Intelligent Assistant for Basketball Coaching - Multiclass classification model comparison

Semir Salkić
Faculty of Computer and Information Science
University of Ljubljana
Ljubljana, Slovenia
Email: ss9343@student.uni-lj.si

***Abstract*—Human activity recognition is done based on the observation and analysis of human behavior to understand the performed activity. This paper is analyzing use case of this approach to AI basketball coaching application via IMU wearable sensors. Here, we present collected data and extracted features used in prediction models estimation. Prediction includes multiclass classification used for prediction of human activity. Extracted features from raw measurements is presented. Comparison analysis between SVM (Support Vector Machines) and MLP (Multi Layer Perceptron) is included.**

I. Introduction

As standardized and easily accessible computing platforms find their place in the multidisciplinary research of human activities in various applications, the popularity of wearable sensor technology is rising. In terms of the comprehensiveness of data collection and the meticulous measurement of human motion, the use of wearable devices is providing support for human activity recognition applications.

In kinesiology research, wearable sensors can be used to quantify the movements performed and provide insights into performance. Machine learning algorithms provide the possibility to quantify movement and provide reasoning assistance to field experts. Using the Internet of Things, you can explore large-scale databases and cloud-based data analysis and knowledge generation tools. Ideally, using these technologies for kinesiology applications will lead to truly autonomous artificial intelligence (AI) assistants. Core of this assistant is prediction model, which is capable to asses and accurately analyze certain human activity. In this paper we are going to showcase and compare performance of two multiclass classification models. SVM (Support Vector Machines) and Artificial neural networks - MLP (Multi Layer Perceptron). To showcase performance and estimations we are going to use data collected for research purposes in *Acikmese et al.* [1]. In this paper we are going to present comparison between selected predictive models and showcase extracted features. In Data section we are going to give detailed overview of raw data and our selected features which form the dataset. In Results model comparison is given and prediction results are given. In conclusion, reasoning for this topic is given and future plans of development are discussed.

II. Data

The data set is collected to identify basketball movements during basketball training. Ultimately, the idea is to track the long-term progress of basketball players based on the type and intensity of exercise they perform in training. During the data collection process, the wearable motion sensor was placed on the trainee's dominant hand, and the trainee was instructed to repeat 6 different types of basketball exercises, and each exercise took about 30 seconds. The exercises are carried out by 4 trainees. Edge computing devices receive sensor signals, mark and store them. Table I shows the types of exercises performed and their labels. In this layout, 6 classes are assigned to the exercises. With this in mind, a multi-class classification is realized. For the model training purposes, accelerometer (3 axes) and gyroscope (3 axes) signals are filtered using a low-pass filter (1) with $20\,Hz$ cut-off frequency. The signals are normalized to sensors range so that values fall between -1 and 1, where signals are segmented into 3 seconds"windows" with 50% overlap. Further, values are converted to be positive for the purpose of feature selection. In total 11 statistical features are then calculated for each window. Calculated features are *maximum, minimum, mean, median, variance, skewness, 25th percentile, 75th percentile, root mean square, mean absolute deviation, mean crossing rate*. The training dataset is formed such that rows are formed from the calculation of features for

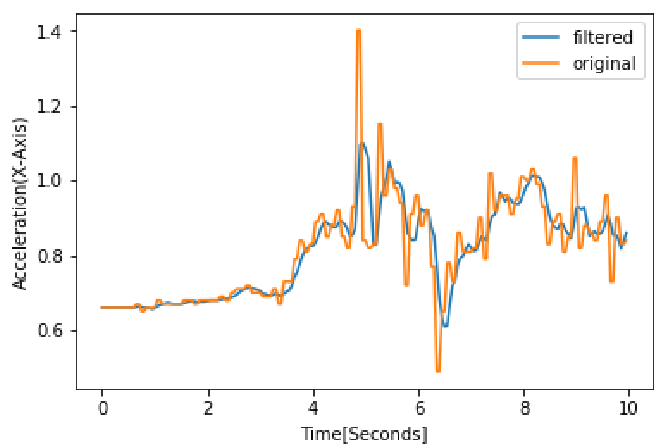

Fig. 1. Example of Filtered raw data

each of $3\,s$ windows for 6 signals. This results in a dataset with 66 columns (6 signals x 11 features) and class label for each of the exercises from Table I. Data contains 1055 data points (rows). Train (70%) - test (30%) set was created with random sampling.

