JumpXClass: Explainable AI for Jump Classification in Trampoline Sports

Lucas Woltmann[1] Katja Ferger[2] Claudio Hartmann[1] Wolfgang Lehner[1]

Abstract: Movement patterns in trampoline gymnastics have become faster and more complex with the increase in the athletes' capabilities. This makes the assessment of jump type, pose, and quality during training or competitions by humans very difficult or even impossible. To counteract this development, data-driven solutions are thought to be a solution to improve training. In recent work, sensor measurements and machine learning is used to predict jumps automatically and give feedback to the athletes and trainers. However, machine learning models, and especially neural networks, are black boxes most of the time. Therefore, the athletes and trainers cannot gain any insights about the jump from the machine learning-based jump classification. To better understand the jump execution during training, we propose JumpXClass: a tool for automatic machine learning-based jump classification with explainable artificial intelligence. Using elements of explainable artificial intelligence can improve the training experience for athletes and trainers. This work will demonstrate a live system capable to classify and explain jumps from trampoline athletes.

Keywords: machine learning; applied AI; explainable AI; sports; trampoline

1 Introduction

Over the years, professional sports have experienced a boost in athletic capabilities. Athletes are able to reach new heights of performance in their respective areas. This has led to the need for better training tools including digital ones to better capture the athletes' performance. As a part of gymnastics, trampoline sport has experienced the same developments. Here, one central point is the capture of jumps via near-body sensors and their automatic classification to support the athletes and trainers via auxiliary digital tools [He11, Ca18, Wo22b]. The idea encompassed by all publications is that athletes can improve their performance and training experience by quantifying the exercise through near-body sensor measurements.

In recent research, Woltmann et al. promote using ML in a feedback system for trampoline athletes to improve training experience [Wo22a]. The work enables athletes and trainers to visualize the sensor measurements and automatically classify their jumps using a deep feed-forward neural network (NN). All shown data can be manipulated interactively. Part of the presented work in this demo is based on this preliminary work.

[1] Technische Universität Dresden, Dresden Database Research Group, Dresden, Germany, firstname.lastname@tu-dresden.de
[2] Justus Liebig University Giessen, Institute of Sport Science, Giessen, Germany, katja.ferger@sport.uni-giessen.de

doi:10.18420/BTW2023-34

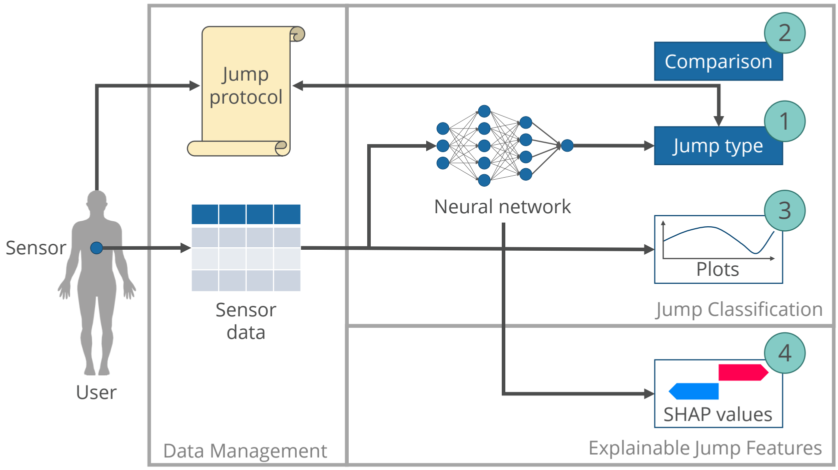


Fig. 1: System overview of JumpXClass. The numbers represent the different use cases.

The high volume of time series data and their black box characteristic are major drawbacks of data-driven ML-based approaches. Decisions, like the jump classification, are generally not traceable by or explainable to the user. This makes it challenging to analyze the interaction between data and model for the subsequent discussions of model quality and application. The scientific field researching this phenomenon and its solutions is called eXplainable Artificial Intelligence (XAI). In this area, ML black boxes are analyzed to conclude about the inner decision-making. For NNs, feature influences are one part of this analysis [LL17, KL17]. These influences quantify the input of an NN according to its influence on the decision. One representative are *shapley additive explanations* (SHAP) values [LL17]. SHAP values are a game theoretic approach modeling the exclusion of features and, therefore, their influence on the model's output. With SHAP, any ML model input feature is assigned a numerical value, representing the absolute influence on the decision. For the jump classification NN, these values represent the probability that the input feature adds to or subtracts from all possible classes.

In this demonstrator, we combine the approaches of a feedback system, automated ML-based classification, and XAI to build an ML-driven exercise assessment tool for trampoline athletes and trainers. During the live demonstration, we will show that combining these concepts brings several advantages to athletes and trainers by incorporating digital technologies into the training.

2 Demonstrator Features

In this section, we present the features and modules of our demonstrator JumpXClass. Three core concepts are implemented in three equivalent modules called *Data Management* to handle user data and meta data, *Jump Classification* to make the ML components accessible, and *Explainable Jump Features* to include the XAI parts. The three general modules of the demonstrator are presented in Figure 1 indicated by gray frames. Each module will add up on the following and improve the feedback given to the athletes and trainers. The four use cases, indicated by numbers, are detailed in Section 3.

