Automatic Detection of Cutting Maneuvers Using Smartphone Sensor Data

Applied computer science in sports

Mohammed Hussam Al Turjman

Computer Science Department, Faculty 3, University of Bremen, alturjmm@uni-bremen.de

Joel Egharevba

Computer Science Department, Faculty 3, University of Bremen, jo_eg@uni-bremen.de

Ana Garzon

Computer Science Department, Faculty 3, University of Bremen, garzon@uni-bremen.de

ABSTRACT

The need to detect cut maneuvers in ultimate frisbee is rising. The popularity this game gained in recent years, opened a door for researchers to help players enjoy and master playing their game. In this study we focus on detecting cutting maneuvers in ultimate frisbee based solely on motion data, recorded by a smartphone. We show that it is possible to classify cut movements with high accuracy using different machine learning models. Furthermore, the models can be utilized to build complex systems that can work in real-time.

CCS CONCEPTS

• Applied computing • Machine learning • Human-centered computing

KEYWORDS

Time series classification, sensor data, marginal sports, human activity recognition, sports science

1 Introduction and Motivation

Ultimate frisbee is a common marginal sport that can be played in teams. It is a non-contact and self-regulated game with a flying disc played; it is played between 2 teams of 7 players each in an American Football field (100x37 meters), it has two score areas of 16 m. The objective of the game is catching the disc to score points on the score area. Today Ultimate Frisbee is considered a recognized sport at the Olympics and it is possible for it to form part of 2028 L.A [1]. Olympic Games. Due to its rising popularity, more research is starting to appear on the matter, mainly regarding the difference of performance among professional players and amateurs. One of the important aspects in ultimate frisbee is performing the so-called “cuts” without injuries. A cut can be described as a running maneuver to trick other players by changing directions suddenly. A study in the field of medicine conducted by Ou-Suet Pang, Florence et al, found that the injury prevalence was 62.7% and from this percentage lower limb was the most common injury with a 61.1%, which is caused mainly by cutting; this study was conducted in Hong Kong [2]. It is therefore essential to carefully learn how to execute a cutting maneuver. It should be mentioned that cutting is an integral part of many field-based team sports, making this study applicable to an even wider field. Furthermore, the ability to do well-executed cuts after a short reaction time, provides a considerable advantage in most of them.

In this paper we propose an out of the box solution to automatically detect cuts using smartphone sensor data. For that we conducted an experimental study to collect sensor data from ultimate frisbee players with experience ranging from beginners to experts. The collected data was then used to train different models. Our aim is to provide a simple way for athletes to detect their cuts with minimal additional gear

2 Related Work

Due to the rise of wearable smart devices, equipped with capable movement sensors, research in human activity recognition (HAR) has increased during the last years [3]. Studies in this field often instruct their participants to perform certain activities in a controlled environment ([4], [3]), but also equip athletes with sensors during competitions in field studies [5]. The sensors used for this purpose, most of the time, include accelerometer and gyroscope ([4], [5]) and sometimes magnetometer [3]. The raw data is then either further processed with transforming techniques like the Fast Fourier Transformation [6], or even fed directly into specifically designed convolutional neural networks (CNNs) [3]. While CNNs seem to be the predominantly used classificator ([4], [3]), more conventional models fed with carefully extracted features can still compete with high accuracies [6]. Besides CNNs, LSTM (long short-term memory), a form of recurrent neural networks (RNN) have been used successfully as well [7].

While there are studies regarding the more strategic aspects [8], research related to Ultimate Frisbee is mainly focusing on two topics. Dynamics of throwing the disc ([4]) and injuries [2]. In order to reach an optimal performance and reduce the injury levels, proper training of cutting techniques for the players is needed; that is where we would like to continue developing this project, which is currently in its first stage: Detecting cuts through motion sensors.

A research study conducted by Johannes Link, Timmur Prest, Maike Stoeve (2022) [4] developed activity recognition for Ultimate Frisbee. They focused on classifying disc-throwing actions of participants with experience in the sport. One key objective of the study was utilizing research of more popular sports through transfer learning. The participants of the study wore inertial measurement units (IMUs) on their dominant hand's wrist. Sensors were sampled at 512 Hz and recorded acceleration and gyroscope in the 3 axis. The recorded signal was manually annotated with the help of video review. They used a peak detection algorithm based on Z-Scores ([9]) to detect activities that were not throws, like running to catch a disk. For this they also disregarded the gyroscope data. For the classification they used two layers for the input, one for acceleration data and one for gyroscope. The sample data was augmented with three different techniques. One being rotating the samples in 3D space. A CNN with two input layers, one for each accelerometer and gyroscope, was developed, which uses the Adam's algorithm [10] for optimisation during training. As a result, they were able to reach a 66.6% of accuracy when classifying 9 different actions and 89.9% with five actions, after generalizing six different throws into two main techniques. They also mentioned that CNNs have a good performance for time series classification and for small training data sets with transfer learning. In the case of this particular research 14 participants took part.

