International Research Journal of Engineering and Technology (IRJET) e-ISSN: 2395-0056
Volume: 07 Issue: 06 | June 2020 www.irjet.net p-ISSN: 2395-0072

Identifying Suitable Sport for Beginners using Data Mining Approach

Perez Bastin

Student, Dept of Dual Degree Computer Applications, Sree Narayana Guru Institute of Science and Technology N. Paravur, Kerala, India

---

Abstract - *Data Mining is a technique which used in various kinds of fields in the industry and they are helpful for collecting information and make use of it in fields. Trainers may use data mining techniques for tactical drill sessions in sports transfiguring data into actionable knowledge coaches, and decrease the impact of activity testing on athletes. Anthropometric measurements are commonly used to evaluate and predict accomplishments in different sports. Anthropometric and physical characteristics of an athlete can fulfil important preconditions for successful participation in any particular sports. Further, anthropometric profiles indicate whether the player would be suitable for the competition at the highest level in a specific sport. Recently, more researches have been carried out on Sport Data mining. In this study, we propose an approach to identify the most suitable sport for beginners using data mining and anthropometric profiles. We propose neural network in data mining approach. We apply a neural network in data mining technique and find the position of sportsman in graphical presentation. Empirical study of our approach has proved the effectiveness through Neural network.*

Key Words: Anthropometric Measurements, Neural Network, Morphological characteristics, Data Mining

1. INTRODUCTION

Anthropometric measurements are generally wont to determine and predict achievement in several sports. Anthropometric measurements and morphological characteristics play significant roles in determining the success of a sports person. An athlete's anthropometric and physical characteristics may perform important preconditions for successful participation in any given sport. Indeed, it can be assumed that an athlete's anthropometric characteristics, can, in some way power his/her level of performance, at the same time helping to determine a suitable physique for a certain sport. It has been well established that specific physical characteristics or anthropometric profiles indicate whether the player would be suitable for the competition at absolutely the best level during a selected sport .Data mining is the method of analyzer the concealed pat- terns of facts according to different outlooks for classification into beneficial information, which is composed and accumulated in areas as data warehouses, for effective analysis, data mining procedures, simplifying business decision production and other information necessities to ultimately reduce price and increase profits.

It was trusted that domain professionals (coaches, managers and scouts) could effectively possess their collected data into valuable knowledge. As various types of data collected increase in scope, more practical methods are found to extract helpful knowledge using data mining techniques. Properly leveraged Sports data processing techniques may result in better team performance by matching players to certain situations, identifying individual player contribution, evaluating the tendencies of opposition, and exploiting any weakness. Finally, the beginners are assigned for suitable sport according their Anthropometric measurements using Neural network in data mining.

2. Related Works

Data Mining has as of late become one of the most dynamic and promising fields for the extraction and control of information to create helpful information. Thousands of organizations are utilizing information mining applications consistently so as to control, recognize, and remove valuable data from the records put away in their databases, information vaults, and information stockrooms. With this sort of information, companies have had the option to improve their organizations by applying the patterns, relationships, and patterns that have lain covered up or unfamiliar inside Goliath measures of information. For instance, information mining has delivered data that empowers organizations to make profiles of current and forthcoming clients to help in picking up and holding their clients. Different employments of information mining incorporate advancement of strategically pitching and advertising methodologies, introduction of potential wrongdoings or frauds, finding designs in the entrance of clients to their sites, and procedure improvement. Normalization is scaling strategy or a planning method or a prepossessing stage. We can discover new range from a current one territory. It very well may be useful for the expectation or estimating reason a great deal. As we probably are aware there are such a large number of approaches to foresee or conjecture yet all can shift with one another a ton. So, to keep up the enormous variety of expectation and determining the Normalization method is required to make them closer. In any case, there is some current standardization methods as referenced in my theoretical segment to be specific Min-Max, Z score and Decimal scaling barring these strategies we are introducing new one procedure called Integer Scaling procedure.

