nalazi se u 'projekt\_CTG.ipynb'.

Redoslijed ćelija u pojedinim bilježnicama odgovara potrebnom redoslijedu izvršavanja koda.

|  |  |
| --- | --- |
| **Biblioteka** | **Opis korištenja** |
| pandas [1] | učitavanje i korištenje podataka |
| numpy [2] | spremanje podataka |
| sklearn [3] | algoritmi strojnog učenja (SVC i Random Forest) |
| imblearn [4] | oversampling metode |
| xgboost [5] | algoritam strojnog učenja (XGBoost) |
| pickle [6] | spremanje i učitavanje modela |
| matplotlib [7] | vizualizacija rezultata |
| seaborn [8] | vizualizacija rezultata |

Table Opis korištenih biblioteka

Reference:

[1] <https://pandas.pydata.org/>

[2] <https://www.numpy.org/>

[3] <https://scikit-learn.org/stable/>

[4] <https://imbalanced-learn.readthedocs.io/en/stable/>

[5] <https://xgboost.readthedocs.io/en/latest/>

[6] <https://docs.python.org/3/library/pickle.html>

[7] <https://matplotlib.org/>

[8] <https://seaborn.pydata.org/>