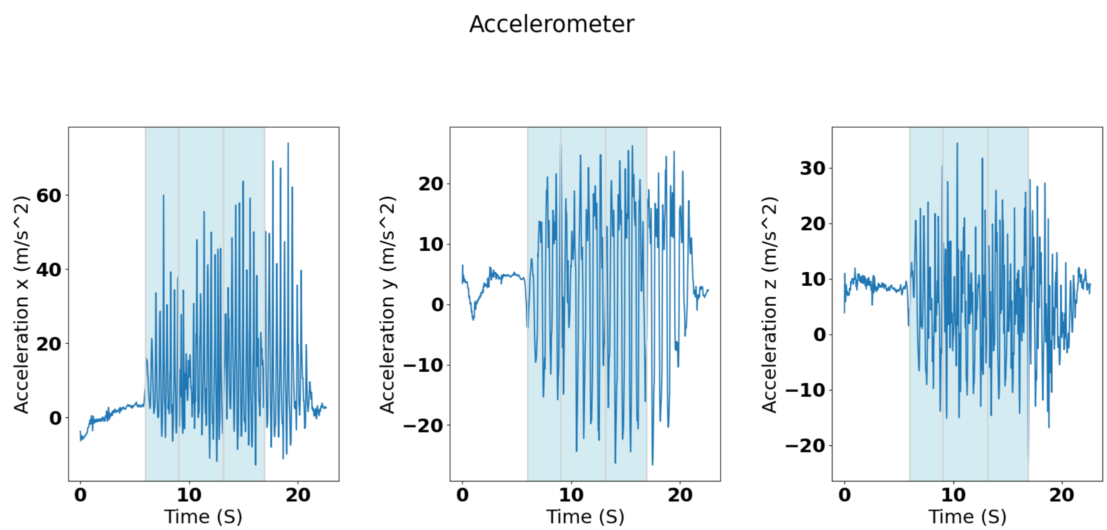
Another study focused on biomechanical analysis and paid special importance on the cutting movements, with injury prevention in mind. Slaughter, Paul and Adamcyzk, Peter (2020) [5] attached an IMU system with seven sensors on the participants lower body. Cuts were identified by footfall patterns. Footfalls were defined as the period of a foot's first to last contact with the ground and were detected by the IMU system. Additionally, the degree of knee flexions was recorded, to possibly correlate them with risks of ACL (anterior cruciate ligament) injuries. Although participants were in games during the experiment this research could be also done under controllable environments. The players in the study averaged 90º cuts and it is stated that filming could have helped in separating cuts by possession. According to this research wearable sensors are accurate enough for distinguishing cuts of players.

While sports like soccer or basketball have plenty of research to improve athletes performance on the field, we observed that marginal sports usually may not have that many opportunities, due to lower popularity and the resulting fewer interest and high quality resources and samples.

3 Description of the Approach

The problem at hand can be formulated as a classification problem and this is due to the simple fact that we want to distinguish between a cut maneuver and other kinds of movements. In this study we use a sequence-based approach to reasonably classify cuts. We regard the samples of each data frame in the data-set as a frame of features. The data frames can be observed in time windows.

To better illustrate the used approach, we can observe the following figure:



Accelerometer recording by phyphox

It shows the acceleration data of a person performing multiple maneuver cuts in ultimate frisbee. The X-Axis shows the time in seconds and the Y-Axis shows each coordinate axis for each dimension X,Y and Z respectively. The areas marked in light-blue are the actual time windows, where cuts were performed. Furthermore, the data were recorded using a smartphone-app called PhyPhox [11].

The idea basically is to label each sequence of data samples as CUT or NOT-CUT and then train different models on a dataset that we gathered from ultimate frisbee players. After evaluating all the trained models, the best model can be then employed to detect cut maneuvers based solely on motion data from sensors such as accelerometer and gyroscope.

4 Study

In this section we describe the process of gathering the data and the classification pipeline.

