DOUGLAS COLLEGE

Winter 2025

Midterm Report: Sports Management Portal for Injury Prevention, Performance, and Athlete Management Using Machine Learning

Prepared for

CSIS4495 Applied Research Project
Section 002

Link to the demo: https://youtu.be/HI5ohnRODJM

Presented to
Padmapriya Arasanipalai Kandhadai
CSIS4495 Section 002

Presented by
Juan Carlos Katigbak
300366535

Table of Contents

Introduction

I. Domain, Overview of Background, and Context of the Research

Sports injuries are a normal thing in everyday life, arising from interactions between modifiable factors (e.g., training load, strength) and non-modifiable factors (e.g., age, injury history) that fluctuate over time. Sports injuries happen to the so-called “weekend warrior” trying to squeeze in some form of exercise which could be some form of sports activity in an attempt to improve fitness in spite of such a busy schedule. The “weekend warrior” is usually at a high risk for injuries that are just waiting to happen. Then you have the athletes, both amateur and professional, who play their respective sports at a high-level which can be demanding and taxing to their bodies to the point of exhaustion and that too can lead to injuries. It is a fact of life indeed in spite of all the conditioning, training, and nutrition present nowadays to be able to avoid such sports injuries.

With the advent of Artificial intelligence (AI) and Machine Learning (ML), sports injuries can be predicted which can lead to injury prevention. The application of AI and ML in sports has grown significantly in recent years which can lead to advantages for both athletes and coaches in their respective sport. AI has been a boon in performance forecasting and injury prevention in various sports. Examples include optimizing training regimens and evaluating injury risks through advanced algorithms.

II. Framing the Problem

Here are some questions that the student has come up with to frame the problem:

1. How can ML improve injury prediction and performance optimization in sports?
2. What methods can provide data-driven insights for athletes?

Injuries significantly impact athletes’ performance thus a reliable injury prediction system can reduce risks and improve overall athlete performance.

III. Relevant Literature and Research

In a study done by researchers Robin Owen, Julian A. Owen, and Seren L. Evans[1], they mentioned that preventing injuries is important to be successful in sports and a lot of investment is made to be able to do this. Despite all of these, injuries still remain a significant challenge. Millions of athletes around the world suffer from injuries every year even if there are scientifically backed prevention strategies. The researchers point out that this sport injury prediction is something that can still be improved and this is where Artificial Intelligence (AI) comes into play. AI offers a promising path forward to provide a comprehensive view of the complexity of sports injuries. Recent advancements of AI and Machine Learning (ML) have become a pivotal part of sports medicine research which provides a better analysis of huge quantities of data to be able to make predictive models of injury. AI and ML can handle imbalanced datasets common in sports injury studies and account for risk factors, assessing their role in predicting injuries.

The same AI and ML technologies are used in one of the major sports leagues in the world which is based in North America and this is the National Basketball Association or the NBA for short. In another study done by researchers Yining Lu, Ayoosh Pareek, Ophelie Z. Lavoie-Gagne, Enrico M. Forlenza, Bhavik H. Patel, Anna K. Reinholz, Brian Forsythe, and Christopher L. Camp[2], they wanted to know about time-loss lower extremity muscle strains (LEMSs) in the NBA and to determine the validity of a ML model in predicting injury risk. Their study involved using various ML models some of which outperformed other ML models by means of achieving high accuracy and reliability. The ML models that did well provided actionable insights which can help physicians assess individual injury risks and inform prevention strategies through a web-based application. The findings highlight the potential of ML to revolutionize injury prevention and management in sports, though further validation and improvement are needed for broader application.

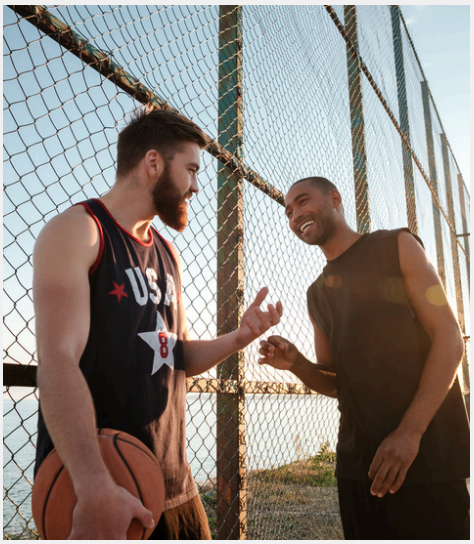
IV. Hypotheses, Assumptions, and Potential Benefit of Research

An integrated sports management portal powered by ML will significantly improve the accuracy of injury prediction and provide insights to optimize athlete performance. The portal will leverage algorithms to analyze both modifiable (e.g., training load, strength) and non-modifiable factors (e.g., age, injury history), enabling athletes to make informed decisions.

