

Machine Learning Approach for Pitch Type Classification Based on Pelvis and Trunk Kinematics Captured with Wearable Sensors

Gomaz, Larisa; Bouwmeester, Celine; van der Graaff, Erik; van Trigt, Bart; Veeger, DirkJan

DOI

[10.3390/s23239373](https://doi.org/10.3390/s23239373)

Publication date

2023

Document Version

Final published version

Published in

Sensors

Citation (APA)

Gomaz, L., Bouwmeester, C., van der Graaff, E., van Trigt, B., & Veeger, D. (2023). Machine Learning Approach for Pitch Type Classification Based on Pelvis and Trunk Kinematics Captured with Wearable Sensors. *Sensors*, 23(23), Article 9373. <https://doi.org/10.3390/s23239373>

Important note

To cite this publication, please use the final published version (if applicable).
Please check the document version above.

Copyright

Other than for strictly personal use, it is not permitted to download, forward or distribute the text or part of it, without the consent of the author(s) and/or copyright holder(s), unless the work is under an open content license such as Creative Commons.

Takedown policy

Please contact us and provide details if you believe this document breaches copyrights.
We will remove access to the work immediately and investigate your claim.

Article

Machine Learning Approach for Pitch Type Classification Based on Pelvis and Trunk Kinematics Captured with Wearable Sensors

Larisa Gomaz ^{1,2,*} , Celine Bouwmeester ², Erik van der Graaff ³, Bart van Trigt ²  and DirkJan Veeger ² 

¹ Delft Institute of Applied Mathematics, Delft University of Technology, 2628 CD Delft, The Netherlands

² BioMechanical Engineering, Faculty of Mechanical, Maritime and Materials Engineering, Delft University of Technology, 2628 CD Delft, The Netherlands; b.vantrigt@tudelft.nl (B.v.T.); h.e.j.veeger@tudelft.nl (D.V.)

³ PITCHPERFECT, 4814 GA Breda, The Netherlands; erik.vandergaaff@cir.nl

* Correspondence: l.gomaz@tudelft.nl

Abstract: The large stream of data from wearable devices integrated with sports routines has changed the traditional approach to athletes' training and performance monitoring. However, one of the challenges of data-driven training is to provide actionable insights tailored to individual training optimization. In baseball, the pitching mechanics and pitch type play an essential role in pitchers' performance and injury risk management. The optimal manipulation of kinematic and temporal parameters within the kinetic chain can improve the pitcher's chances of success and discourage the batter's anticipation of a particular pitch type. Therefore, the aim of this study was to provide a machine learning approach to pitch type classification based on pelvis and trunk peak angular velocity and their separation time recorded using wearable sensors (PITCHPERFECT). The Naive Bayes algorithm showed the best performance in the binary classification task and so did Random Forest in the multiclass classification task. The accuracy of Fastball classification was 71%, whilst the accuracy of the classification of three different pitch types was 61.3%. The outcomes of this study demonstrated the potential for the utilization of wearables in baseball pitching. The automatic detection of pitch types based on pelvis and trunk kinematics may provide actionable insight into pitching performance during training for pitchers of various levels of play.

Keywords: baseball; pitching; wearables; classification; pitch types



Citation: Gomaz, L.; Bouwmeester, C.; van der Graaff, E.; van Trigt, B.; Veeger, D. Machine Learning Approach for Pitch Type Classification Based on Pelvis and Trunk Kinematics Captured with Wearable Sensors. *Sensors* **2023**, *23*, 9373. <https://doi.org/10.3390/s23239373>

Academic Editor: Arnold Baca

Received: 27 September 2023

Revised: 10 November 2023

Accepted: 15 November 2023

Published: 23 November 2023



Copyright: © 2023 by the authors. Licensee MDPI, Basel, Switzerland. This article is an open access article distributed under the terms and conditions of the Creative Commons Attribution (CC BY) license (<https://creativecommons.org/licenses/by/4.0/>).

1. Introduction

Data-driven decision-making is establishing itself in training and high-level sports performance. Data made available through game statistics and technology integrated with training routines serve as the input for big data analytics in sports. Data analysis started in many sports disciplines with some form of video analysis. Currently, a variety of different metrics can be extracted and analyzed not only from videos, but also sensors integrated into sleeves, straps, watches, rings, and smart fabrics. For instance, in baseball, for over 100 years, the difference between a slider and a curveball was defined based on previous experience. Following the technological advancements in pitch tracking, the concept of pitch types is quantified and explained by the speed, spin rate, and spin axis of the ball. Information on the ball (Rapsodo), the bat (Blast), and body movement (PITCHPERFECT) has become widely accessible, creating a new flow of data, which are valuable for performance assessment and pitchers' overall success.

The advancements in wearable technology are changing the traditional approach to athlete training and performance monitoring. Wearables enable measurements in a wide range of settings during training and matches. This removes any practical limitation compared to a lab and offers unlimited athlete availability, which results in high numbers of recorded repetitions. While biomechanical measurements in the lab as well as coaching sessions during training are often limited to one athlete at a time, the utilization of wearables

ensures that every pitch thrown by the pitchers is recorded, even the ones during warm-up sessions. The use and collection of data from wearables can be performed by any motivated team that might lack the resources available to professional sports teams, and this enables coaches to retrospectively provide feedback to every pitcher. Such performance tracking in terms of pitch counts enables players to pitch without fatigue, directly adhering to the pitch count limit regulated by the federations in order to limit the workload and prevent shoulder and elbow injuries [1].

Next to the pitch count, the pitching mechanics and pitch type are considered the main factors in pitching training, which are relevant not only for pitchers' performance, but also for the prevention of injuries [2–5]. As the pitcher's response to a given training stimulus is highly individualized [6], continuous and prospective individual monitoring is crucial in managing the effect of the intense training and competition schedule on the pitcher's performance and health. The use of wearable sensors may provide the opportunity to achieve this.