4.1 Data Acquisition

For the purpose of solving the problem of automatically detecting cut movements, we conducted an experiment to gather sensor-data. For this, a rather novel approach was developed. The goal was to allow the participants to record training data remotely, thus possibly limiting the entry barrier and opening the study to more people. Initially, every participant received an audio and a video clip. as well as a file containing the experiment configurations. With the video, participants were informed what the data would be used for and how to properly participate. The audio clip was to be played during the experiment and gave instructions on which action to take at exactly which moment of the recording, effectively synchronizing the participants with the experiment. Since every time-sensitive action was preceded with a three second countdown, the reliance of the participants' quick reactions was eliminated. Every participant was asked to use PhyPhox [11] to record the sensor data while holding the phone in whichever hand they felt most comfortable with. They were also told to swing their arms as naturally as possible. The recorded experiment consisted of following sensor data: acceleration, gyroscope and linear acceleration at 60 Hz each. Each participant performed three cuts with 180° angles. Although a controlled environment could not be guaranteed, this process could have provided enough control over the experiment execution, as the results might suggest. The resulting data was then exported and sent to us. The synchronicity of the experiment with our audio then allowed automatic labeling based on the timestamps.

The experiments' participants mainly consisted of attendees of an ultimate frisbee class at the University of Bremen and had different levels of cutting skill. The sample size of 17 was roughly split 50/50 between men and women in the age range of 19 to 28. We were able to gather 17 recordings, each one of them is 22 seconds on average.

4.2 Classification

As mentioned in previous sections, we experiment with various models based on two strategies. The first one is to only use motion features and the second one is to use high level features. For the first approach we employ an LSTM model to classify the gathered data in fixed size windows. Furthermore, for the high level features approach, we utilize five faier classifiers. Namely the stochastic gradient descent, random forest, linear support vector machine, ada-Boost and gradient boosting classifiers. To be able to train the model, the data have to go through preprocessing steps. Initially, the compressed files of the sensor data are extracted automatically and they contain comma separated values for each sensor. All the sensors used have a frequency of 60 Hz. Thus, eliminating the need to align different sensor data.

Since the files can be loaded as data frames, they can be easily manipulated and used. After merging the all data frames and adding the correct respectiv label .i.e CUT/NOT-CUT for every row in the data frame, we obtain 9 spatial features, a single time feature and a label. The following table exhibits an instance of two rows of the raw features:

Time (s)	[12.12, 28.79]
Acceleration x (m/s^2)	[-3.75, -3.76]
Acceleration y (m/s^2)	[3.43, 4.65]
Acceleration z (m/s^2)	[3.88, 7.43]
Gyroscope x (rad/s)	[-2.16, -1.1]
Gyroscope y (rad/s)	[0.42, 0.28]
Gyroscope z (rad/s)	[-0.89, -0.56]
Linear Acceleration x (m/s^2)	[0.0, -0.07]
Linear Acceleration y (m/s^2)	[0.0, -0.19]
Linear Acceleration z (m/s^2)	[0.0, -0.83]
Label	[0, 0]