Assumptions include the availability of high-quality datasets such as those from the NBA containing relevant variables like athlete demographics, performance metrics, and injury histories. It also assumes that the adoption of the portal by athletes will align with the growing reliance on data-driven approaches in sports.

The potential benefits of this research are immense. First of all, the system is expected to reduce athlete downtime by accurately predicting injury risks and recommending preventive measures tailored to individual needs. Furthermore, the portal will improve training outcomes by enabling people like coaches to adjust regimens based on real-time performance data and risk assessments. Lastly, the research will contribute to the development of a user-friendly and customizable platform, empowering athletes to track their progress and coaches to manage multiple profiles efficiently.

By addressing the gaps in current injury prevention and performance tracking tools, this research aims to advance sports management technology and contribute to safer and more efficient athletic practices. In the long term, the portal could serve as a model for broader adoption across various sports and training environments.



Summary of Initial Proposed Research Project

I. Research Design and Objectives

Design

Develop a web-based portal incorporating machine learning (ML) models for injury prediction, performance analytics, and training customization.

Objectives

a. Create a database to store athlete profiles and performance metrics.
b. Train an ML model to predict injury risks using historical data.
c. Build interactive dashboards for visualizing performance trends.

II. Methodology

Data Collection

Sources include publicly available datasets from the NBA and synthetic data. Possible features include age, previous injuries, physiological metrics. Surveys and interviews of some individuals that are involved with basketball will also be conducted as part of the data collection.

Data Analysis

Use machine learning algorithms/models for injury prediction and trend analysis.

Justification

ML models done using NBA data such as in the study of Yining Lu, Ayoosh Pareek, Ophelie Z. Lavoie-Gagne, Enrico M. Forlenza, Bhavik H. Patel, Anna K. Reinholz, Brian Forsythe, and Christopher L. Camp[2] show that this proposed research project can be done as well as make it more holistic by means of including performance and training with injury prediction thus making the application broader.

III. Technologies

Operating System/Platform
MacOS and Windows 11

Programming Languages/Frameworks
For the backend, it will be Python (machine learning algorithms to be decided)
For the frontend, it will be HMTL, CSS, and JavaScript
Possibility of using new programming languages/frameworks not yet learned by the student

Database
Database Management System (still deciding if MS SQL Server or MySQL will be used)

IV. Expected Results

The expected results are:

1. Accurate injury risk predictions with recommendations
2. Practical applications which people like athletes and coaches can use

Changes to the Proposal

II. Key Changes and Justifications

1. Changes in Technology Stack

MS SQL Server or MySQL was originally considered but SQLite is now being used due to the Django web framework being used. Given that the project is still in development, SQLite is a lightweight and effective choice for prototyping and testing. Frontend now using Bootstrap so that it looks cleaner and more responsive. The current proof of concept of the app has no machine learning model per se, only using predefined conditions akin to the Decision Tree algorithm but will definitely use machine learning stack consisting of libraries such as scikit-learn for predictive modeling, pandas & NumPy for data processing, and Matplotlib & Seaborn for data visualization. The current proof of concept is using these libraries with scikit-learn as the next one to be integrated.

2. Changes in Timelines

Machine learning model implementation was expected in early phases but is the next one after database refinement and frontend development. The proof of concept shows that foundational elements like database and UI need further improvement before integrating machine learning models. Some parts of PHASE 2: Crash Course Coding, Some Implementation, and Data Collecting especially in the data collecting and crash course coding had to be pushed to PHASE 3: Build Minimum Viable Product and Testing because there was not enough time to do both as this researcher has just begun diving into machine learning and is still learning which algorithms to use while looking for good datasets to make the proof of concept useful in the final phase.

