

1. Instalacja:

Do uruchomienia wymagany jest python. W celu uruchomienia projektu należy sklonować repozytorium git. Następnie należy zainstalować wymagane zależności. Można w tym celu wykorzystać plik `environment.yml` znajdujący się w folderze głównym repozytorium lub ręcznie zainstalować następujące biblioteki:

- pandas
- unidecode
- matplotlib
- mplsoccer
- xgboost
- dalex

Wymagane jest również podanie ścieżki do folderu danych w pliku `path.py` w folderze `data_processing`.

2. Użytkowanie

W celu wytrenowania modelu konieczne jest najpierw uruchomienie skryptów *generateTrainingData* oraz *players_minutes* z folderu `data_processing`, które wygenerują pliki z danymi pozwalających przygotować dane treningowe. Możliwe jest również uruchomienie kodu z pliku *read_merge* z wizualizacją sytuacji boiskowej.

Następnie można uruchomić notatniki jupyter i wygenerować model oraz dodatkowe pliki do wizualizacji oraz analizy.

Dodatkowo, konieczne jest umieszczenie plików z przetworzonymi odbiorami oraz przechwytyami o nazwach odpowiednio *complete_data_tackles.json* oraz *complete_data_interceptions.json*, a także wygenerowanego pliku *players_minutes_played.json* w folderze `notebooks`.

UWAGA: Notatniki powstały na podstawie znacznie większej ilości danych w celu wytrenowania i ewaluacji modelu, dlatego ponowne uruchomienie komórek notatnika może prowadzić do innych wyników.

Dostępna jest również aplikacja PowerBI umożliwiająca wizualizację wyników. Część wizualizacji wymaga pełnego zbioru danych w związku z czym nie będzie widoczna.