TalentFinder - Architektura Rozwiązania

# Spis treści

1. Wprowadzenie

2. Ogólna Architektura Systemu

3. Moduły Systemu

4. Zależności i Narzędzia

5. Komunikacja między Modułami

# 1. Wprowadzenie

TalentFinder to projekt mający na celu analizę danych piłkarskich oraz ocenę talentów zawodników na podstawie zgromadzonych informacji z różnych źródeł. Dokumentacja ta opisuje architekturę system, w celu zrozumienia struktury projektu oraz dalszy rozwój i konserwację systemu.

# 2. Ogólna Architektura Systemu

System TalentFinder składa się z kilku modułów, z których każdy odpowiada za inny aspekt przetwarzania danych i analizy. Moduły te współpracują ze sobą, aby zbierać, przetwarzać i analizować dane piłkarskie. Moduły te to *scrappers*, *edit\_data*, *database* i *ml*.

# 3. Moduły Systemu

**1. Moduł Zbierania Danych (Scrappers)**  
- Odpowiedzialny za pobieranie danych z zewnętrznych źródeł, takich jak sofascore i transfermarkt, wraz z oczyszczeniem danych ze zbędnych informacji.  
- Skrypty:  
 - scrappers/sofascore/get\_players\_ids.py  
 - scrappers/sofascore/get\_ratings\_from\_sofascore.py  
 - scrappers/transfermarkt/transfermarkt\_scrapper.py  
  
**2. Moduł Przetwarzania Danych (Edit\_data)**  
- Odpowiedzialny za przetwarzanie zebranych danych i przygotowanie ich do analizy.  
- Skrypty:  
 - database/add\_flags.py  
 - database/flags\_KM.py  
 - database/merge\_two\_csv\_with\_flags.ipynb

**3. Baza Danych (Database )**  
- Baza danych, w której przechowywane są zgromadzone dane.  
  
**4. Moduł Analizy Danych (ML)**  
- Odpowiedzialny za analizę danych i trenowanie modeli predykcyjnych.  
- Skrypty:  
 - ml/random\_forest/random\_forest.ipynb   
 - *ml/* *clasification /* *Clasification\_transfer.ipynb*  
 - *ml/* *regresion /* *Regresion\_transfers.ipynb*

# 4. Zależności i Narzędzia

Projekt TalentFinder wykorzystuje szereg narzędzi i bibliotek, które są niezbędne do jego prawidłowego działania. Wszystkie zależności są wymienione w pliku *requirements.txt*.  
  
**Przykładowe zależności:**  
- pandas  
- numpy  
- requests  
- beautifulsoup4  
- scikit-learn  
- jupyter  
- matplotlib  
  
Aby zainstalować wszystkie zależności, wykonaj:  
*pip install -r requirements.txt*

# 5. Komunikacja między Modułami

Komunikacja między modułami odbywa się poprzez wymianę danych w formacie CSV oraz poprzez połączenie z bazą danych. Każdy moduł zapisuje swoje wyniki do odpowiednich plików, które następnie są wykorzystywane przez inne moduły.