Project Planning and Timeline

PHASE	DURATION	MILESTONES
PHASE 3: BUILD MINIMUM VIABLE PRODUCT AND TESTING	5 weeks	1. Crash Courses on Coding (Machine Learning Algorithms) 2. Finalizing Data Methodology (datasets, API, survey, training data for prediction) 3. Building athlete profiles and injury prediction 4. Building integration of performance and training features 5. Conducting testing
PHASE 4: SPORTS MANAGEMENT PORTAL PILOT AND DEPLOYMENT	2 weeks	1. Refine 2. Gather user feedback 3. Ready pilot and deployment

Implemented Features

1. Player Profile Management

Description

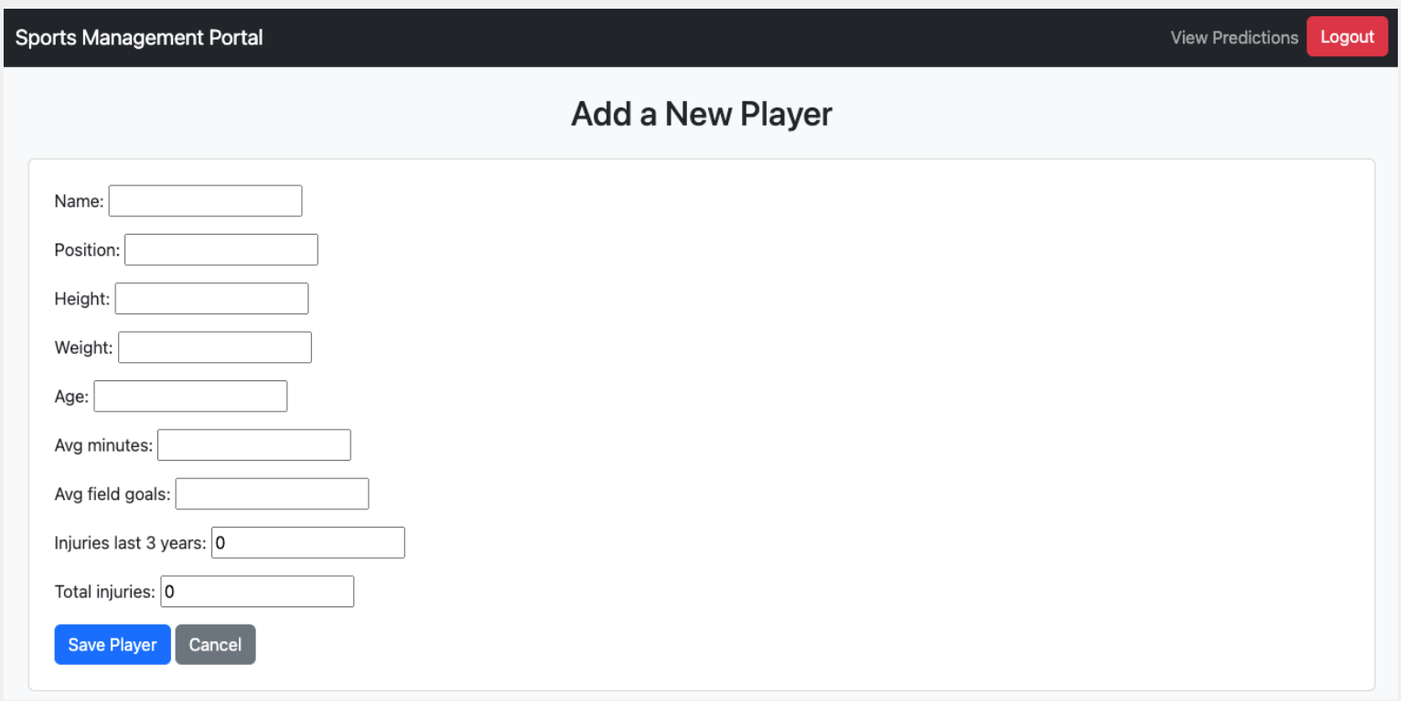
The system allows users to create and manage player profiles, tracking key performance metrics and injury history. Each player record includes details such as name, position, height, weight, average minutes played, and injury history.