Information extracted from wearables creates the opportunity to understand the body mechanics of each pitcher on an individual level. Detailed pitch-to-pitch information can help the pitcher learn safe and efficient pitch mechanics. In general, pitching mechanics follow the kinetic chain principle in which the pelvis and trunk serve as a link in the transfer of the momentum generated by the lower extremities to the upper extremities. Efficient proximal-to-distal timing between the pelvis and trunk allows momentum transfer to the ball, resulting in increased throwing velocity [7–9]. On the contrary, poor pitching mechanics in combination with the repetitive mechanical strain of throwing through a high pitch count can negatively affect pitching performance and, at the same time, put the pitcher at risk of shoulder and elbow injuries [1,3–5].

To translate training success into game success, pitchers need to translate their movement skills into a variation of pitch trajectories. A successful pitcher alters the velocity and trajectory of the ball to keep the batters off balance and discourage their anticipation of a particular pitch type. To obtain a variation of ball trajectory, in theory, the pitcher manipulates the grip on the ball at the release point, which results in different rotations of the ball out of the hand of the pitcher. The particular seams of a baseball lead to air pressure variations around the ball, which creates the bending, curving, or sliding motions of the pitch. It should be noted though that multiple studies have reported differences in the pelvis and trunk kinematics between pitch types [3,10–13]. From a strategic point of view, a pitcher may want to achieve similar kinematics among all pitch types to make pitch identification difficult for the batters [11]. If that were the case, it would be unlikely that the pitch type could be distinguished from the body mechanics alone. However, the aforementioned studies acquired their data in a lab setting with highly trained individuals. It can be expected that, at lower levels of play, the movement variation within the individual is even higher.

Except the skill difference, there are obvious differences in financial resources and staff availability as well. Although it is common in youth baseball that a volunteer manually counts the amount of pitches, the tracking of the pitch types is very limited, and in particular, off-speed pitches lead to wildly inaccurate manual classifications given the skill level of the person performing the tagging. Therefore, the automatic detection of pitch types might be extremely beneficial, especially for baseball players who cannot afford expensive camera systems and rely on the manual tracking of pitch types. In this context, it should also be noted that off-speed pitches are associated with an increased risk of shoulder and elbow injuries in youth baseball pitchers. In combination with the increased number of pitches per game and the full baseball calendars, pitchers are at risk of not only acute problems, but also overuse injuries in the later stages of their careers [1].

Translating collected wearables data into actionable insights may bridge the gap between scientific knowledge from biomechanical studies and daily practice. We provide a machine learning approach to the utilization of wearables data through pitch type classification based on the pelvis and trunk peak angular velocity and their separation time recorded using body-worn motion sensors. Machine learning methods showed promising results in pitch type classification investigated in similar contexts [14–20]. Opposed to predicting

the next pitch thrown based on the information available prior to that pitch [14–16], our approach relies on inclusion of post-delivery features to detect which pitch was thrown purely based on pitching mechanics. Having pitch type readily available on every pitch, in combination with kinematic data, might help us provide insight into pitching technique to baseball pitchers of various levels. On top of that, overview of such performance metrics can be presented to the athletes in real time, enabling players to track their progress throughout the whole season and empowering them to shape the training accordingly.

To the best of our knowledge, this is the first study investigating baseball pitch type detection based on pelvis and trunk kinematics during pitching and, moreover, based on such data obtained from wearables. This approach allows for workload monitoring, which is important for maintaining safe and efficient pitching performance during the full course of the season. Therefore, this study aims to establish the methodology for pitch type classification based on biomechanical input from wearables by comparing performance of the various classification algorithms.

2. Materials and Methods

2.1. Participants

Out of 24 pitchers initially participating in the measurements, 19 pitchers were included in this study (age 18.5 ± 3.7 years, height 178.3 ± 11.1 m, weight 71.9 ± 18.3 kg, experience 7.3 ± 3.7 years). The participants were members of the elite youth academies of the Royal Dutch Baseball and Softball Federation (KNBSB). The included pitchers were pain- and injury-free during the course of the measurements. This research was conducted in accordance with the Declaration of Helsinki, and the Ethics Committee of the Delft University of Technology approved the measurement protocol (approval no. ETC_TUdelft_1394). Informed consent was signed by the participants or the general manager of the respective baseball academy.

2.2. Data Collection and Data Pre-Processing

The data were collected during the pitchers' regular training at the training facilities of the affiliated baseball academy. To maintain pitching-specific routines, warm-up and pitch count were not standardised. After performing their standard warm-up, the pitchers were instructed to throw a selection of pitch types they usually throw during the game, containing a minimum of three different pitch types. The pitchers followed their own training routine in accordance with the training program set by their pitching coach. The bullpen session consisted of a minimum of 20 pitches from mound toward a catcher at the official distance of 18.45 or 16.45 m, depending on the pitcher's age.

The pitching motion was recorded using the PITCHPERFECT system (PITCHPERFECT, Breda, The Netherlands) consisting of two synchronised 3-DOF IMUs (Gyroscope ± 2000 ($^{\circ}$ /s)) showed on Figure 1. Sensors were taped with Leukoplast FixoMull® stretch (BSN Medical GmbH, Hamburg, Germany) on processus Xiphoides on the chest and in the middle of the left and right posterior superior iliac spine on the lower back of the pitcher before starting the bullpen (Figure 2). Pitch types were manually coded by experienced off-field staff members based on the visual inspection, hand signal and pitcher–catcher agreement prior to each throw. The ball velocity (mph) was measured from behind the pitcher with a Pocket radar Ball coach, Model PR1000-BC (Pocket Radar Inc., Santa Rosa, CA, USA). The accuracy of the pitch was noted, distinguishing only between a wild pitch or not, wherein a wild pitch was noted if the catcher was unable to catch the ball with reasonable effort.

The outcome of the PITCHPERFECT system consists of the pelvis and trunk peak angular velocity and the separation time between them. Pre-processing of the raw sensor signal and computing Euclidean norms from the raw data were conducted by the algorithm developed by the manufacturer (PITCHPERFECT, The Netherlands). Details of the algorithm are property of the manufacturer.

