Athletic Runner Injury Prediction System

T Akhila,A Divya Sri,K Sunidhi,B Amulya

Under the esteemed guidance of

Dr. P. Kayal

Associate Professor

Bachelor of Technology

Department of Information Technology

BVRIT HYDERABAD College of Engineering for Women

April 5, 2024

Contents

Overview

- By leveraging advanced machine learning techniques, we empower athletes to proactively manage their well-being, optimize performance, and contribute to the broader goals of health, innovation, and sustainability in sports and society.

Problem Statement

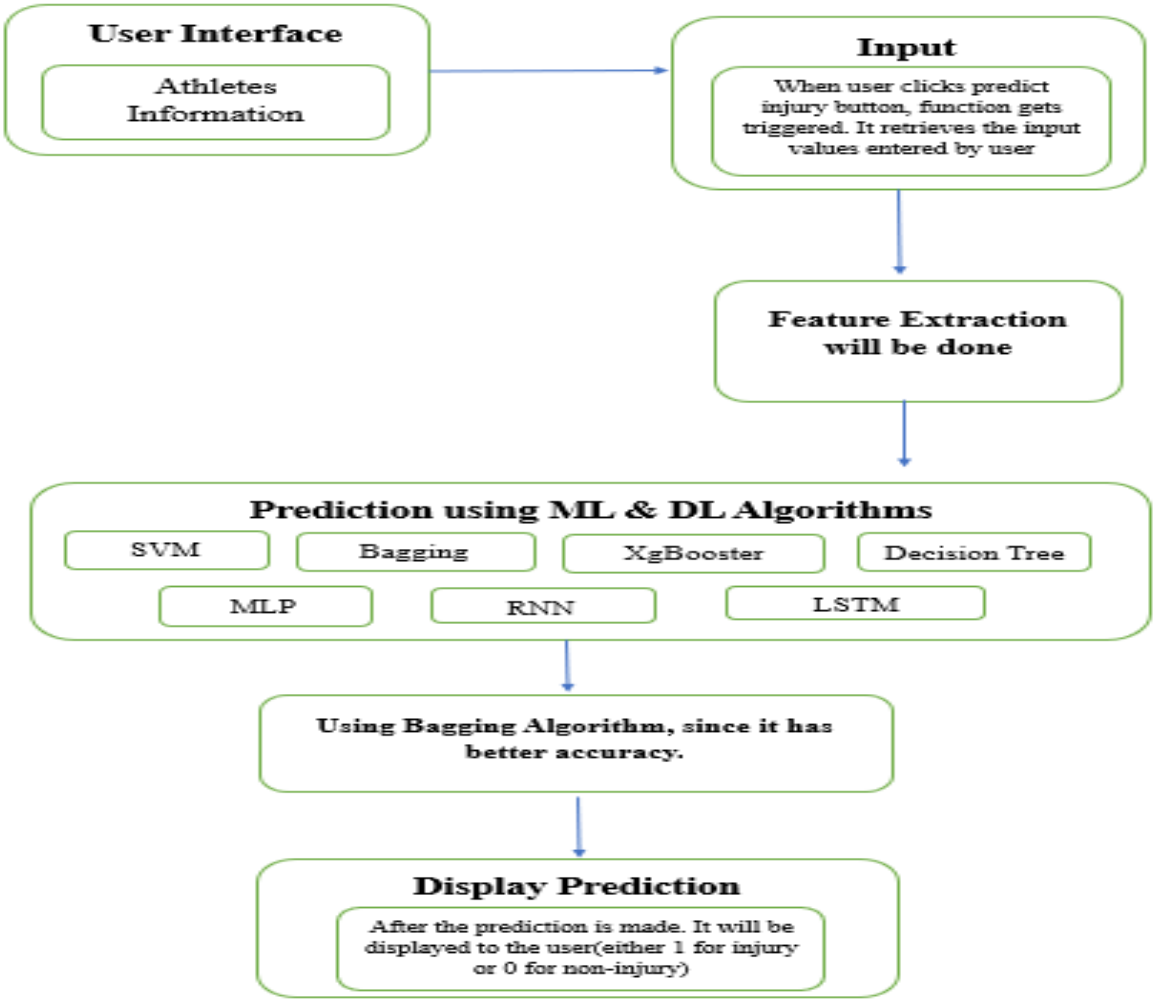
- Running is one of the most popular sports in the world.
- 60 million people participated in jogging, running, or trail running in America alone in 2017. But it is reported that 50 percent of runners get injured every year.

Objectives

Model Development

- Performing feature extraction and identifying the feature importance for Injury Prediction For Athletic Runners.
- Building model for injury prediction for athletic runners using machine learning techniques.