Implementation Details

- **Model:** Player (Defined in models.py)
- **Key Fields:**
  name, position, height, weight, age
  avg_minutes, avg_field_goals
  injuries_last_3_years, total_injuries
  user (linked to Django User model)

Code

```
class Player(models.Model):
    name = models.CharField(max_length=100)
    position = models.CharField(max_length=50)
    height = models.IntegerField()
    weight = models.IntegerField()
    age = models.IntegerField()
    avg_minutes = models.FloatField()
    avg_field_goals = models.FloatField()
    injuries_last_3_years = models.IntegerField(default=0)
    total_injuries = models.IntegerField(default=0)
    user = models.ForeignKey(User, on_delete=models.CASCADE)
```



2. Authentication

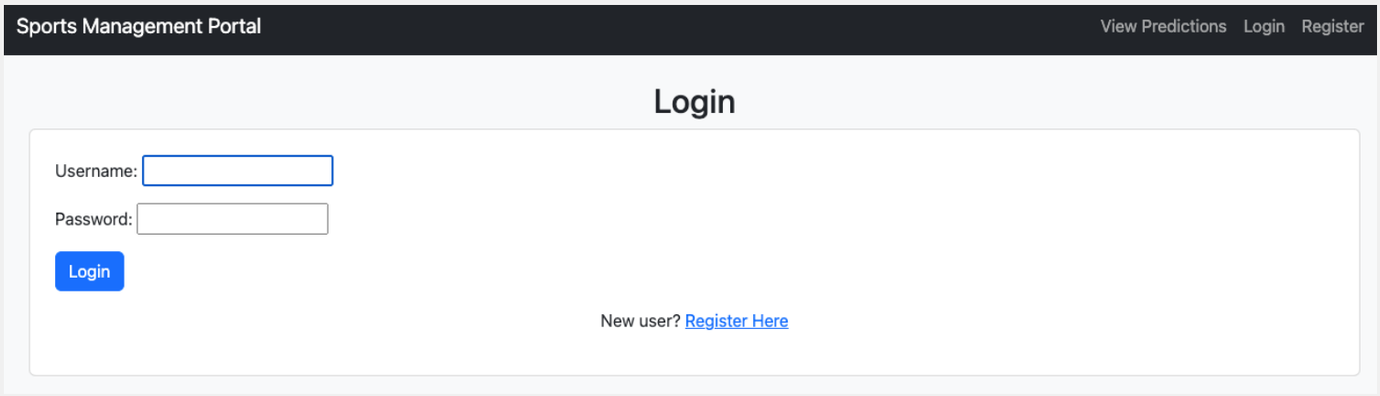
Description
The system allows authenticated users to manage their profiles securely. It utilizes Django's built-in authentication framework to ensure secure login and registration.

Implementation Details
- **User Registration:** Uses Django's UserCreationForm.
- **Login Required:** Ensures only authenticated users can add players.

Code
```
import matplotlib.pyplot as plt
import seaborn as sns
import io
import urllib, base64

def injury_trend_chart():
    plt.figure(figsize=(8, 5))
    sns.histplot(Player.objects.values_list('injuries_last_3_years',
flat=True), bins=5, kde=True)
    plt.xlabel('Number of Injuries (Last 3 Years)')
    plt.ylabel('Player Count')
    plt.title('Injury Distribution Among Players')

    buf = io.BytesIO()
    plt.savefig(buf, format='png')
    buf.seek(0)
    string = base64.b64encode(buf.read()).decode()
    return f'{string}'
```



3. Injury Risk Visualization (Matplotlib & Seaborn)
Description
The system generates injury risk insights using visualizations. It leverages Matplotlib and Seaborn to display player statistics and injury trends.