In this study, we used a database created by PITCHPERFECT that characterizes each pitch with three features used directly from the system (Table 1). Data were pre-

processed and analyzed using the R programming language (version 4.3.1). Data of five players were excluded from the analyses because their peak angular velocity was below the threshold of 400 ($^{\circ}/s$) of the PITCHPERFECT system. Individual pitches were included based on three inclusion criteria: (1) the pitch type is a Fastball (FB), Curveball (CU) or Change-up (CH), as they were the most occurring pitch types among the included pitchers; (2) the thrown ball was not a wild pitch; and (3) all three kinematic parameters (*Pelvis*, *Trunk*, *Separation*) were recorded (i.e., sensor clipping did not occur). All continuous features were scaled and centered.



Figure 1. Pitch Perfect sensor system for measuring pelvis and trunk kinematics and separation time between them.

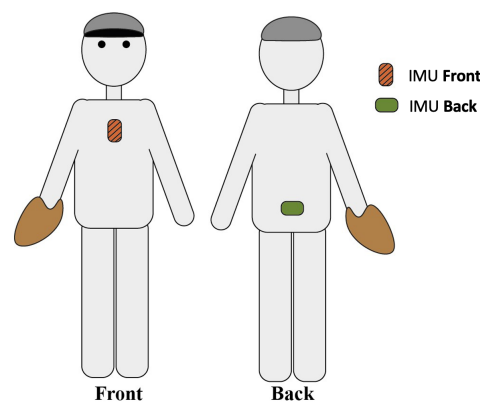


Figure 2. Placement of the sensors. Figure adopted from the study of Gomaz et al. [21].

Table 1. Included features for pitch type classification.

Features	Definitions
Pelvis ($^{\circ}/s$)	Pelvis peak angular velocity available directly from PITCHPERFECT.
Trunk ($^{\circ}/s$)	Trunk peak angular velocity available directly from PITCHPERFECT.
Separation (ms)	The timing between pelvis and trunk peak angular velocity, available directly from PITCHPERFECT.

2.3. Data Analysis

The automatic detection of pitch types from sensor data is a classification problem. The goal is to learn a mapping from inputs x to outputs y , where $y \in \{1, \dots, C\}$, with C being the number of classes. Inputs x are the features (Table 1) and outputs y are pitch types, where C denotes number of different pitch types.

This study utilized classifiers integrated in the *caret* package [22] including K-Nearest Neighbors (KNN), Naive Bayes (NB), Random Forest (RF) and Support Vector Machine (SVM). We investigated the performance of the classifiers in both binary and multiclass classification, including additional Logistic Regression (LOGREG) for binary and Multinomial Logistic Regression (MNOM) for the multiclass classification task.

Binary classification is a classification task that has two class labels. In this study, it is used to detect whether the pitch was Fastball or not by classifying recorded pitches in one of the two classes—FB and Other (Figure 3 (left)). Among the recorded pitches, 48.7% were originally labelled as FB and 51.3% as Other.

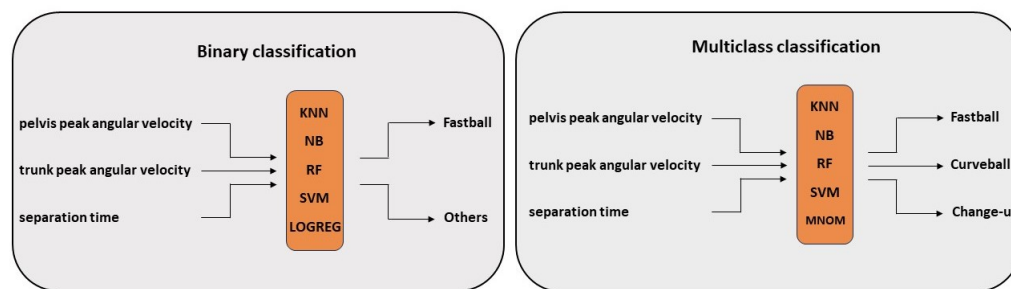


Figure 3. The baseball pitch type classification approaches. **(Left)** The binary classification approach classifies pitch types into two categories—Fastball and Others—based on input from wearables (pelvis and trunk peak angular velocities and separation time). **(Right)** The multiclass classification approach classifies pitch types into three categories—Fastball, Curveball and Change-up—based on input from wearables (pelvis and trunk peak angular velocities and separation time). Both approaches used four classifiers—K-Nearest Neighbors (KNN), Naive Bayes (NB), Random Forest (RF) and Support Vector Machine (SVM)—to assess their classification performance, including additional logistic regression (LOGREG) for binary and multinomial logistic regression (MNOM) for the multiclass classification task.

Multiclass classification refers to classification tasks that have more than two class labels. Unlike binary classification, it classifies non-fastball pitches in different classes and therefore detects whether the pitch was Fastball (FB), Curveball (CU) or Change-up (CH) (Figure 3 (right)). Among the recorded pitches, 48.7% were originally labelled as FB, 26.4% as CH and 24.9% as CU. Due to variations in the number and type of off-speed pitches (CU and CH) among pitchers, the collected data show unequal distribution between classes. Such disparity in the frequencies of the observed classes can have a negative impact on model fitting. A possible solution is to subsample the training data in such a way that mitigates the issue (e.g., under- and oversampling). Hence, to address this issue, the minority classes (CU and CH) were up-sampled so that each class was of equal size.

We set up our training and testing cases following the 80% (training) and 20% (testing) split. To achieve a fair understanding of the generalizability of the classifiers, in the designated training set, Leave-One-Group-Out Cross-Validation (LOGO-CV) was carried out. LOGO-CV is a specific type of k -fold cross-validation that utilizes data from each individual pitcher as a test set. The number of folds therefore equals the number of pitchers. For every fold, the model is trained on data from $J - 1$ pitchers and tested on the data from the one left-out pitcher.

