

Examining the impact of NBA players' injuries on their individual basketball performance through data science

Milan Sandor^a, Tibor Tajti^b

^aInstitute:

Eszterhazy Karoly Catholic University
sandor.szoda.milan@gmail.com

^bInstitute:

Eszterhazy Karoly Catholic University
tajti.tibor@uni-eszterhazy.hu

Abstract

The aim of this research is to use machine learning techniques to project the consequences of basketball injuries on player statistics. In particular, this paper focuses on injuries to key players and their impact on player performance. We collected data on player statistics, injury information, from publicly available sources. The machine learning models will be trained on these data to predict the effect of injuries on player statistics such as points per game, rebounds per game, and assist per game. Domain experts, managers, technical staff, owners and players pay attention to advanced analytics with serious consideration for key decisions based on data, advanced metrics, Artificial Intelligence (AI) and technology.[6] Additionally, we analyze the impact of injuries on team performance measures such as point differential. This research also investigates the relationship between the severity and type of injuries. Little is known about the incidence of lower extremity bone stress injuries and their impact on return to play and performance in these athletes.[2, 3, 5] We explore the use of various machine learning algorithms, such as decision trees, random forests, and neural networks, to model this relationship. Additionally, we examine the role of player age, experience, and position in predicting the consequences of injuries. Another important factor after injury player impact that players could get injured again. A player who has had a serious injury in the past

is thought to be more likely to have another ailment in the future than a player who has not had an injury. This is referred to as "injury recurrence" or "injury proneness" in the medical literature.[1]

The results of this research provides valuable information for coaches, trainers, and team managers to make informed decisions about injuries different effect on players' careers. Additionally, this research can contribute to the understanding of the relationship between injuries' and player performance in the sport of basketball, and also to get a clearer picture of the extent of the influence of the type of injuries suffered by players, based on their statistical data, how much of an impact a minor/major injury had on their performance in some cases and how much of an impact it could have in similar cases. It also provides a framework for similar studies in other sports.

The present study aims to investigate the relationship between injuries, player statistics, and team performance in a sports setting. To achieve this goal, a combination of descriptive statistics and visualization techniques are utilized to identify patterns and trends in the data. Data Mining is the huge domain to study which is all about extracting patterns, classifying huge and uncertain data, where data is of the heterogeneous forms like text, audio, video, etc.[7] The most significant features are selected from the data using various feature selection methods, including correlation analysis, feature importance, and recursive feature elimination.

Subsequently, a range of machine learning algorithms are applied to the data, including decision trees, random forests, neural networks, and logistic regression. The performance of the selected models is evaluated using a set of standard evaluation metrics, such as accuracy, precision, recall, F1-score, and ROC curve. The results are statistically analyzed to support the conclusions of the study.

In addition to traditional machine learning algorithms, the study also employs T-SNE and SOM algorithms to visualize high-dimensional data points in a 2D plane. T-SNE is particularly useful for data with many dimensions, as it can help reveal the underlying structure of the data, while SOMs are well suited for data with complex, non-linear correlations and varying value ranges.

Machine learning (ML) and Data mining (DM) have been applied in various fields. As a result of this movement, Sport Analytics, a field where ML methods and its implementations are used to gain useful insights from sport data, has been emerging as one of the favourable areas for both business and academic research.[4] The findings of this study provide valuable insights into the relationship between injuries, player statistics, and team performance, and have important implications for sports organizations, coaches, and athletes. The results of this study contribute to a growing body of knowledge in this field and have the potential to inform future research in this area.

References

- [1] C. BITCHELL, J. VARLEY-CAMPBELL, G. ROBINSON: *Recurrent and Subsequent Injuries in Professional and Elite Sport: a Systematic Review*, en, *Sports Med - Open* 6 (2020), p. 58, DOI: [10.1186/s40798-020-00286-3](https://doi.org/10.1186/s40798-020-00286-3).
- [2] T. HUYGHE, P. E. ALCARAZ, J. CALLEJA-GONZALEZ, S. P. BIRD: *The underpinning factors of NBA game-play performance*, en, a systematic review (2001–2020)., 2022, DOI: [10.1080/00913847.2021.1896957](https://doi.org/10.1080/00913847.2021.1896957).
- [3] M. KHAN, K. MADDEN, M. BURRUS: *Epidemiology and Impact on Performance of Lower Extremity Stress Injuries in Professional Basketball Players*, en, *Sports Health* (2018), DOI: [10.1177/1941738117738988](https://doi.org/10.1177/1941738117738988).
- [4] N. H. NGUYEN, D. T. A. NGUYEN, B. MA, J. HU: *The application of machine learning and deep learning in sport: predicting NBA players' performance and popularity*, en, , *Journal of Information and Telecommunication* 6.2 (2022), pp. 217–235, DOI: [10.1080/24751839.2021.1977066](https://doi.org/10.1080/24751839.2021.1977066).
- [5] C. RUGG, A. KADOOR, B. FEELEY, N. PANDYA: *The effects of playing multiple high school sports on National Basketball Association players' propensity for injury and athletic performance*, en, 2018, DOI: [10.1177/0363546517738736](https://doi.org/10.1177/0363546517738736).
- [6] V. SARLIS, V. CHATZILIAS, C. TJORTJIS, D. MANDALIDIS: *A Data Science approach analysing the Impact of Injuries on Basketball Player and Team Performance*, en, 2021, DOI: [10.1016/j.is.2021.101750](https://doi.org/10.1016/j.is.2021.101750).
- [7] M. SOMVANSHI, P. CHAVAN, S. TAMBADE, S. SHINDE: *A review of machine learning techniques using decision tree and support vector machine*, fr, in: 2016 International Conference on Computing Communication Control and automation (ICCUBEA, Pune, India, 2016, pp. 1–7, DOI: [10.1109/ICCUBEA.2016.7860040](https://doi.org/10.1109/ICCUBEA.2016.7860040).