Implementation Details

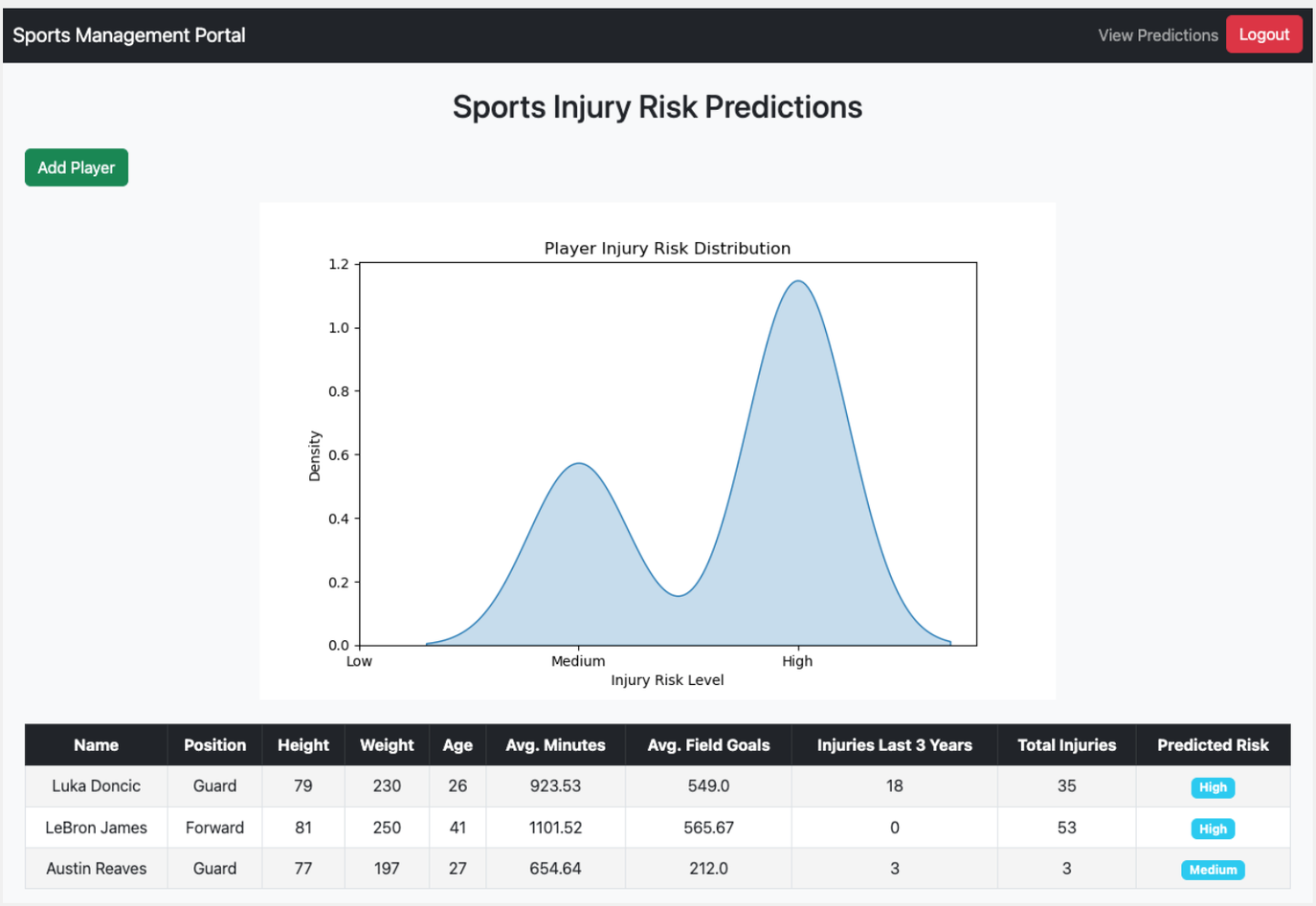
- **Graphs and Charts:** Generate insights on injury history and player performance.
- **Uses:** matplotlib.pyplot & seaborn.

Code

```
import matplotlib.pyplot as plt
import seaborn as sns
import io
import urllib, base64

def injury_trend_chart():
    plt.figure(figsize=(8, 5))
    sns.histplot(Player.objects.values_list('injuries_last_3_years',
flat=True), bins=5, kde=True)
    plt.xlabel('Number of Injuries (Last 3 Years)')
    plt.ylabel('Player Count')
    plt.title('Injury Distribution Among Players')

    buf = io.BytesIO()
    plt.savefig(buf, format='png')
    buf.seek(0)
    string = base64.b64encode(buf.read()).decode()
    return f'{string}'
```



Name	Position	Height	Weight	Age	Avg. Minutes	Avg. Field Goals	Injuries Last 3 Years	Total Injuries	Predicted Risk
Luka Doncic	Guard	79	230	26	923.53	549.0	18	35	High
LeBron James	Forward	81	250	41	1101.52	565.67	0	53	High
Austin Reaves	Guard	77	197	27	654.64	212.0	3	3	Medium