3. Proposed system

Anthropometric measurements are commonly used to evaluate and predict accomplishments in different sports. Anthropometric and physical characteristics of an athlete can fulfill important preconditions for successful participation in any particular sports. Further, anthropometric profiles indicate whether the player would be suitable for the competition at the highest level in a specific sport. The weight, height, BMI, Skin folds, Arm circumference and Calf circumference and hip abdominal were selected as features and tested by Anthropocentric measurements. The improved system provides efficient analyzing for a player is suitable for this sport. The player which provides a data set which can be used to predict the player strength. A neural network in data mining technique is used to find the position of sportsman data in graphical presentation.

4. Methodology

The advent of knowledge mining confined, sports organizations almost exclusively on human proficiency. it had been trusted that domain professionals (coaches, managers and scouts) could effectively possess their collected data into valuable knowledge. As various sorts of data collected increase in scope, more practical methods are found to extract helpful knowledge using data processing techniques. Properly leveraged Sports data processing techniques may result in better team performance by matching players to certain situations, identifying individual player contribution, evaluating the tendencies of opposition, and exploiting any weakness. An athlete's anthropometric and physical characteristics may perform important preconditions for successful participation in any given sport. Indeed, it can be assumed that an athlete's anthropometric characteristics, can, in some way power his/her level of performance, at the same time helping to determine a suitable physique for a certain sport.

4.1 Data Set

Data set might also be a series of understanding. Most primarily an understanding set corresponds to the contents of one database, the place each and every column of the desk represents a precise variable, and each and every row corresponds to a member of the dataset.

Normalization is a scaling technique, a mapping technique or a preprocessing stage where we can discover a new range from current range. It can be helpful for giving prediction or forecasting purpose. The techniques which supply linear transformation on original range of data are called Min-Mix Normalization. The procedure which keeps relationship among unique information is called Min-Mix Normalization.

Here, we normalized values in all the features using following.

$$ai = \frac{vi - \min vi}{\max vi - \min vi}$$

Here, ai is normalizing value between land 0 and vi represents i-th attributes value of player. Min vi is minimum value of the i-th attributes. Max vi is maximum value of the i-th attributes.

5. Implementation

The system architecture consists of sports dataset and different data mining techniques and normalization techniques. In the datasets, we use technique form given chosen data set using normalization. It is a classification method primarily based on Bayes Theorem. A Naive Bayes classifier assumes that the presence of a unique in a category is unrelated to the presence of any different feature. It performs classification by finding the hyper plane that completely separates the vector into two non-overlapping classes. The weight, height, BMI, Skin folds, Arm circumference and Calf circumference and hip abdominal were selected as features and tested by Anthropometric measurements. The improved system provides efficient analyzing for a player is suitable for this sport. The player which provides a data set which can be used to predict the player strength.

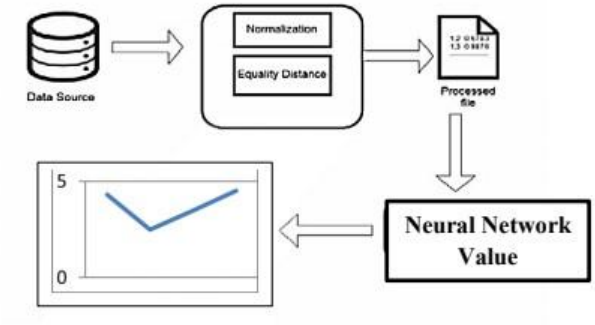

Figure 1: Implementation model

5.1 Algorithm for Support Vector Machine (SVM)

Input: Taking some of the records as a sample data and also, we took averages of Performance rating, Monthly Income and Training Times Last Year.

Output: Classified records

Step 1: We calculate the Anthropometric measurements in the sample data.
Step 2: Repeat Step 1 for Weight, height and BMI.
Step 3: Repeat Step 1 for Skin folds.
Step 4: Repeat Step 1 for Arm and calf circumference. Step 5: Compare each user's Weight, height, BMI, Skin folds, Arm

circumference and calf circumference with the Anthropometric measurements.
Step 6: Now take each record under that particular column and add them to the output.
Step 7: Remaining columns are also calculated according to step 4 and step 5.
Step 8: Now all these records are taken into a single class.
Step 9: We display these records.