TABLE I
EXERCISE TYPES AND CORRESPONDING LABELS

Label	Exercise Type
1	Forward-Backward Dribbling
2	Left-Right Dribbling
3	Regular Dribbling
4	Two Hands Dribbling
5	Shooting
6	Layup

III. RESULTS

We have used three kernels for SVM classification. Linear kernel performed poorly, while polynomial and radial basis (RBF) kernel performed similarly. Using grid search RBF kernel yielded best performance on our dataset. For ANN model, grid search procedure produced optimal configuration of 1 hidden layer of 100 perceptrons, while used optimization solver was ADAM. Optimal L2 regularization parameter is $\alpha = 0.05$. Having in mind that we have 66 features we have based our analysis on selecting best features which will increase our model performance. Table II presents top 10 features. These features are scored by univariate statistical tests performed on each feature. Scoring function used here is Chi-squared distribution. We can notice that most significant general feature for our models is acceleration, while most significant single feature is 25^{th} percentile Y-axis acceleration. After acceleration, gyroscope is important feature for our predictions. This result can be argued in terms of our classes. Each of these actions are sensitive in terms of acceleration and its direction in 3D space. This supports a claim that gyroscope features is in top 10 because each movement is strictly connected with orientation and angular speed. The highest average exercise recognition accuracy of the ANN model is 97.476%, while SVM model gets 98.422%. Figures 2 and 3 provides more clarity about the recognition capabilities of the trained models. The confusion matrix shows the correctly predicted instances in the main diagonal. All other elements of the matrix represent the misclassified data instances. It is obvious that the trained model cannot make a clear distinction between two similar dribble types (Forward-Backward vs Left-Right). These dribbles in terms of movements differ in the change of perpendicular directions respectively. Misclassification of the specified movements relies on the fact that certain labeled actions are practically the same movements with minor deviations in terms of sensor data which are limited by current hardware scope of sensing (differences can be viewed in slight changes between acceleration and gyroscope axes). Further improvements can be made by the introduction of direction information in the form of Euler angles obtained from the

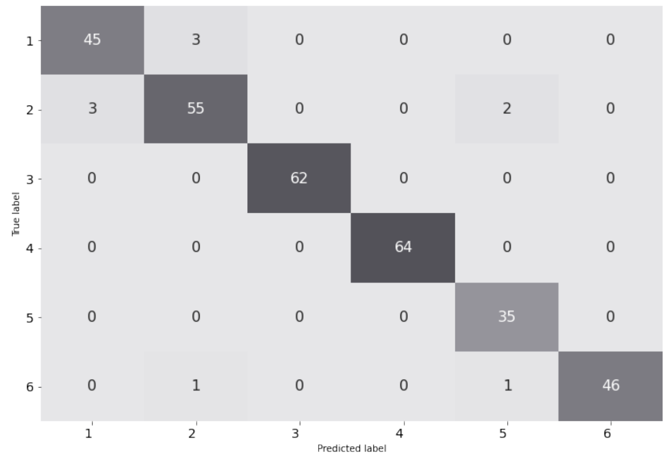

Fig. 2. Confusion matrix (ANN)

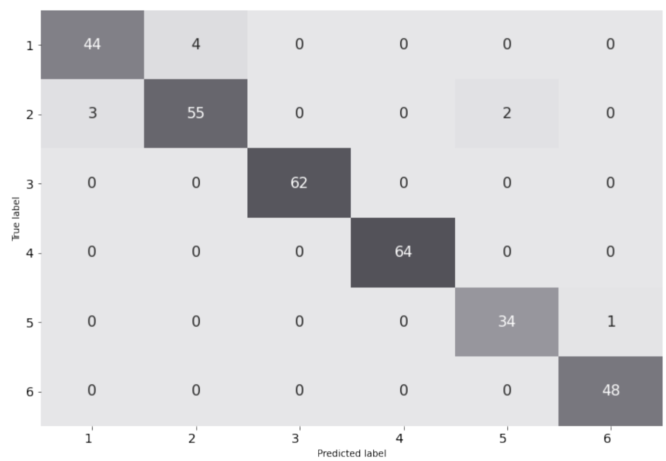

Fig. 3. Confusion matrix (SVM)

combination of acceleration and gyroscope signals [1]. We can also notice that both models are quite similar. Both of them are precise for prediction of Regular (3) and Two Hands (4) dribbling, while both of them fail to classify same two instances of Left-Right dribbling, where they are misclassified as shooting.