Work Date/ Hours Logged

Juan Carlos Katigbak		
Date	**Number of Hours**	**Description of work done**
February 14, 2025	1	Continued Proof of Concept with zipped folder named "learning_log.zip" containing a learning log app using Django Python web framework as a proof of concept for my project and this is found under NEW - Proof of Concept Part 2 of 4 folder and added it to Repo under Implementation folder. The Implementation folder also contains a folder called "NEW - Supporting Docs Proof of Concept Part 2 of 4" which contains file "Proof of Concept Part 2 of 4.png" which is a screenshot that contains a stopwatch log. I give acknowledgement to one of my book references [Matthes, E. (2023). Phyton crash course 3rd edition: A hands-on, project-based introduction to programming. No Starch Press.] for providing me with a step-by-step approach how I can create this full stack for my project.
February 17, 2025	8	Finished Proof of Concept Part 2 of 4 with zipped folder named "learning_log.zip" containing a learning log app using Django Python web framework as a proof of concept for my project and this is found under FINISHED - Proof of Concept Part 2 of 4 folder and added it to Repo under Implementation folder. The Implementation folder also contains a folder called "UPDATED - Supporting Docs Proof of Concept Part 2 of 4" which contains file "Proof of Concept Part 2 of 4 - February 17.png" which is a screenshot that contains a stopwatch log. I give acknowledgement to one of my book references [Matthes, E. (2023). Phyton crash course 3rd edition: A hands-on, project-based introduction to programming. No Starch Press.] for providing me with a step-by-step approach how I can create this full stack for my project.
February 18, 2025	5.5	Studied for Proof of Concept 3 of 4 which is basically trying to combine Django with Machine Learning to show data visualization (Proof of Concept 1 of 4 and 2 of 4 were learning and implementing Django to be used for my project) found in The Implementation folder containing a folder called "NEW - Supporting Docs Proof of Concept Part 3 of 4" which contains "NEW - Proof of Concept Part 3 of 4 - Feb 18.png" which is a screenshot that contains a stopwatch log - added to Repo. Updated file "Updated - References.docx" under Misc folder. I give acknowledgement to two of my website references [1. Braun, N. (2025). Fantasy Math: Fantasy sports predictions and analytics. Fantasy Math. https://fantasymath.com/; 2. Broni, K. (2020). KenBroTech. YouTube. https://www.youtube.com/@KenBroTech] for teaching me how to integrate Machine Learning data visualization with Django.
February 19, 2025	2	Further studying for Proof of Concept 3 of 4 to learn more about Numpy, Pandas, Matplotlib, and Seaborn found in The Implementation folder containing a folder called "Supporting Docs Proof of Concept Part 3 of 4" which contains "NEW - Proof of Concept Part 3 of 4 - Feb 19.png" which is a screenshot that contains a stopwatch log - added to Repo. I give acknowledgement to two of my website references [1. Braun, N. (2025). Fantasy Math: Fantasy sports predictions and analytics. Fantasy Math. https://fantasymath.com/; 2. Broni, K. (2020). KenBroTech. YouTube. https://www.youtube.com/@KenBroTech] for teaching me how to integrate Machine Learning data visualization with Django.
	4	Started Proof of Concept Part 4 of 4 doing the Machine Learning Process still in the Data Acquisition and Data Cleaning parts using datasets that I will be using for my proof of concept submission for the Midterms found in the Implementation folder containing folder called "NEW - Proof of Concept Part 4 of 4" which contains folder "NEW - NBA Los Angeles Lakers" which contains csv data of injury and stats of the Los Angeles Lakers players as well as a Jupyter notebook called "Lakers.ipynb" - added to Repo. Also found in the Implementation folder is a folder called "NEW - Supporting Docs Proof of Concept Part 4 of 4" which contains "NEW - Proof of Concept Part 4 of 4.png" which is a screenshot that contains a stopwatch log - added to Repo.
February 21, 2025	3	Still currently working on Proof of Concept Part 4 of 4 doing the Machine Learning Process still in the Data Acquisition and Data Cleaning parts using datasets that I will be using for my proof of concept submission for the Midterms.
February 22, 2025	4.5	A continuation of February 21, 2025, still currently working on Proof of Concept Part 4 of 4 doing the Machine Learning Process still in the Data Acquisition and Data Cleaning parts using datasets that I will be using for my proof of concept submission for the Midterms and whatever I have done so far I added in Repo and uploaded to folder "NEW - NBA Los Angeles Lakers" and deleted files there from February 19, 2025 because they are not going to be used anymore. Also found in the Implementation folder under folder called "NEW - Supporting Docs Proof of Concept Part 4 of 4" a file "NEW - Proof of Concept Part 4 of 4 - Feb 22.png" which is a screenshot that contains a stopwatch log - added to Repo.
February 23, 2025	7	Still currently working on Proof of Concept Part 4 of 4 integrating Django with Machine Learning to bring out injury risk prediction with a data visualization.
February 24, 2025	7	Finished Proof of Concept Part 4 of 4 having integrated Django with Machine Learning to bring out injury risk prediction with a data visualization and added to Repo in Implementation folder under FINISHED - PROOF OF CONCEPT in a zipped file called "lal_injury_risk.zip". Added 2 csv files in folder Proof of Concept Part 4 of 4 with datasets used for "lal_injury_risk.zip" - added to Repo.
	3	Finished Video Demo of Product and uploaded on YouTube with link saved in front page of Midterm Report. Finished "JuanCarlosK_MidtermReport.pdf", uploaded it on Blackboard and added it to Repo including finishing ReadMe file in Github.

