Classification in football

Activity classification using sensor data in football

Georgios Kaketsis

TU Delft

Classification in football

Activity classification using sensor data in football

by

Georgios Kaketsis

to obtain the degree of Master of Science in Applied Mathematics with specialization in Stochastics at
the faculty of Electrical Engineering, Mathematics and Computer Science
at the Delft University of Technology.
This Thesis Project will be defended publicly on July 1, 2020 .

Student number:	4380894	
Project duration:	October 1, 2019 – June 13, 2020	
Thesis committee:	Dr. Jakob Sohl,	TU Delft, Daily Supervisor
	Professor Dr. ir. Geurt Jongbloed ,	TU Delft, Responsible Professor
	Professor Dr. ir. Kaspar Jansen,	TU Delft, External Expert

TU Delft

Preface

The goal of this report is to present and describe the effort surrounding the completion of the Master Thesis Project of classification in football. Classification is a procedure which belongs in the field of Statistics. The objective is to capture, detect and distinguish certain actions relative to the environment of analysis. In our case we are focused on classification for actions related to football, using sensor data. As a first step we introduce the motivation that enforces our project and also the nature of the sensors that provide us with the information for our analysis. We follow with some enlightening review of previous research on motion recognition related projects, in order to have some supplemental information that will provide us with experience that will serve as guidance and further direction. Afterwards we introduce the methodology of the classification. The methodology includes all the models and tools needed to achieve a precise and robust classification outcome. As a next step, we dive through the details of the experiment we are going to analyse, while we explain the process followed in order to create a refined data set that will consist the input of our models. After the description of the data and the preprocessing procedure, we present the results obtained by the analysis along with the evaluation and relative comparisons. Finally, we give the most important conclusions we reached in the whole process along with some proposals for future improvements.

Acknowledgements

I would like to thank my Supervisors Dr. Jakob Sohl and Professor Dr. ir. Kaspar Jansen for their support and understanding through this long and extensive procedure. I also thank Mister Erik Wilmes who kindly provided the data utilized in this project. It was an effort with a lot of challenges, but i tried to rise to the occasion as a student. I did my best to sustain my motivation and dream for mathematical research as fuel for my stamina and perseverance in this academic year. The strength and prayers of my family kept me going till the very last moment while trying hard to meet and exceed the requirements for a successful project. I wish from the bottom of my heart for this research to contribute and prove to be enlightening for science and open paths for future scientists that search for guidance and assistance. At last, i dedicate this effort and project to my dearly beloved family who stood by my side all this time and especially to my Mother Eleni Gkika Kaketsi, my Father Konstantinos Kaketsis and my Grandmother Euthalia Provou.

Contents

1

Introduction

1.1. Motivation behind the classification of acceleration signals

In the soccer world the coaches and athletes put their interest in statistics like number of shots and passes during training sessions and tournaments. The main reason to use acceleration signals is that cameras or GPS sensors cannot detect the forces on the upper leg, knee and hip. Monitoring these forces can help to prevent injuries. The primary motivation for this project is to develop an efficient classifier that will be able to recognize and classify acceleration signals produced by sensors the football players wear, while doing specific soccer related activities. The athletes that participate in the experiment wear five sensors (right thigh, left thigh, pelvis, right shank, left shank). The mentioned method has the potential to be a low-cost inertial sensor based approach for activity classication in soccer, which can be used by teams with low budgets and elite teams as well. This approach will be able to revolutionize football, and the amount of competition will keep rising because every kind of team will have access to premium quality information, that will enable them to improve their club and also the skills of their players. An additional advantage will be the enlightenment about the strengths and weaknesses of the players of a team based on the quantified information of their movements and ways of handling the ball in the field. Furthermore the coaches will be enabled to efficiently plan the appropriate training schemes in order to strengthen and improve the weak points of their players.

