IE 7275 Data Mining in Engineering

Weightlifting Performance Monitoring

Group No : 2

Student name : Karan Parikh , Kumaran Nehru Uthra

1. Background and Introduction:

The approach we propose for the Weight Lifting Exercises dataset is to investigate "how (well)" an activity was performed. The "how (well)" investigation has only received little attention so far, even though it potentially provides useful information for a large variety of applications, such as sports training. We first define quality of execution and investigate three aspects that pertain to qualitative activity recognition: the problem of specifying correct execution, the automatic and robust detection of execution mistakes, and how to provide feedback on the quality of execution to the user. Class A corresponds to the specified execution of the exercise, while the other 4 classes correspond to common mistakes. Participants were supervised by an experienced weight lifter to make sure the execution complied to the manner they were supposed to simulate.

It is significant because performing the exercise with the wrong posture leads to injuries. Also, this generation is moved towards virtual monitoring and training , so our model can be used to help athletes train and the coaches to monitor the perform of their athletes without any susceptibility to injuries and posture of their exercise for maximum efficiency and utilization of their energy in the most efficient way. It can be used by newbies/ beginners to learn and monitor their performance so that they are not prone to injuries with zero knowledge about weightlifting. It can be developed to a proper business model in the form of a virtual or a new physical product incorporating with mobile application, google assistant etc, to earn profit.

This study focusses on classifying the errors into four categories in

1. Data Origin :

Six young health participants were asked to perform one set of 10 repetitions of the Unilateral Dumbbell Biceps Curl in five different fashions: exactly according to the specification (Class A), throwing the elbows to the front (Class B), lifting the dumbbell only halfway (Class C), lowering the dumbbell only halfway (Class D) and throwing the hips to the front (Class E).

Class A corresponds to the specified execution of the exercise, while the other 4 classes correspond to common mistakes. Participants were supervised by an experienced weightlifter to make sure the execution complied to the manner they were supposed to simulate. The exercises were performed by six male participants aged between 20-28 years, with little weight lifting experience. We made sure that all participants could easily simulate the mistakes in a safe and controlled manner by using a relatively light dumbbell (1.25kg).

1. Data collection:

• Sample is collected and stored using on-body sensing approach and ambient sensing approach, which is then logged.

• If the database has too many data information, PCA might be used to reduce the number of features.

• The output from the logged files are saved into a CSV file, containing 39243 rows and 158 columns

Following are some applications where the weightlifting monitoring can be helpful:

* • Help personal trainer and training assistants to correct and gauge the weightlifting performance of the athletes more accurately and seek for future injuries
* • Develop a device for beginners to check weather they are training effectively or not
* • Gauging one’s performance in terms of correct posture and technique

 The outcome of this model is Class A, Class B, Class C, Class D and Class E that tells us what kind of error the person deals with. The plan is to apply various statistical methods and machine learning algorithms to build the model. The performances among different methods is then compared for observing the factors that might affect the performance of the prediction

Data mining techniques and Implementation:

1. Classification 1 :

* Classification was done based on K-nearest neighbour algorithm.
* Converted the variable Classes into factors.
* Split the sample data into 70% training data and 30% validation data.
* Target variable was determined, and the predictors were chosen.
* Trained the model using K-nearest neighbour with predictors and target variable.
* We predicted this model on the validation data.
* Confusion matrix was made and the accuracy of the model was determined with different K-values and the best value is chosen from that.

Summary:

This case study determines the classification of different kinds of errors when an individual performs bicep curl in weight lifting using dumbbells and classify the errors with the help of data driven models. We obtained the dataset from UCI Machine learning website and it had 42896 and 53 columns with class being the errors made by the individual. We first performed PCA and built our train and test models with 70% training data and 30% being the validation data. We used four techniques to classify and predict the new dataset K-Nearest- Neighbours , Random forest, Support vector machine and Decision tree. We got the highest accuracy for random forest and recommend using that for our model based on the computation and accuracy and the lowest accuracy was obtained for X .

Data Exploration

1. Missing map
2. PCA – Dimension reduction
3. Bar chart
4. Normalize values