

Dokumentace

Název projektu: GOAL

Autor: Nguyen Hoang Lan

Třída: C4c

Kontaktní údaje:

- školní email: nguyen4@spsejecna.cz
- osobní email: honzik.nguyen@post.cz

Datum vypracování: 5.4.2025

Střední průmyslová škola elektrotechnická, Praha 2, Ječná 30

Školní rok 2024/2025

Aplikace zobrazuje zápasy italské Serie A – jak odehrané, tak nadcházející – získané z Football-Data.org API. Pro odehrané zápasy lze zobrazit detaily včetně statistik. Pomocí předtrénovaného modelu `RandomForest` lze predikovat výsledek zápasu.

Data Collection

Pro analýzu byly vybrány zápasy pěti nejlepších evropských lig, tedy Premier League, La Liga, Serie A, Ligue 1 a Bundesliga. Vybrané zápasy se hrály od začátku sezóny 2023/2024 do sezóny 2024/2025. Tyto ligy byly vybrány, protože se jedná o největší ligy na světě a ze zápasů odehraných v těchto ligových soutěžích je k dispozici obrovské množství dat.

Ke sběru dat byla použita data z <https://www.football-data.org>, která poskytla většinu údajů, jež se skládaly ze základních informací o zápasech, jako je datum, místo konání a skóre, ale také z podrobnějších údajů o střelbě, držení míče a počet faulů.

Závislosti

- `tkinter`
 - `requests`
 - `threading`
 - `pickle`
 - `scikit-learn` (pro vytvoření modelu, ne nutné pro běh)
 - `football-data.org` API účet (pro získání API klíče)
-

API

- **Klíč:** `API_KEY = '9be9843eb46e4df7b8bb5b8ca44adb00'`
 - **Base URL:** `https://api.football-data.org/v4/`
 - **Hlavní endpointy:**
 - `GET /competitions/SA/matches` – Zápasy Serie A
 - `GET /matches/{match_id}` – Detaily zápasu
-

Struktura kódu

1. `FootballAPIClient`

Zajišťuje komunikaci s API.

- `get_serie_a_matches()` – Získá všechny zápasy Serie A.
- `get_match_details(match_id)` – Detailní info o konkrétním zápasu.

2. `SerieAModel`

Zpracovává získaná data.

- `get_match_groups()` – Rozdělí zápasy na odehrané a budoucí.
- `get_match_detail(match_id)` – Vrací podrobnosti o konkrétním zápasu.

3. `MainView` (GUI)

Tkinter GUI pro zobrazení zápasů a interakci s uživatelem.

- Listboxy pro odehrané a budoucí zápasy
- Detailní okno s informacemi o zápasu
- Tlačítko pro spuštění predikce
- `predict_match()` – Vytáhne příznaky a použije `RandomForest` model k předpovědi výsledku

4. **AppController**

Řídí logiku aplikace.

- Načítá zápasy do GUI
- Na požádání zobrazuje detaily o zápasu

Predikční model

- Uložen v souboru `random_forest_model.pkl`
- Načítá se v hlavní třídě `MainView` při startu
- Vstup: 24 statistických příznaků zápasu (držení míče, střely, fauly, karty atd.)
- Výstup: `0` (výhra domácích), `1` (remíza), `2` (výhra hostů)

Uživatelské rozhraní

- **Hlavní okno:**
 - Seznam odehraných a nadcházejících zápasů
 - Dvojklik → otevře detaily zápasu
- **Okno detailu zápasu:**
 - Textové okno se statistikami
 - Tlačítko pro spuštění predikce

Spuštění aplikace

```
python app.py
```

Soubor `random_forest_model.pkl` musí být přítomen ve stejném adresáři jako aplikace.

Random Forest Classifier

Logistic Regression

Gradient Boosting Classifier

K-Nearest Neighbors Classifier