Closing and References

I give acknowledgement to one of my book references [Matthes, E. (2023). Phyton crash course 3rd edition: A hands-on, project-based introduction to programming. No Starch Press.] for providing me with a step-by-step approach how I can create this full stack for my project. I also give acknowledgement to two of my website references [1. Braun, N. (2025). Fantasy Math: Fantasy sports predictions and analytics. Fantasy Math. https://fantasymath.com/;
2. Broni, K. (2020). KenBroTech. YouTube. https://www.youtube.com/@KenBroTech] for teaching me how to integrate Machine Learning data visualization with Django.

Here are my current references:

I. Research Journals and Studies
1. Owen, R., Owen, J. A., & Evans, S. (2024). Artificial intelligence for sport injury prediction. In Artificial Intelligence for Sport Injury Prediction. Springer.

2. Lu, Y., Pareek, A., Lavoie-Gagne, O. Z., Forlenza, E. M., Patel, B. H., Reinholz, A. K., Forsythe, B., & Camp, C. L. (2022). Machine learning for predicting lower extremity muscle strain in National Basketball Association athletes. The Orthopaedic Journal of Sports Medicine, 10(7), 23259671221111742.

3. Papageorgiou, G., Sarlis, V., & Tjortjis, C. (2024). Evaluating the effectiveness of machine learning models for performance forecasting in basketball: A comparative study. Knowledge and Information Systems, 66, 4333–4375.

4. Chmait, N., & Westerbeek, H. (2021). Artificial intelligence and machine learning in sport research: An introduction for non-data scientists. Frontiers in Sports and Active Living, 3, Article 682287.

5. Dindorf, C., Bartaguiz, E., Gassmann, F., & Fröhlich, M. (2022). Conceptual structure and current trends in artificial intelligence, machine learning, and deep learning research in sports: A bibliometric review. Preprint.

6. Li, S., & Zhang, W. (2022). Evaluation method of basketball teaching and training effect based on wearable devices. Frontiers in Physics, 10, Article 900169.

7. Yang, X. (2024). Construction of measurement index system of basketball players' specific physical fitness training based on AI intelligence and neural network. Molecular & Cellular Biomechanics, 21(1), 250.

8. Georgievski, B., & Vrtagic, S. (2021). Machine learning and the NBA game. Journal of Physical Education and Sport.

9. Simenz, C. J., Dugan, C. A., & Ebben, W. P. (2005). Strength and conditioning practices of National Basketball Association strength and conditioning coaches. The Journal of Strength and Conditioning Research, 19(3), 495–504.

II. Books
1. Braun, N. (2023). Learn to code with basketball.

2. Matthes, E. (2023). Phyton crash course 3rd edition: A hands-on, project-based introduction to programming. No Starch Press.

3. Foran, B. (2026). Complete conditioning for basketball: National Basketball Strength & Conditioning Association. Human Kinetics.

4. Gillett, J. & Burgos, B. (2020). Strength Training for Basketball: National Strength and Conditioning Association. Human Kinetics.

III. Websites
1. Braun, N. (2025). Fantasy Math: Fantasy sports predictions and analytics. Fantasy Math. https://fantasymath.com/

2. Broni, K. (2020). KenBroTech. YouTube. https://www.youtube.com/@KenBroTech

3. Gilermo, D. (n.d.). NBA players stats [Dataset]. Kaggle. https://www.kaggle.com/datasets/drgilermo/nba-players-stats/data

4. Lauton, L. (n.d.). NBA injury stats (1951-2023) [Dataset]. Kaggle. https://www.kaggle.com/datasets/loganlauton/nba-injury-stats-1951-2023

5. Lounias, T. (n.d.). NBA stat scrape and analysis [Kaggle Notebook]. Kaggle. https://www.kaggle.com/code/themelissoulounias/nba-stat-scrape-and-analysis