SOCCER PLAYER VALUATION AND PREDICTION PROJECT - by TRUNG TRAN

CREATE A DATA PIPELINE TO UPDATE DATASET.

Create an account and get an API key from API-Football.com.
Install the requests library in Python using pip. Requests is a popular library for making HTTP requests to APIs.
Write a Python script that calls the API to get the data you need.
*** Free subscriptions do not allow many downloads. Dataset for this project will be downloaded from Kaggle.
Raw Data File was uploaded to Amazon S3.

DATA TRANSFORMATION USING PANDAS AND NUMPY

Reads the CSV dataset: reads the FIFA player data from CSV files and stores it in a dictionary named dataset.
Cleans up the dataset: The code drops irrelevant columns from the dataset, including player ID, photo, nationality, club logo, etc.
Adds a new column named BMI and drops some more columns.
Creates a dictionary of preferred foot values.
Merges and groups the data: merge the data from all years into one dataset and group it by player name to get the mean values for each player.
Combined file is saved on S3

DATA VISUALIZATION USING MATPLOTLIB AND SEABORN

Visualizes the data: use Seaborn and Matplotlib libraries to create various visualizations of the data, such as bar charts, box plots, and line plots.
Creates a radar plot: create a radar plot using Matplotlib to visualize the average values of some player attributes like age, overall, potential, value, and wage.
All files are saved on Amazon S3

MACHINE LEARNING USING PYSPARK

Apply linear regression, random forest, and decision tree model to find the best fit.
Do some more data cleaning and tuning to reevaluate the models

BUILD A WEB APPLICATION WITH FLASK

User interface is a website, hosted on AWS.
This website has information about the project.
Also, data visualization will be accessible
User can play a little game, where they enter player's name, it will print out the player's value

USING DOCKER TO PACK THE APPLICATION

USING AMAZON KUBERNETES SERVICES TO HOST THE APPLICATION AND RUN THE APPLICATION