The performance of the classifiers is evaluated by four evaluation criteria—Accuracy (1), Sensitivity (2), Precision (3) and F1-score (4)—which can be calculated from the confusion matrix. The confusion matrix provides a summary of the prediction results of a classification algorithm. In the matrix, the numbers of correct and incorrect predictions are summarised with count values and broken down by each class. The output True Positive (TP) represents the number of positives classified correctly, whereas True Negative (TN) represents the number of correctly classified negatives. False Positive (FP) shows the num-

ber of negatives that are classified as positives, whereas False Negative (FN) indicates the number of positives classified as negatives.

$$Accuracy = \frac{TP + TN}{Total\ sample} \quad (1)$$

$$Sensitivity = \frac{TP}{TP + FN} \quad (2)$$

$$Precision = \frac{TP}{TP + FP} \quad (3)$$

$$F1 = \frac{2TP}{2TP + FP + FN} \quad (4)$$

The hyper-parameters were tuned using grid search, a default method for optimizing tuning parameters in the caret package [22]. Feature selection was performed using correlation analysis. Since the correlation between the features was low, the models were trained and tested using all variables derived from the PITCHPERFECT system (Table 1).

3. Results

A total of 353 pitches thrown by 19 pitchers met the inclusion criteria and were included in the study. Descriptive statistics for binary and multiclass classification is presented in Table 2 and Table 3, respectively. A total of 284 pitches were used for training the models and 69 pitches were used for their testing.

Table 2. Descriptive statistics for binary classification.

Features	FB (<i>n</i> = 172)		Other (<i>n</i> = 181)	
	Mean	SD	Mean	SD
Pelvis (°/s)	737	138	695	120
Trunk (°/s)	799	228	827	262
Separation (s)	0.03	0.13	0.06	0.13
Speed (m/s)	33.1	3.82	28.6	3.81

Table 3. Descriptive statistics for multiclass classification.

Features	CH (<i>n</i> = 93)		CU (<i>n</i> = 88)		FB (<i>n</i> = 172)	
	Mean	SD	Mean	SD	Mean	SD
Pelvis (°/s)	708	129	681	109	737	138
Trunk (°/s)	831	277	823	247	799	228
Separation (s)	0.06	0.15	0.06	0.11	0.03	0.13
Speed (m/s)	29.9	3.72	27.2	3.40	33.1	3.82

3.1. Binary Classification

The performance of the K-Nearest Neighbors, Naive Bayes, Random Forest, Support Vector Machine and Logistic Regression algorithms in the binary classification problem was evaluated using four performance metrics (1)–(4). Among the trained classifiers, the Naive Bayes algorithm performed the best in classifying fastballs among the recorded pitches. The confusion matrix seen in Figure 4 shows the summary of the prediction performance for Naive Bayes (Accuracy = 71.0%, Precision = 71.9%, Sensitivity = 67.6%, F1-score = 69.7%). The accuracy of the NB algorithm was 7.2% higher than for KNN, 1.4% higher than for RF, 5.8% higher than for SVM and 20.3% higher than for LOGREG. The sensitivity of the RF algorithm is 11.8% higher than for KNN, 3% higher than for NB, 5.9% higher than for

SVM and 17.7% higher than for LOGREG. The precision of the NB algorithm was 7.4% higher than for KNN, 3.3% higher than for RF, 7.2% higher than for SVM and 21.9% higher than for LOGREG. The F1-score of the NB algorithm was 8.2% higher than for KNN, 0.1% higher than for RF, 5.0% higher than for SVM and 18.3% higher than for LOGREG. The confusion matrices with corresponding performance metrics of the remaining algorithms are shown in Appendix A.

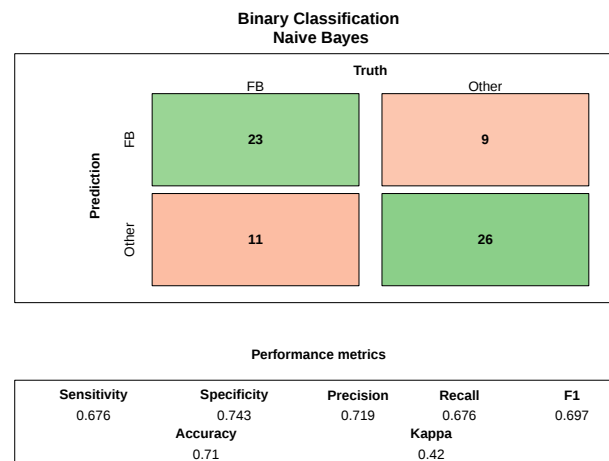


Figure 4. Two-class confusion matrix summarizing the performance of Naive Bayes in classification of fastballs.

3.2. Multiclass Classification

The four metrics are used to evaluate the performance of the K-Nearest Neighbors, Naive Bayes, Random Forest, Support Vector Machine and Multinomial logistic regression algorithms in the multiclass classification problem. Among the trained classifiers, the Random Forest algorithm performed the best in classifying pitches in three different classes of pitch types (FB, CH and CU). The confusion matrix seen in Figure 5 shows the summary of prediction performance for Random Forest. The accuracy of the RF algorithm was at 52.2%, which is 7.2% higher than for KNN, 7.2% higher than for NB, 11.6% higher than for SVM and 8.7% higher than for MNOM. The confusion matrices with corresponding performance metrics of the remaining algorithms are shown in Appendix B. Performance metrics of the Random Forest algorithm are reported in Table 4.

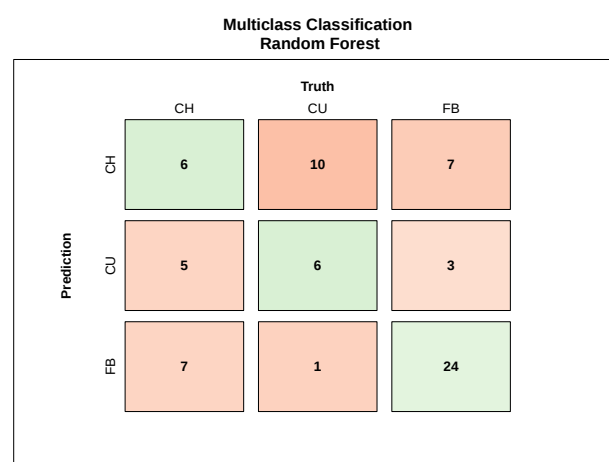


Figure 5. Three-class confusion matrix summarizing the performance of Random Forest by class in classification of baseball pitch types.