5.2 Algorithm for Normalization

Input: Values of x over a mini-batch: $\mathcal{B} = \{x_{1...m}\}$;
Parameters to be learned: γ, β
Output: $\{y_i = \mathrm{BN}_{\gamma,\beta}(x_i)\}$

$$\mu_{\mathcal{B}} \leftarrow \frac{1}{m}\sum_{i=1}^{m} x_i \qquad \text{// mini-batch mean}$$

$$\sigma_{\mathcal{B}}^2 \leftarrow \frac{1}{m}\sum_{i=1}^{m}(x_i - \mu_{\mathcal{B}})^2 \qquad \text{// mini-batch variance}$$

$$\widehat{x}_i \leftarrow \frac{x_i - \mu_{\mathcal{B}}}{\sqrt{\sigma_{\mathcal{B}}^2 + \epsilon}} \qquad \text{// normalize}$$

$$y_i \leftarrow \gamma \widehat{x}_i + \beta \equiv \mathrm{BN}_{\gamma,\beta}(x_i) \qquad \text{// scale and shift}$$

Figure 2: Normalization Method

6. Experimental study and discussion

Normalization algorithm was implemented. The measurements of the players and the Anthropometric measurements were compared, which showed that the players are the most suitable for same game. The players were represented graphically by making a comparison, as is represented in figure 3.

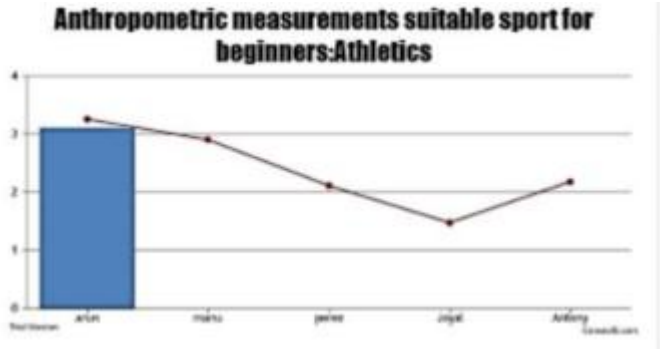

Figure 3: Graphical Representation

6. Conclusion

Anthropometric measurements are generally used to determine and predict achievement in different sports. Further, anthropometric profiles indicate whether the player would be suitable for the competition at the highest level in a specific sport. The research shows Anthropometric measurements are generally used to determine and predict achievement in different sports. Further, anthropometric profiles indicate whether the player would be suitable for the competition at the highest level in a specific sport. In this paper we have presented an approach that uses anthropometric profiles and data mining techniques to identify the most suitable sport for the beginners. Here we used Neural network technique. Thus, in Sports data mining proved useful in analyzing the data and gathering information and providing report and necessary training sessions to support and improve the efficiency and win rate of the sports.

7. Refrences

[1] Banage T. G. S. Kumara, Yuichi Yaguchi, Incheon Paik and Wuhui Chen, “Clustering and spherical visualization of Web services,” in Proceedings of the IEEE International Conference on Services Computing, 2013.
[2] Sun Ya Xue, “Research on the Influence of Big Data on School Physical Education Development,” Sports frontier, vol. 8, no. 6, 2016, pp. 1-3.
[3] S. Patro and K. Sahu, "Normalization: A Preprocessing Stage”, arXiv.org,2019. [Online]. [Accessed: 12-Oct-2018],
[4] S. Mining, R. Schumaker, O. Solieman, H. Chen and S. US, "Sports Data Mining | Robert P. Schumaker Springer", Springer.com, 2018
[5] J. Solarte, "A Proposed Data Mining Methodology and its Application to Industrial Engineering", Trace: Tennessee Research and Creative Exchange, 2018.
[6] D. Santos et al., "Reference Values for Body Composition and Anthropometric Measurements in Athletes", 2019.
[7] Joumal of sports. com, 2019