TABLE II
TOP 10 FEATURES (CHI-SQUARE SCORING FUNCTION)

Name	Value
Acc. Y $25^{th}\%$	155.809
Acc. Y min.	121.776
Acc. X max.	104.764
Acc. Z max.	103.430
Gyro. Y med.	97.465
Gyro. X max.	84.530
Gyro. X $75^{th}\%$	84.013
Gyro. Z $25^{th}\%$	81.175
Gyro. X skew.	80.657
Acc. X skew.	75.799

In Figures 4 and 5 precision and recall scores of models are presented.

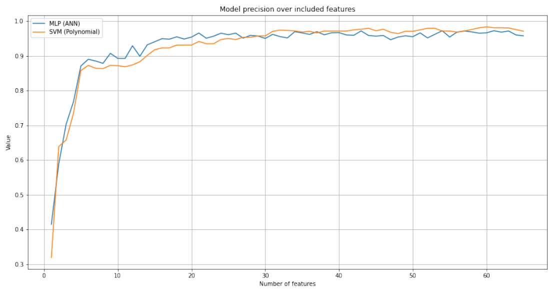

Fig. 4. Precision Score of models

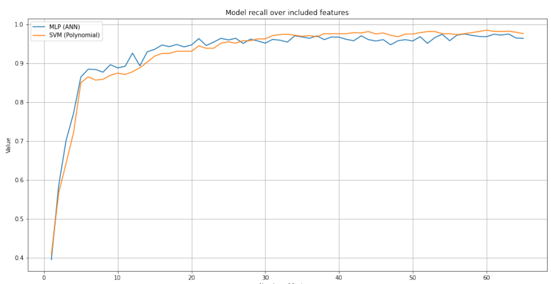

Fig. 5. Recall Score of models

X-axis shows number of included sorted labels by chi-squared scored, while Y-axis shows value of model scores. Expected behavior is that Recall score is following the curve of precision and vice-versa. It is important to mention that these values are average macro scores from model predictions. Figure 6 shows F1 Score.

We can see that common thing between all three scores is obvious superiority and stability of SVM model. We can see that for low values of included feature ANN model is better, but after 27 included features SVM shows consistent growth with increasingly included number of features. In this case we can see that generally after best 30 features models are have peak performance. For each extra included feature we have small increase in model performance. This can be viewed like in terms of accuracy and model complexity trade off.

IV. Conclusion

In the presented results we obtained meaningful insights. We can see that model gets most of the prediction information from the acceleration data, while some movements are problematic for classification due to its similarity. Besides that, we have presented results for our multiclass classification, where we can see that SVM with RBF kernel is better overall model for this application. Future plans include generalization of predictive modelling to further classification of human movement.

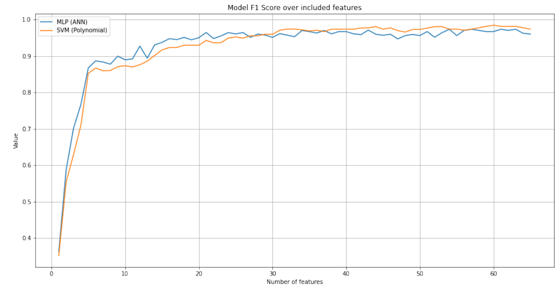

Fig. 6. F1 Score of models

References

[1] Acikmese, Y., Ustundag, B. C., Golubovic, E. Towards an artificial training expert system for basketball. 10th International IEEE Conference on Electrical and Electronics Engineering (ELECO), pp. 1300–1304. (2017)