2.1 Data Management

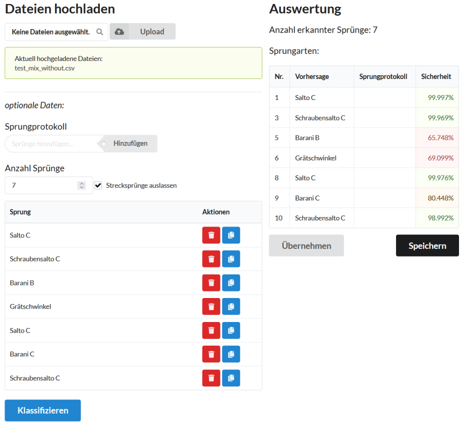
The tool works on the sensor data collected by an accelerometer and a gyroscopic sensor on the back of the athlete. After exporting the data from the sensor, the athletes upload their exercise data, usually containing ten consecutive jumps, into the tool. These are usually CSV files exported from the sensors with a measurement resolution of 500 Hz. Additionally, athletes need to document their exercise in a jump protocol manually. Our tool allows for the creation of such a protocol. Therefore, the athletes can store both their data and meta data for further use. This is important for the reproducibility of a training session or competition and the self-assessment of the athletes. The combined sensor data and jump protocols can also be used as labeled training data within the tool.

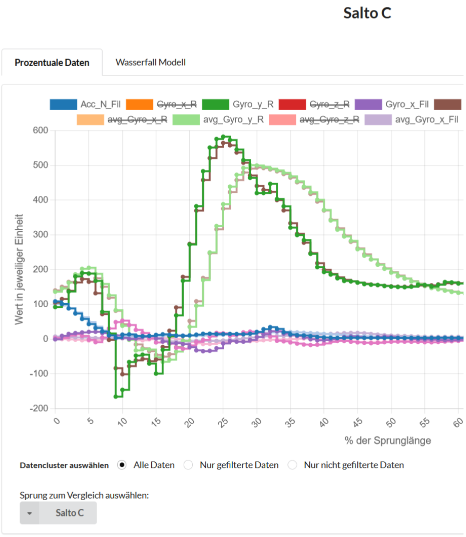
2.2 Jump Classification

Another aspect of this work is the direct use of a model that takes the sensor data and classifies all jumps according to their jump type. There are 148 jump types that our model needs to distinguish. The model is trained and tested on data containing around 2,500 jumps. The data was labeled by athletes (and their trainers) competing on a national level.

The classification gives direct feedback to the athletes about their high-level performance, i.e., if they performed the jump correctly. This can be useful for training by giving the athletes hints on the jumps they need to improve their performance. Another capability is the direct checking of the jump protocol. Mislabeled or mixed-up entries in the protocol can be spotted more quickly if the NN contradicts a protocol entry.

All sensor channels are visualized as time series plots, as shown in Figure 2b. Athletes and trainers can analyze individual jumps from the jump data in more detail. The sensor data contains fine-grained acceleration and orientation information. This enables the analysis of wrong movements during individual jumps. As an additional feature, a comparison to previous jumps or an averaged jump for a jump type is possible via an inlay plot as presented by the lines with a lighter color in Figure 2b. This provides a comparative movement analysis





(a) Use Cases (1) and (2): The jump protocol (left) and the automatic jump classification (right).

(b) Use Case (3): The time series plot for a specific jump including a reference jump (lighter colors).

Fig. 2: Example screens for three use cases.

to spot improvements or errors. Specifically, trainers can have a more detailed look at the jump performance and give quantitative and qualitative feedback to the athlete.

2.3 Explainable Jump Features

The jump classification gives a high-level assessment of a jump but no detailed feedback about the ML model's decision. With XAI, the ML model can give athletes and trainers feedback. To achieve this, we analyze the jump classification NN with SHAP and report the SHAP values to the user via a waterfall plot. With the SHAP influences for each gyroscopic measurement, the athlete can see what movement patterns define a specific jump type according to the NN. For the first time in trampoline sports, XAI helps to identify the part of a jump most influential for good execution. This opens up new possibilities to support training and enhance athletes' performances with ML and XAI. Additionally, the trainers can identify the measurements leading the NN to a wrong classification and give feedback to the athlete about a changed movement pattern for the jump. The athletes can get manifold feedback from the ML and XAI components to improve their training.

