**Sports Injury Analysis**

**Overview**

This project focuses on analyzing sports injury data to predict whether an athlete is likely to be injured based on various features such as workload, position, and date. The analysis is performed using a Jupyter Notebook (finalMac.ipynb), which includes data preprocessing, model implementation, training, and evaluation.

The goal is to build a machine learning model that can accurately predict injuries, helping teams and coaches make data-driven decisions to prevent injuries and optimize player performance.

**Table of Contents**

1. [Dataset](https://chat.deepseek.com/a/chat/s/e2b66c6b-394c-4daf-8cc1-3234524dae65#dataset)
2. [Requirements](https://chat.deepseek.com/a/chat/s/e2b66c6b-394c-4daf-8cc1-3234524dae65#requirements)
3. [Installation](https://chat.deepseek.com/a/chat/s/e2b66c6b-394c-4daf-8cc1-3234524dae65#installation)
4. [Usage](https://chat.deepseek.com/a/chat/s/e2b66c6b-394c-4daf-8cc1-3234524dae65#usage)
5. [Models Implemented](https://chat.deepseek.com/a/chat/s/e2b66c6b-394c-4daf-8cc1-3234524dae65#models-implemented)
6. [Results](https://chat.deepseek.com/a/chat/s/e2b66c6b-394c-4daf-8cc1-3234524dae65#results)
7. [Contributing](https://chat.deepseek.com/a/chat/s/e2b66c6b-394c-4daf-8cc1-3234524dae65#contributing)
8. [License](https://chat.deepseek.com/a/chat/s/e2b66c6b-394c-4daf-8cc1-3234524dae65#license)

**1.Dataset**

The dataset used in this project is stored in a CSV file named final2.csv. It contains the following columns:

* athlete\_id(footballer\_id): Unique identifier for each footballer.
* date: Date of the record.
* postion: Position of the footballer (e.g., midfielder, attacker).
* value: A numerical value associated with the footballer
* game\_workload: Workload of the footballer in the game.
* injury: Indicates whether the footballer was injured (yes or No).

**2.Requirements**

To run the Jupyter Notebook or Python script, you need the following libraries installed:

* Python 3.8+
* pandas
* numpy
* scikit-learn
* matplotlib
* seaborn
* xgboost
* tensorflow (optional, for neural networks)

**3.Usage**

**Running the Jupyter Notebook**

1. Open the Jupyter Notebook:
2. Execute the cells in the notebook to perform data preprocessing, model training, and evaluation.
3. Running the Python Script

If you have converted the notebook to a Python script (finalMac.py), you can run it directly:

**4.Models Implemented**

The following machine learning models are implemented and evaluated in this project:

1. **Random Forest Classifier**
2. **Logistic Regression**
3. **Support Vector Machine (SVM)**
4. **XGBoost**
5. **Neural Network (using TensorFlow/Keras)**
6. **K-Nearest Neighbors (KNN)**

**7.Contributing**

Contributions to this project are welcome! If you'd like to contribute, please follow these steps:

1. Fork the repository.
2. Create a new branch for your feature or bugfix.
3. Commit your changes and push them to your fork.
4. Submit a pull request with a detailed description of your changes.

**Acknowledgments**

* The dataset used in this project is sourced from [source-name].
* Special thanks to the open-source community for providing the libraries and tools used in this project.