Table 4. Performance metrics of multiclass Random Forest in classification of three different pitch types.

Class	Accuracy	Sensitivity	Precision	F1
CH	0.500	0.333	0.261	0.293
CU	0.600	0.353	0.429	0.387
FB	0.739	0.706	0.750	0.727

4. Discussion

The aim of this study was to establish a methodology for pitch type classification based on biomechanical input from wearables. We used pelvis and trunk peak angular velocity and separation time between them as an input and evaluated the performance of five machine learning classifiers in the binary and multiclass classification task. The Naive Bayes algorithm showed the best performance in classifying Fastballs with an accuracy of 71%. Furthermore, in the classification of pitch types as Fastball, Curveball or Change-up, the Random Forest algorithm performed the best with an average accuracy of 61.3% over those three pitch types.

Binary classification was used to detect whether the pitch was Fastball or not. Fastball can be considered a "normal" throw. Fastball is the most common pitch type thrown, specifically among youth pitchers. This has to do with the physical development of youth pitchers where the Fastball pitch is used to learn proper body mechanics and throwing accuracy before learning more demanding off-speed pitches. Therefore, to explore the possibility of pitch type classification based on pitching mechanics, it makes sense to first investigate whether we can detect fastballs. Previous studies that used a binary approach for pitch type prediction focused on predicting whether the next pitch will be Fastball rather than detecting whether Fastball was thrown [15,19]. They used pre-pitch ball data as an input, which resulted in accuracies of 70% [15] and 77.45% [19]. Even though such approach offers benefits for choosing the right strategy, it does not contribute to the pitch tracking as part of the workload monitoring for an individual pitcher.

The multiclass classification task classified recorded pitch types into three categories—Fastball (FB), Change-up (CH) and Curveball (CU). It serves as a base for pitch tracking and detects different pitch types thrown. The Random Forest algorithm performed the best with a 50.0% accuracy in classifying CH, a 60.0% accuracy in classifying CU and a 73.9% accuracy in classifying FB. The performance metrics reported in Table 4 show the performance of the RF classifier for each pitch type versus the rest. Multiclass classification has been a subject of several studies before, focusing on pitch type classification based on pre-pitch ball data. Compared to the accuracy of the Random Forest algorithm revealed in this paper, those studies reported higher predictive accuracies, from 74.5% [20] for the SVM algorithm to 93.63% for the KNN algorithm with Manhattan distance [17,18]. This may be due to the sensitive nature of wearable data and inconsistent pitching mechanics of different pitchers among various pitch types. The feature importance for the Random Forest multiclass classifier revealed that the pitcher's pelvis peak angular velocity is considered most important for the pitch type classification task, whereas the trunk peak angular velocity is considered the least important (Figure A9).

Although we are confident that the proposed methodology could be key to predict pitch type based on biomechanical data from wearables, the reported accuracies leave much to desire. One limitation of this study was that the amount of collected data was low ($n = 353$). The proof of methodology provided in this paper could serve for a study on a larger scale. Additionally, due to the small sample size of individual pitchers, we were not able to perform the classification of pitch types per individuals. The data from the pitchers have a hierarchical structure, suggesting that pitching mechanics [21] as well as pitch kinematics [23] among different throws are more similar for an individual pitcher compared to others. Therefore, it may be sensible to classify pitch types for individual pitchers. Pitch type prediction by pitch count and by pitcher showed improved performance

in the prediction of the next pitch the pitcher will throw based on features available from the previous throws [15,16,18]. Our study would have benefited from longitudinal data collection including kinematic data during the full season. This would allow us to perform classification tasks for different pitch types for individual pitchers. Moreover, matching pitching kinematics with ball speed data may also increase the accuracy of the model.

To the best of our knowledge, this is the first study that uses biomechanical data from wearables to predict pitch types, and thus enriches the available data from an easy-to-use motion sensor system. It is important to clarify that this method is proposed for the classification of the pitch thrown and not the prediction of the next pitch. Pitch prediction uses information available prior the pitch to judge which pitch can be expected. However, pitch type classification uses information available post pitch to determine which pitch type was thrown. Previous studies used post-pitch data from PITCHf/x describing the characteristics of the ball from when it leaves the hand of the pitcher until it crosses the home plate [17,18]. Defining the pitch type from ball flight data is related to the inherent need of redefining pitch types. Traditional pitch type description is not sufficient any longer, with the newly available data in professional pitching. Our methodology aims to expand this knowledge to situations such as youth baseball, where expensive PITCHf/x systems are not prevalent.

The proposed classification method, based on a limited amount of data from youth baseball pitches, shows promising performance in predicting Fastball vs. off-speed pitches. Application of this binary classification method in youth baseball training can create a major advantage for the development of individual players. Since nowadays pitch count is the only variable that is noted, and mostly manually recorded, the automatic tracking of pitch counts, biomechanical data and pitch types can be of great value to coaches and players. Given that youth players are still learning how to throw different pitch types and their susceptibility to injuries is higher when throwing off-speed pitches [1], implementing the proposed methods in baseball practice may provide a wealth of information relevant for both pitchers and coaches in those situations.

Implementing similar technologies for elite athletes' training could benefit from the aforementioned suggestions to improve the accuracy of the multiclass classification model. However, further studies should determine the necessity of such a system since high-level players often have access to other resources that can measure or calculate pitch trajectory. Indirect pitch type prediction may thus not be needed for players at a high level with many resources at their disposal.