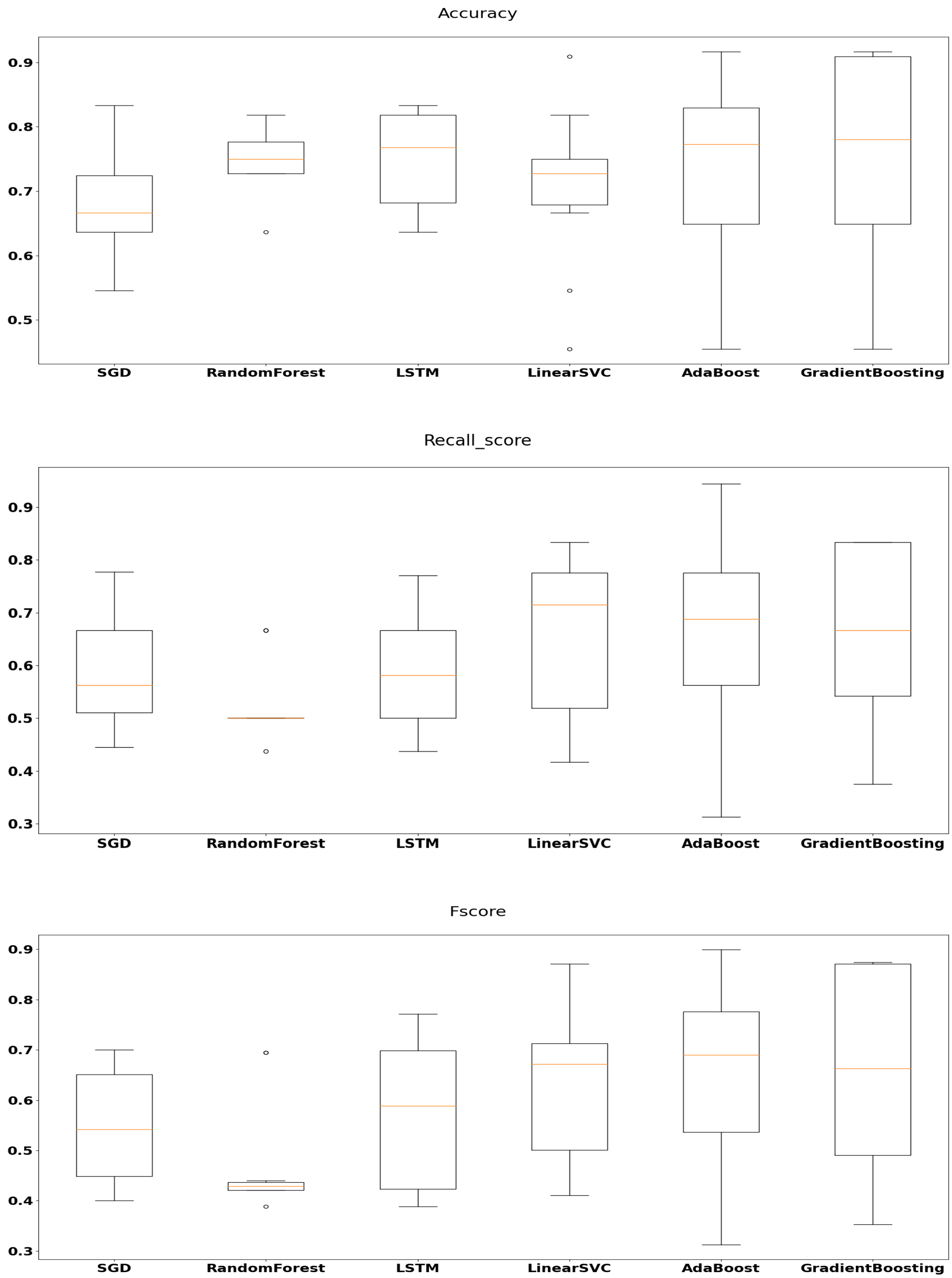
Features of the sensor data with example values

The next step after preparing the desired format of the features, is to create a dataset consisting of $\boldsymbol{n}$ data frames for every recording we gathered. Afterwards, we slice each data frame into windows with a fixed length. At the end the model is trained with the k-fold cross validation approach, where in every training iteration the dataset is split into three different sub-datasets. The resulting sub-datasets are then : train, test and validation sets. Since our dataset does not contain a lot of data, we restricted the test and validation sets to only containing a single data frame.

Whereas the low level features approach uses the spatial features as they are without modification except for scaling, the high level features approach extracts high-level features out of the raw signals. Those high level features are selected peaks of the fast Fourier-transform (FFT) and the power spectral density function (PSD) out of the scipy library. Those two functions similarly decompose a window of sample signals into its frequency spectrum. Expecting a significant change of the distribution of frequencies, the amplitude and frequency of the three highest peaks within a certain distance (in frequency) of each other are selected. As another, more raw feature, the real mean square error of these windows is also added to the feature set. Since these features are extracted from each of the nine sensor signals, the feature space's size grows to 117. The sample size, depending on the overlap of the windows, decreases in size.

5 Results

The different models examined in this study have been trained on the gathered data-set and evaluated with 10 folds of the k-fold validation approach. In the following figures, we show a comparison of the performance of the models with respect to three metrics, namely accuracy, recall and F-score. All models except LSTM have been trained with the frame-based approach. The confusion matrices shown in the appendix, further visualizes the performances for each class.



Evaluation metrics of 10-folds on all models

6 Discussion and Limitations

The results with our test samples show that, except for RandomForest, the accuracy metrics are highly dependent on the distribution of test and training sets. However, while showing robustness under the k-fold cross validation, RandomForest underperforms on average compared to the other algorithms. It is unclear which sensors and features contributed to correct predictions and which only introduced distracting noise. In theory, the linear acceleration should be the most indicative of a cut and the gyroscope the least, but further investigation would need to be done to confirm. Furthermore, there were a lot of parameters for mainly the frame based models that were not thoroughly optimized, due to time limitations. Such parameters include the window size and offset, peak selection in FFT and PSD, sample frequency and model-specific hyperparameters.