Graphical Representation



Dataset

- The data set consists of a detailed training log from a Dutch high-level running team over a period of seven years (2012-2019).

	nr. sessions	nr. rest days	total kms	max km one day	total km Z3-Z4-Z5-T1-T2	nr. tough sessions (effort in Z5, T1 or T2)	nr. days with interval session	total km Z3-Z4	max km Z3-4 one day	total km Z5-T1-T2	...	max training success.2	avg recovery.2	min recovery.2	max recovery.2	Athlete ID	injury	rel total kms week 0_1	rel total kms week 0_2	rel total kms week 1_2	Date
0	5.0	2.0	22.2	16.4	11.8	1.0	2.0	10.0	10.0	0.6	...	0.0	0.18	0.16	0.20	0	0	0.718447	1.378882	1.919255	0
1	5.0	2.0	21.6	16.4	11.7	1.0	2.0	10.0	10.0	0.5	...	0.0	0.18	0.16	0.20	0	0	0.683544	1.018868	1.490566	1
2	5.0	2.0	21.6	16.4	11.7	1.0	2.0	10.0	10.0	0.5	...	0.0	0.17	0.16	0.18	0	0	0.683544	1.018868	1.490566	2
3	5.0	2.0	21.6	16.4	11.7	1.0	2.0	10.0	10.0	0.5	...	0.0	0.18	0.16	0.18	0	0	0.683544	1.018868	1.490566	3

Methodology

- Data Preprocessing:
  - Reading the data
  - Splitting data
  - Dimensionality reduction
- Exploratory data analysis:
  - This involves visualizing the distribution of injury and non-injury.
  - Exploring correlations between features and the target variable(Injury).

- Model building:
  - SVM
  - Bagging
  - XgBooster
  - Decision tree
  - MLP
  - RNN
  - LSTM
- Evaluation:
  - Confusion matrices to visualize the performance of each model, particularly in terms of true positives, true negatives, false positives, and false negatives.
  - Accuracy for each model.