5. Conclusions

The accessibility of wearable sensors for performance tracking during both training and games represents a new source of large amounts of data that need powerful algorithms for their analysis, resulting in actionable insights relevant for pitchers' performance and injury risk management. This study established machine learning methods for the detection of the pitch type that was thrown based on pitching mechanics recorded with wearables. The Naive Bayes algorithm showed the best performance in the detection of fastballs, whereas the Random Forest algorithm performed best in the multiclass (FB vs. CH vs. CU) classification task. While these findings demonstrate the potential for the utilisation of wearables in baseball pitching, further development of the classification algorithm, as well as longitudinal data collection, is required. Providing insight into pitch count, pitching mechanics and pitch type enables pitchers to throw safely and efficiently. Through automatic tracking of pitch types, every pitch is counted. Thus, monitoring pitching mechanics and providing an informative feedback to the pitchers may lead to safe and efficient pitching and increase a pitcher's chances of success.

Author Contributions: Conceptualization, C.B., E.v.d.G., B.v.T. and D.V.; methodology, C.B. and E.v.d.G.; software, L.G.; validation, L.G.; formal analysis, L.G. and C.B.; investigation, C.B. and E.v.d.G.; resources, C.B., E.v.d.G. and B.v.T.; data curation, L.G., C.B. and E.v.d.G.; writing—original draft preparation, L.G. and E.v.d.G.; writing—review and editing, L.G., E.v.d.G. and D.V.; visualiza-

tion, L.G.; supervision, D.V.; project administration, D.V.; funding acquisition, D.V. All authors have read and agreed to the published version of the manuscript.

Funding: This research was supported by the NWO Domain Applied and Engineering Sciences (AES) under project number [R/003635]. The NWO-funded projects, named Breaking the High Load—Bad Coordination Multiplier in Overhead Sports Injuries (Project 7) and Data Science for Injury Prevention and Performance Improvement (Project 2), are part of the research program Perspectief CAS and a cooperative effort between the Royal Dutch Baseball and Softball Federation (KNBSB), Royal Dutch Tennis Federation (KNLTB), Vrije Universiteit Amsterdam, Delft University of Technology, Milé Fysiotherapy, PitchPerfect and PLUX.

Institutional Review Board Statement: This research was conducted in accordance with the Declaration of Helsinki and the Ethics Committee of the Delft University of Technology approved the measurement protocol (approval no. ETC_TUdelft_1394).

Informed Consent Statement: Informed consent was obtained from all subjects involved in the study.

Data Availability Statement: The data presented in this study are openly available in 4TU. Research-Data repository at <https://data.4tu.nl/datasets/f86ba220-08a1-4fa0-89a9-d8995790675b> (accessed on 6 November 2023). The code is available at <https://data.4tu.nl/datasets/e339176b-0ecd-48e5-bc7e-9b587c0a8959> (accessed on 9 November 2023).

Acknowledgments: We would like to thank the pitchers and their coaches for participating in the study and for being so hospitable. Special thanks to the staff and players of Twins Oosterhout for letting us test the sensors at their training sessions.

Conflicts of Interest: Author Erik van der Graaff was employed by the company PITCHPERFECT. The remaining authors declare that the research was conducted in the absence of any commercial or financial relationships that could be construed as a potential conflict of interest.

Abbreviations

The following abbreviations are used in this manuscript:

FB	Fastball
CH	Change-up
CU	Curveball
KNN	K-Nearest Neighbors
NB	Naive Bayes
RF	Random Forest
SVM	Support Vector Machine
LOGREG	Logistic Regression
MNOM	Multinomial Logistic Regression

Appendix A. Binary Classification

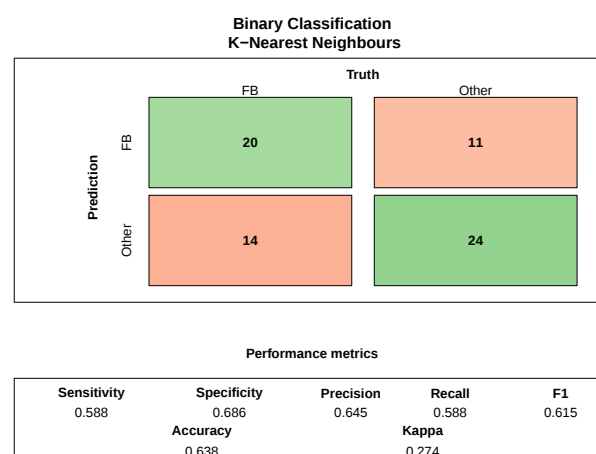


Figure A1. Confusion matrix for binary K-Nearest Neighbors algorithm.

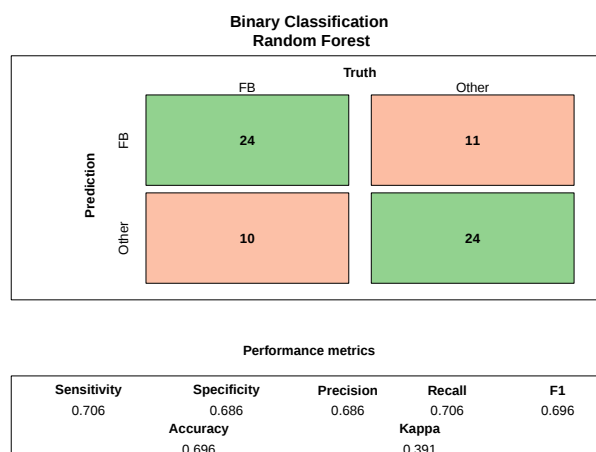


Figure A2. Confusion matrix for binary Random Forest algorithm.

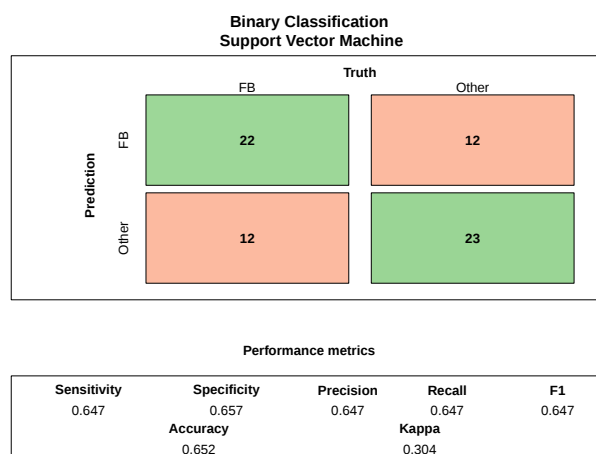


Figure A3. Confusion matrix for binary Support Vector Machine algorithm with radial basis kernel function.