We are aware that the high variance of the results may be mainly due to the small training data. The developed process for data acquisition, while in theory convenient, needed a couple revisions and additional instructions until usable data could be recorded. This resulted in only 17 data frames, which is, in our opinion, too little to train any machine learning model. However, a conventional lab study with participants in presence, while resulting in more reliable data, likely would not have increased the number of samples. A more convenient way, i.e. an app specifically for the purpose of recording audio-synchronized training data, would have probably resulted in significantly more samples.

Still, the results are somewhat suggesting that a robust model with high accuracy is in the realm of possibilities.

7 Conclusion and Future Work

In conclusion it can nevertheless be said that cut detection using smartphone sensor data is by all means possible, provided a larger sample size, if the features are thoroughly evaluated. In addition we see that using high level features lead to better results with a subset of the tested models. Specifically, the AdaBoosting and GradientBoosting classifiers delivered higher accuracy than the approach with the low level features. On the top of that, it is worth mentioning that adjusting the hyperparameters of the individual models can play a big role in their performance.

With a little improvement, the developed models could be used for base level training applications, like measuring cutting reaction time. Future work could furthermore focus on detecting indicators for specific techniques to help improve cutting performance.

REFERENCES

[1] Nelsen, M. (2021, December 9). *Flying Disc hoping to glide into spot on LA28 sports program - Infobae*.
https://www.infobae.com/aroundtherings/articles/2021/12/09/flying-disc-hoping-to-glide-into-spot-on-la28-sports-program/

[2] Ou-Suet Pang, Florence, Chi-Wai Man, Gene, Ka-Kann Ling, Samuel, Shu-Hang Yung, Patrick(2021). Injury Epidemiology of Ultimate Frisbee in Hong Kong, *Asia-Pacific Journal of Sports Medicine, Arthroscopy Rehabilitation and Technology,*26, 27-31. https://doi.org/10.1016/j.asmart.2021.07.006

[3] Qi, W., Su, H., Yang, C., Ferrigno, G., de Momi, E., & Aliverti, A. (2019). *A Fast and Robust Deep Convolutional Neural Networks for Complex Human Activity Recognition Using Smartphone*. https://doi.org/10.3390/s19173731

[4] Link, Johannes, Prest, Timur, Stoeve, Maike & Eskofier, Bjoern(2022). Wearable Sensors for Activity Recognition in Ultimate frisbee Using Covolutional Neural Networks and Transfer Learning.*Sensors* 22,2560. https:// doi.org/10.3390/s22072560

[5] Slaughter, Paul & Adamczyk, Peter (2020). Tracking Quantitative Characteristics of Cutting maneuvers with Wearable Movement Sensors during Competitive Women's Ultimate Frisbee Games. *Sensors.* 20,6508. https://doi.org/10.3390/s20226508

[6] Mashita, T., Shimatani, K., Miyamoto, H., Komaki, D., Hara, T., Kiyokawa, K., Takemura, H., & Nishio, S. (n.d.). *Human Activity Recognition for a Content Search System Considering Situations of Smartphone Users*.

[7] Bock, M., Hölzemann, A., Moeller, M., & van Laerhoven, K. (2021). Improving Deep Learning for HAR with shallow LSTMs. *Proceedings - International Symposium on Wearable Computers, ISWC*, 7–12. https://doi.org/10.48550/arxiv.2108.00702

[8] Lam, Hilary, Kolbringer, Otto, Lames, Martin & Russomanno, Tiago (2021). State Transition Modeling in Ultimate Frisbee: Adaptation of a Promising Method for Performance Analysis in Invasion Sports. 12, 664511.https://doi.org/ 10.3389/fpsyg.2021.664511