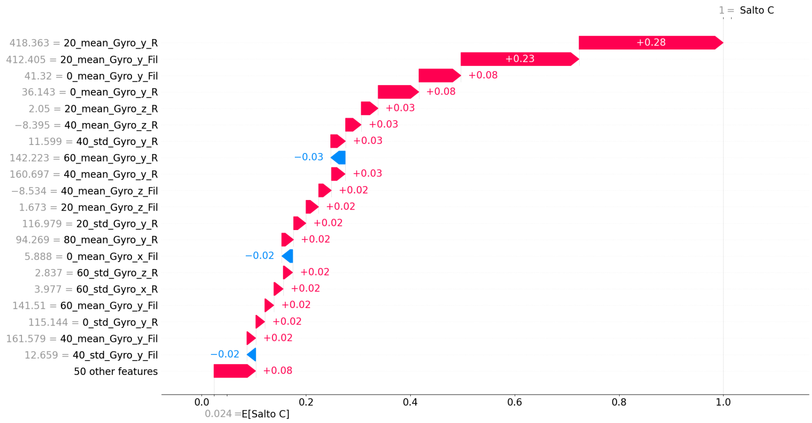


Fig. 3: Use Case (4): The SHAP waterfall plot for a specific jump.

3 Demo Description

In the live demo session, the users will see four use cases: (1) the competition mode, (2) the jump protocol, (3) the measurement plots, and (4) the XAI jump features. The cases are depicted and highlighted according to their number in Figure 1. Every use case has a dedicated target group. Whereas use cases (2), (3), and (4) are meant for athletes and trainers, use case (1) is aimed at competition judges. In (1), only the NN's jump classification from an exercise data set is shown. Therefore, the judges can assess the exercise's difficulty score.

In use case (2), the athletes upload their data for an exercise. The gyroscopic sensor produces time series data sets containing all required measurement channels. Additionally, the athletes can add their jump protocols and store them as meta data for the sensor data. Use case (1) is a subpart of use case (2) since it uses the same sensor data. Both use cases (1) and (2) are presented in Figure 2a. After uploading the data and adding the protocol, the NN automatically analyzes and classifies the jumps into jump types according to [Wo22b]. Every classification is annotated by a percentage detailing the confidence of the NN for this classification. Lower confidence scores usually show that a jump was not cleanly executed. Another point is the direct comparison of the automatic classification and the jump protocol. Color coding shows the users any discrepancies between these two parts and allows for a high-level analysis of the athletes' performance.

Use case (3) plots the uploaded sensor measurement channels as time series plots by clicking on a specific jump in the jump protocol from the first use case, as shown in Figure 2b. The user can (de-)select each channel separately to be plotted or not. Another feature is the comparison view. Here, the athletes or trainers can choose either another jump from the jump protocol or an averaged representative for the specific jump type to be plotted for comparison with the original jump. This gives allows the users to compare their performance to a quasi-standard.

Use case (4) shows the SHAP value waterfall plot for the same selected jump as in use case (3). As detailed in Figure 3, the plot provides feedback regarding the classification and according to which features the NN decided on the jump type. The plot helps scientific experts to assess the jump quality from a technical perspective. The measurements and their influence on the decision can be analyzed and used to find the most important movement of a jump or a deviation of the athlete from a high-quality jump. Additionally, the SHAP values can be used to debug the data and model, either to re-record the jumps or to fine-tune the NN. A re-recording is necessary if a certain feature greatly influences a wrong classification and the corresponding movement is not part of the jump. Here, the athlete would execute the jump again and generate new sensor data for verification. If a feature influences the wrong classification but its corresponding movement is part of the jump, the NN needs to be adjusted through hyperparameter tuning or retraining with new data.

4 Conclusion

In this work, we have shown JumpXClass, a feedback system for trampoline athletes and trainers based on ML models and XAI. The multitude and interplay of features make the tool a valuable asset for athlete quality assessment. It can help to identify errors in jumps or quality metrics of jump types. We argue that using JumpXClass can also highlight the advantages of ML and (X)AI in sports and other applications. For future work, we plan to verify these advantages with actual trampoline gymnasts to obtain feedback for the tool.

Bibliography

[Ca18] Camomilla, Valentina; Bergamini, Elena; Fantozzi, Silvia; Vannozzi, Giuseppe: Trends supporting the in-field use of wearable inertial sensors for sport performance evaluation: A systematic review. Sensors, 18(3):873, 2018.

[He11] Helten, Thomas; Brock, Heike; Müller, Meinard; Seidel, Hans-Peter: Classification of trampoline jumps using inertial sensors. Sports Engineering, 14(2):155–164, 2011.

[KL17] Koh, Pang Wei; Liang, Percy: Understanding black-box predictions via influence functions. In: International conference on machine learning. PMLR, pp. 1885–1894, 2017.

[LL17] Lundberg, Scott M; Lee, Su-In: A unified approach to interpreting model predictions. Advances in neural information processing systems, 30, 2017.

[Wo22a] Woltmann, Lucas; Ferger, Katja; Hartmann, Claudio; Lehner, Wolfgang: Exercise assessment in trampoline sport by automated jump classification. In: Tagungsband zum 14. Symposium der Sektion Sportinformatik und Sporttechnologie der Deutschen Vereinigung für Sportwissenschaft. p. 23, 2022.

[Wo22b] Woltmann, Lucas; Hartmann, Claudio; Lehner, Wolfgang; Rausch, Paul; Ferger, Katja: Sensor-based Jump Detection and Classification with Machine Learning in Trampoline Gymnastics. German Journal of Exercise and Sport Research, 12 2022.