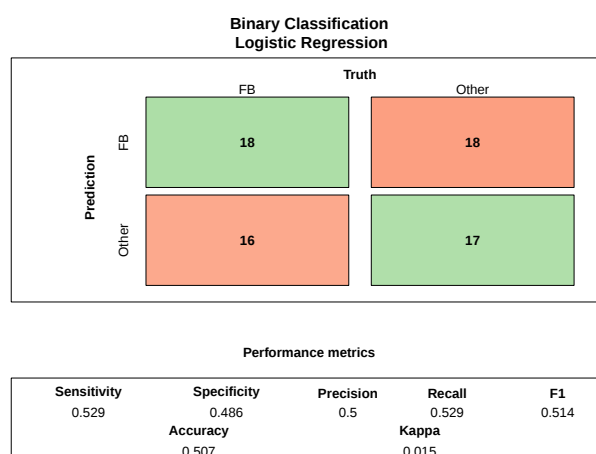


Figure A4. Confusion matrix for binary Logistic Regression algorithm with radial basis kernel function.

Appendix B. Multiclass Classification

Multiclass Classification K-Nearest Neighbours				
Prediction	Truth			
	CH	CU	FB	
	CH	6	5	9
	CU	7	9	9
	FB	5	3	16

Figure A5. Confusion matrix for multiclass K-Nearest Neighbors algorithm.

Multiclass Classification Naive Bayes				
Prediction	Truth			
	CH	CU	FB	
	CH	3	1	6
	CU	6	12	12
	FB	9	4	16

Figure A6. Confusion matrix for multiclass Naive Bayes algorithm.

Multiclass Classification Support Vector Machine				
Prediction	Truth			
	CH	CU	FB	
	CH	6	8	9
	CU	6	6	9
	FB	6	3	16

Figure A7. Confusion matrix for multiclass Support Vector Machine algorithm with radial basis kernel function.

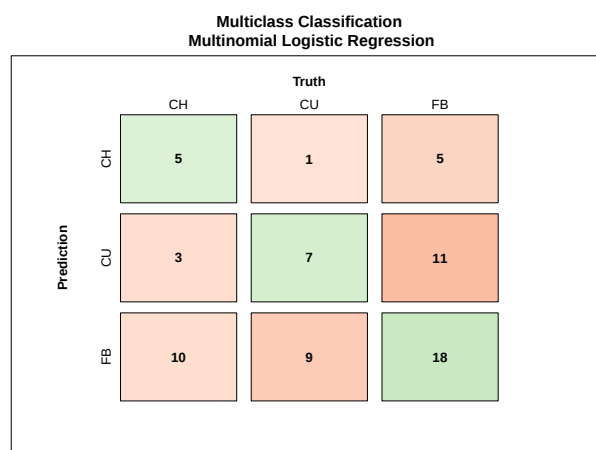


Figure A8. Confusion matrix for multiclass Multinomial Logistic Regression algorithm with radial basis kernel function.

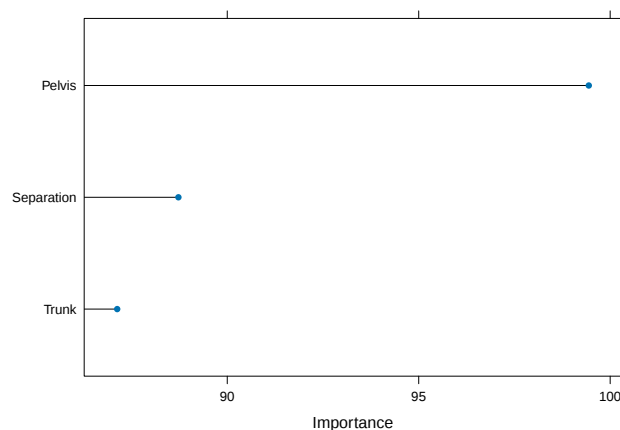


Figure A9. Visual representation of the feature importance for Random Forest multiclass classifier calculated with `varImp` from `caret` package. The horizontal axis should be interpreted as a measure for relative importance of predictive variables. The figure reveals *Pelvis* to be considered as the most important for the multiclass classification task, whereas *Trunk* is considered as the least important.