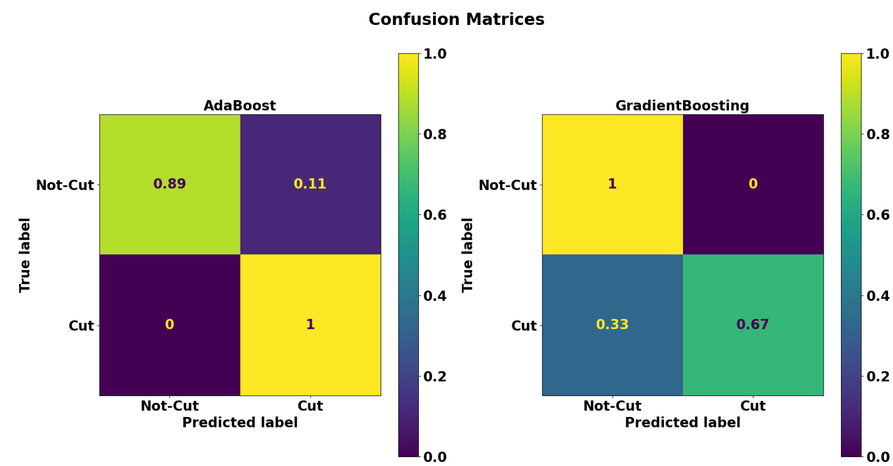
[9] Brakel, J.P.G. van (2014). "Robust peak detection algorithm using z-scores". Stack Overflow. Available at:

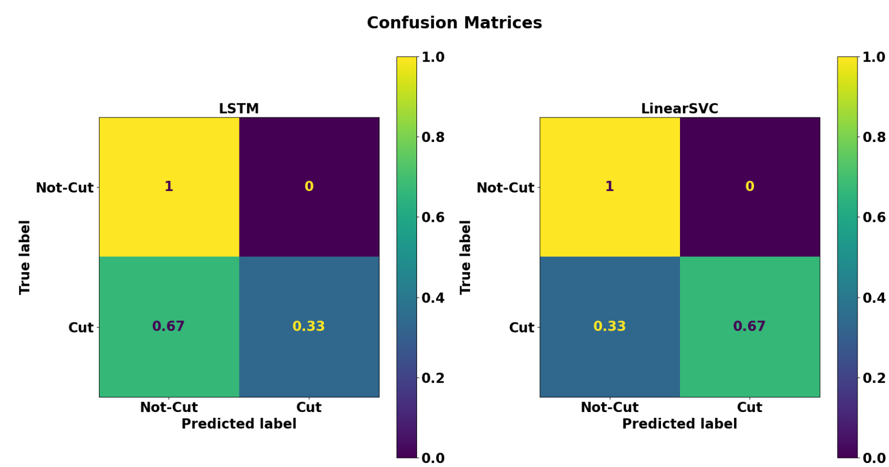
https://stackoverflow.com/questions/22583391/peak-signal-detection-in-realtime-timeseries-data/22640362#22640362 (version: 2022-08-12).

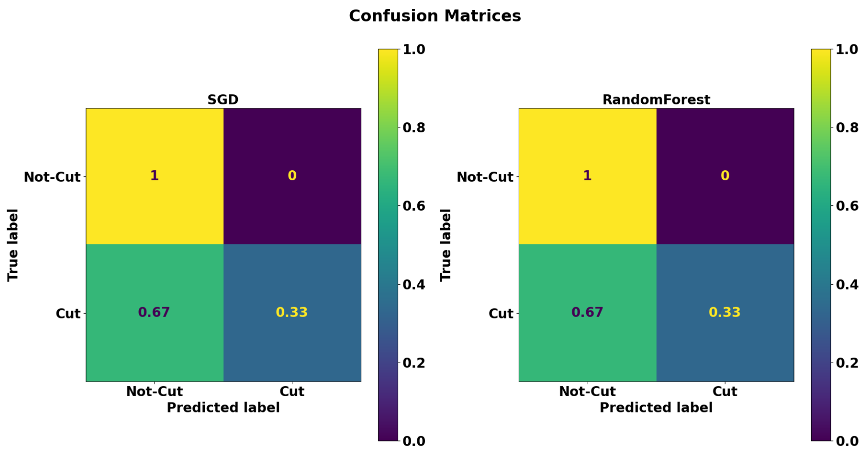
[10] Kingma, D. P., & Ba, J. L. (2014). Adam: A Method for Stochastic Optimization. *3rd International Conference on Learning Representations, ICLR 2015 - Conference Track Proceedings*. https://doi.org/10.48550/arxiv.1412.6980

[11] PhyPhox https://phyphox.org/

A APPENDICES






Confusion Matrices
SGD
RandomForest
Not-Cut
Cut
True label
Predicted label
1
0
0.67
0.33
1.0
0.8
0.6
0.4
0.2
0.0