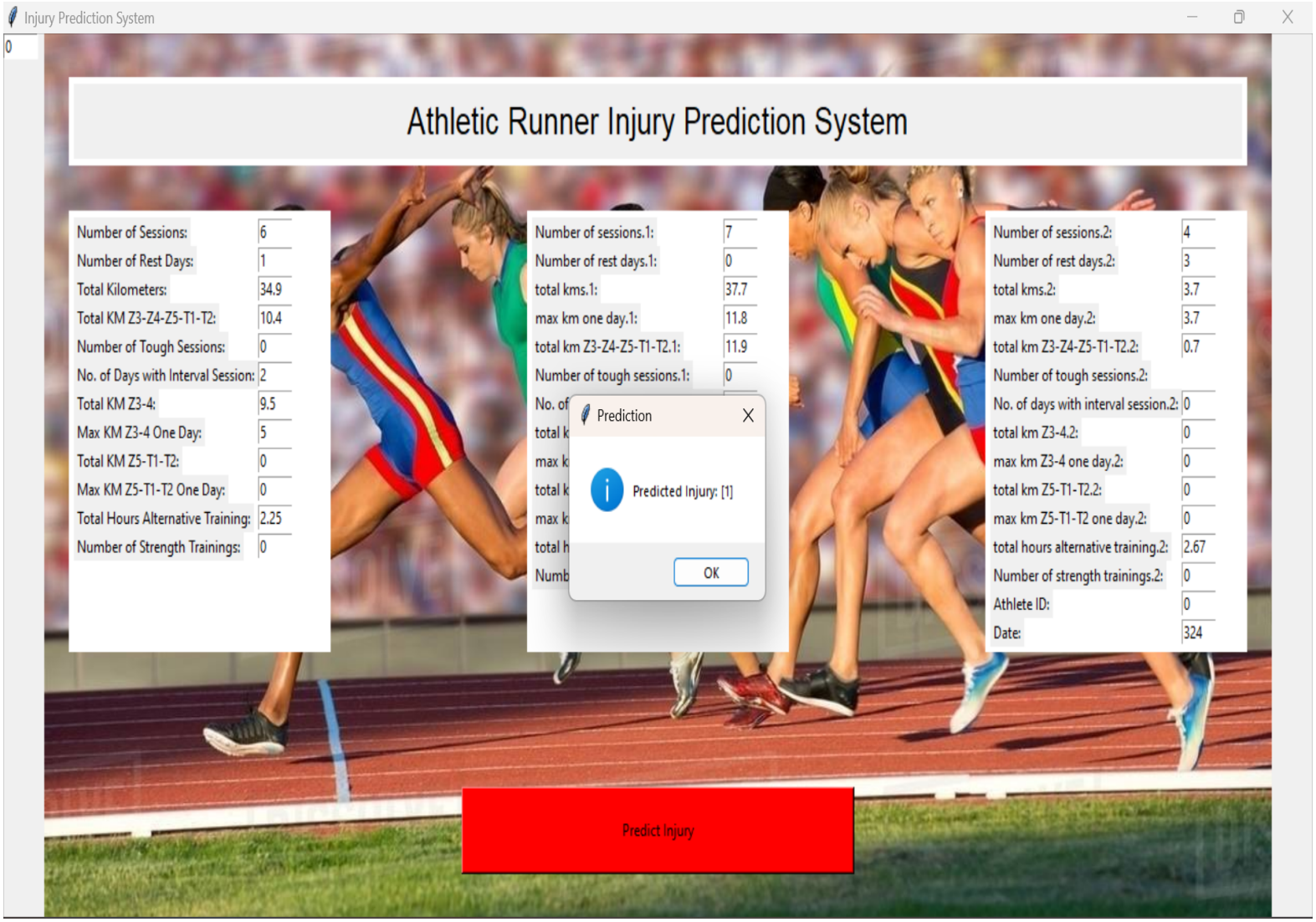
Results and Discussion

DATASET	SVM			Bagging			XgBooster		
	Accuracy	Precision	F1 Score	Accuracy	Precision	F1 Score	Accuracy	Precision	F1 Score
575*42224	**0.94**	0.03	0.05	**0.98**	0	0	**0.98**	1	0.01
575*575	0.58	0.57	0.6	0.57	0.57	0.6	0.68	0.65	0.71
1150*1150	0.65	0.63	0.67	0.78	0.74	0.8	0.71	0.67	0.74
1725*1725	0.7	0.68	0.71	**0.88**	**0.83**	**0.89**	0.72	0.68	0.75
2300*2300	0.7	0.66	0.73	**0.92**	**0.88**	**0.92**	0.73	0.69	0.75
2875*2875	0.73	0.69	0.75	**0.94**	**0.91**	**0.95**	0.71	0.68	0.74

Results and Discussion

DATASET	Decision Tree			MLP			RNN			LSTM		
	Accuracy	Precision	F1 Score	Accuracy	Precision	F1 Score	Accuracy	Precision	F1 Score	Accuracy	Precision	F1 Score
575*42224	**0.97**	0.03	0.04	**0.87**	0.02	0.04	**0.94**	0.01	0.03	**0.95**	0.03	0.05
575*575	0.58	0.53	0.58	0.58	0.57	0.45	0.58	0.55	0.55	0.5	0.4	0.6
1150*1150	0.79	0.76	**0.81**	0.51	0.59	0.3	0.65	0.6	0.73	0.64	0.62	0.7
1725*1725	**0.85**	**0.8**	**0.86**	0.63	0.62	0.66	0.6	0.56	0.72	0.62	0.59	0.71
2300*2300	**0.89**	**0.83**	**0.9**	0.54	0.73	0.28	0.57	0.66	0.42	0.61	0.6	0.65
2875*2875	**0.92**	**0.86**	**0.92**	0.68	0.62	0.74	0.6	0.62	0.57	0.61	0.58	0.69

Results and Discussion



Report Structure

- Introduction
- Related Work
- Dataset
- Proposed Method
- Results and Discussion
- Conclusion

Key Findings

- Data set is biased for non-injured cases.
- Machine Learning models performed better than deep learning models.

References

- Bullock, Garrett S., et al. "Just how confident can we be in predicting sports injuries? A systematic review of the methodological conduct and performance of existing musculoskeletal injury prediction models in sport." Sports medicine 52.10 (2022): 2469-2482.
- Kathan, Alexander, et al. "Investigating Individual-and Group-Level Model Adaptation for Self-Reported Runner Exertion Prediction from Biomechanics." 2022 E-Health and Bioengineering Conference (EHB). IEEE, 2022.
- Huang, Chen, and Lei Jiang. "Data monitoring and sports injury prediction model based on embedded system and machine learning algorithm." Microprocessors and Microsystems 81 (2021): 103654.

References

- Van Eetvelde, Hans, et al. "Machine learning methods in sport injury prediction and prevention: a systematic review." Journal of experimental orthopaedics 8 (2021): 1-15.
- Nejković, Valentina, Maša Radenković, and Nenad Petrović. "Ultramarathon result and injury prediction using PyTorch." 2021 15th International Conference on Advanced Technologies, Systems and Services in Telecommunications (TELSIKS). IEEE, 2021.
- Bai Y, Guan Y, Ng W F. Fatigue assessment using ECG and actigraphy sensors[C]//Proceedings of the 2020 International Symposium on Wearable Computers. 2020: 12-16.

Thank you