Gasser Ahmed
gasser18@vt.edu
ECE 5984, Project Proposal

KickStats: Visualizing Live Football Scores

Purpose

The purpose of this project is to provide football enthusiasts with real-time access to live scores of multiple football matches from various leagues and competitions. It addresses the problem of football fans needing a convenient and centralized platform to track ongoing matches, their scores, and relevant statistics. By offering a streamlined visualization of live scores, this project aims to enhance the football-watching experience and keep fans informed.

Key Questions and Objectives

1. **Real-time Score Updates:** How can we efficiently collect and deliver live scores from multiple football matches?
2. **Data Integration:** How can we seamlessly integrate data from the API-Football API into our data pipeline?
3. **Data Storage:** What is the best way to store the real-time football match data for future analysis and visualization?
4. **Data Transformation:** How can we transform the raw data into a format suitable for analysis and visualization?
5. **Visualization:** What tools and methods can we use to create user-friendly and interactive visualizations of live football match scores?

Visualization

1. **Real-time Scoreboard**:
    - A dynamic Tableau scoreboard displays live match scores.
    - Scores update in real-time using Tableau's live data connections to the football API.
2. **Interactive Dashboards:**
    - Filtering matches by league, date, or team.
    - Sorting matches by score, time remaining, or match status.
    - Viewing detailed match info by clicking on specific matches.
3. **Match Details:**
    - When users click on a specific match in the scoreboard or use filters, display match details in an interactive visualization.
    - Include team logos, goal scorers, statistics (e.g., possession, shots on goal), and any relevant events (e.g., yellow/red cards).

Pipeline and Methods

For this project, we will use the following Stream-Visualization pipeline for the API-FOOTBALL dataset:

1. **Data Ingestion:** We will retrieve real-time football match data from the API-Football API.
2. **Stream Ingestion (Extract and Load):** Using Apache Kafka, we will ingest and stream the live data to ensure we have the latest scores and updates.
3. **Data Storage:** The ingested data will be stored in Amazon S3, forming a data lake that allows for scalable and cost-effective storage.
4. **Data Transformation:** Pandas, a powerful Python library for data manipulation, will be used to transform the raw data into a structured format suitable for analysis.
5. **Data Warehousing:** The transformed data will be further stored in an Amazon S3 bucket, serving as our data warehouse for historical data analysis.
6. **Relational Database:** We will use Amazon RDS or MySQL to maintain a relational database for structured data storage and querying.
7. **Data Analytics:** Tableau will be employed to create interactive dashboards and visualizations, enabling users to access live football match scores with ease.

Expected Results

The final result of this project will be a user-friendly and interactive platform that provides real-time access to football match scores from various leagues and competitions. Users will be able to:

- View live scores of multiple football matches simultaneously.
- Access match details, including team information, goal scorers, and match statistics.
- Customize their view to focus on specific matches or leagues of interest.
- Interact with dynamic visualizations and charts.
- Stay updated with the latest scores without the need for constant manual refreshing.

By achieving these objectives, we aim to offer football fans a convenient and engaging solution for tracking live scores and enjoying the excitement of football matches in real-time.