References

1. Dowling, B.; McNally, M.P.; Chaudhari, A.M.; Oñate, J.A. A review of workload-monitoring considerations for baseball pitchers. *J. Athl. Train.* **2020**, *55*, 911–917. [\[CrossRef\]](#) [\[PubMed\]](#)
2. Lyman, S.; Fleisig, G.S.; Andrews, J.R.; Osinski, E.D. Effect of Pitch Type, Pitch Count, and Pitching Mechanics on Risk of Elbow and Shoulder Pain in Youth Baseball Pitchers. *Am. J. Sport. Med.* **2002**, *30*, 463–468. [\[CrossRef\]](#) [\[PubMed\]](#)
3. Fleisig, G.S.; Kingsley, D.S.; Loftice, J.W.; Dinnen, K.P.; Ranganathan, R.; Dun, S.; Escamilla, R.F.; Andrews, J.R. Kinetic Comparison among the Fastball, Curveball, Change-up, and Slider in Collegiate Baseball Pitchers. *Am. J. Sport. Med.* **2006**, *34*, 423–430. [\[CrossRef\]](#) [\[PubMed\]](#)
4. Fortenbaugh, D.; Fleisig, G.S.; Andrews, J.R. Baseball Pitching Biomechanics in Relation to Injury Risk and Performance. *Sport. Healthc. Multidiscip. Approach* **2009**, *1*, 314–320. [\[CrossRef\]](#) [\[PubMed\]](#)
5. Davis, J.; Limpisvasti, O.; Fluhme, D.; Mohr, K.J.; Yocum, L.A.; ElAttrache, N.S.; Jobe, F.W. The Effect of Pitching Biomechanics on the Upper Extremity in Youth and Adolescent Baseball Pitchers. *Am. J. Sport. Med.* **2009**, *37*, 1484–1491. [\[CrossRef\]](#) [\[PubMed\]](#)
6. Soligard, T.; Schweltnus, M.; Alonso, J.M.; Bahr, R.; Clarsen, B.; Dijkstra, H.P.; Gabbett, T.; Gleeson, M.; Häggglund, M.; Hutchinson, M.R.; et al. How much is too much? (Part 1) International Olympic Committee consensus statement on load in sport and risk of injury. *Br. J. Sport. Med.* **2016**, *50*, 1030–1041. [\[CrossRef\]](#) [\[PubMed\]](#)
7. Aguinaldo, A.L.; Buttermore, J.; Chambers, H. Effects of Upper Trunk Rotation on Shoulder Joint Torque among Baseball Pitchers of Various Levels. *J. Appl. Biomech.* **2007**, *23*, 42–51. [\[CrossRef\]](#) [\[PubMed\]](#)
8. Putnam, C.A. Sequential motions of body segments in striking and throwing skills: Descriptions and explanations. *J. Biomech.* **1993**, *26*, 125–135. [\[CrossRef\]](#) [\[PubMed\]](#)

9. Van Der Graaff, E.; Hoozemans, M.M.; Nijhoff, M.; Davidson, M.; Hoezen, M.; Veeger, D.H. Timing of peak pelvis and thorax rotation velocity in baseball pitching. *J. Phys. Fit. Sport. Med.* **2018**, *7*, 269–277. [\[CrossRef\]](#)
10. Dun, S.; Loftice, J.; Fleisig, G.S.; Kingsley, D.; Andrews, J.R. A Biomechanical Comparison of Youth Baseball Pitches: Is the Curveball Potentially Harmful? *Am. J. Sport. Med.* **2008**, *36*, 686–692. [\[CrossRef\]](#) [\[PubMed\]](#)
11. Escamilla, R.F.; Fleisig, G.S.; Barrentine, S.W.; Zheng, N.; Andrews, J.R. Kinematic Comparisons of Throwing Different Types of Baseball Pitches. *J. Appl. Biomech.* **1998**, *14*, 1–23. [\[CrossRef\]](#)
12. Escamilla, R.F.; Fleisig, G.S.; Groeschner, D.; Akizuki, K. Biomechanical Comparisons Among Fastball, Slider, Curveball, and Changeup Pitch Types and Between Balls and Strikes in Professional Baseball Pitchers. *Am. J. Sport. Med.* **2017**, *45*, 3358–3367. [\[CrossRef\]](#) [\[PubMed\]](#)
13. Fleisig, G.S.; Laughlin, W.A.; Aune, K.T.; Cain, E.L.; Dugas, J.R.; Andrews, J.R. Differences among fastball, curveball, and change-up pitching biomechanics across various levels of baseball. *Sport. Biomech.* **2016**, *15*, 128–138. [\[CrossRef\]](#) [\[PubMed\]](#)
14. Hoang, P. Supervised Learning in Baseball Pitch Prediction and Hepatitis C Diagnosis. Ph.D. Thesis, North Carolina State University, Raleigh, NC, USA, 2015. [\[CrossRef\]](#)
15. Ganeshapillai, G.; Guttag, J.V. Predicting the Next Pitch. In Proceedings of the MIT Sloan Sports Analytics Conference, Boston, MA, USA, 2–3 March 2012.
16. Sidle, G.; Tran, H. Using multi-class classification methods to predict baseball pitch types. *J. Sport. Anal.* **2018**, *4*, 85–93. [\[CrossRef\]](#)
17. Attarian, A.; Danis, G.; Gronsbell, J.; Iervolino, G.; Tran, H. A Comparison of Classification Methods with an Application to Classifying Baseball Pitches. In Proceedings of the International MultiConference of Engineers and Computer Scientists, Hong Kong, 13–15 March 2013.
18. Pane, M.A.; Ventura, S.L.; Steorts, R.C.; Thomas, A.C. Trouble with the Curve: Improving MLB Pitch Classification. *arXiv* **2013**, arXiv:1304.1756.
19. Hamilton, M.; Hoang, P.; Layne, L.; Murray, J.; Padget, D.; Stafford, C.; Tran, H. Applying Machine Learning Techniques to Baseball Pitch Prediction. In Proceedings of the 3rd International Conference on Pattern Recognition Applications and Methods, ESEO, Angers, France, 6–8 March 2014; pp. 520–527. [\[CrossRef\]](#)
20. Bock, J. Pitch Sequence Complexity and Long-Term Pitcher Performance. *Sports* **2015**, *3*, 40–55. [\[CrossRef\]](#)
21. Gomaz, L.; Veeger, D.; Van Der Graaff, E.; Van Trigt, B.; Van Der Meulen, F. Individualised Ball Speed Prediction in Baseball Pitching Based on IMU Data. *Sensors* **2021**, *21*, 7442. [\[CrossRef\]](#) [\[PubMed\]](#)
22. Kuhn, M. Building Predictive Models in R Using the caret Package. *J. Stat. Softw.* **2008**, *28*, 1–26. [\[CrossRef\]](#)
23. Umemura, K.; Yanai, T.; Nagata, Y. Application of VBGMM for pitch type classification: Analysis of TrackMan’s pitch tracking data. *Jpn. J. Stat. Data Sci.* **2021**, *4*, 41–71. [\[CrossRef\]](#)

Disclaimer/Publisher’s Note: The statements, opinions and data contained in all publications are solely those of the individual author(s) and contributor(s) and not of MDPI and/or the editor(s). MDPI and/or the editor(s) disclaim responsibility for any injury to people or property resulting from any ideas, methods, instructions or products referred to in the content.