The big picture and aspiration of this project, is to be able to construct a model that will be able in the near future to provide football teams of any budget, with reliable information, in order to improve their progress. This can be achieved when the coaches, trainers and recruiters of a team utilize the facilities that the method provides. This way they can extract vital knowledge on the strain of the body of the participants on specific exercises and drills. The reliable classified acceleration data, can be studied and correlated to certain injuries that are inflicted on the players, due to excessive pressure on the muscles that are responsible for their exact movements. Based on the experience on certain exercises, the training tactics can prove to be the most efficient they have ever been by minimizing the amount of effort and strain on the subjects and maximizing their endurance, stability and stamina. Subsequently, the career of the football players can be extended while they remain in healthier physical condition, by protecting their bodies from excessive short and long term damage.

1.2. IMU (Inertial Measurement Unit) Sensors

In scientific research that demand data acquisition, the IMU sensors are broadly used. The specific devices are wireless and can be applied on the limbs of the athletes that perform certain experiments that include explosive motions, running and direction changing. The IMU can efficiently detect and capture the acceleration of the body parts they are placed on. Because they are comfortable and easy to wear, they are chosen very frequently in the wide world of sports. Specifically for our case the IMU sensors are responsible for detecting the movements of the soccer players they are placed on. The standard IMU sensors include an accelerometer, a gyroscope and a magnetometer. For the purpose of this research we are going to use the acceleration data, which measure the acceleration of the specific limb they are placed on. We will also use the gyroscope data which measure the angular velocity of a specific body part. The data we have at our disposal come on the three axis scheme relative to the sensor for the three dimensional space. In Figure 1.1 we provide the IMU sensor used in the project.

Figure 1.1: Sensor unit as used in 2020 version of the Sensor Shorts (21*24*8 mm) at the project CAS P6 of TU Delft

1.3. Classification of data

The strategy for the generic classification of data begins with the establishment of the wearable sensors to the subjects relative to the experiment. The next step is to define the features of choice that will consist the input data representative to the activities performed by the subjects, for the classification algorithm. Further we split our input data to a training set and a test set. The purpose of the training set is to assist the classification algorithm to distinguish the patterns of our data, so that it can be able to recognize the nature of the test data set, which is used for the validation of the algorithm. At this point the features for the training and test data are chosen and extracted. After the completion of the training phase, the model is applied on the test data followed by the assessment of the effectiveness and precision of the algorithm. In Figure 1.2 the general classification scheme is presented.

Figure 1.2: Schematic diagram for classification[13]

1.4. Research questions

In this Section we formulate the most important research questions that concern the specific project. After the completion of the procedure the research questions will be answered on Chapter 6, based on the conclusions and remarks that were born from our analysis of classification in football.

1. Should we focus on the analysis of raw three dimensional data or should we consider the Euclidean Norm also? Do we save a significant amount of time by analysing only the norm of the data?

2. Is it more effective to use all of the five sensors or three sensors (pelvis, right thigh, left thigh) are enough? How much loss in accuracy there is for the simplified version of the sensor shorts?

3. Do we need the settings S2 (gyroscope and spatial domain features) and S5 (gyroscope with spatial and spectral domain features) for the raw data case or the settings S1 (acceleration with spatial features) and S4 (acceleration with spatial and spectral features) also?

4. Should we utilize only spatial domain features or mixed features (Spatial together with spectral domain features)?

5. Which of the six settings works best and which one gives the lowest accuracy?

6. Does the classification accuracy increase or decrease if seven instead of four actions are considered and why?

2

Literature overview

2.1. History and development on classification in sports

Nowadays tactics in sports are becoming strongly dependent on the mathematical field of data science. Machine Learning methods are utilized in order to capture and detect complicated patterns in the data[6]. In the early stages of this model based direction, the researchers were mainly focused on sports related to specific and easily identifiable activities and moves. Such a sport is baseball as it encompasses explosive and agile movements of the players arms to hit the ball[6]. As the advances in Stochastic mathematics and specifically machine learning touched a high point the research became way more broad and focused on more complex sports[6] that include a wider variety of actions like football or basketball. Those sports for example, demand continuous movements and extensive ball handling[18]. Machine learning models that are up to date are capable of classifying these connected movements of the players effectively. Ways of capturing the data from the movements of the players usually come from video imaging or on-body sensor detecting. The reason behind the classification involved in sport related activities is to monitor the physical performance of athletes[6] and also to acquire enlightening information on the pressure concentrated on the bodies of the athletes to prevent injuries, which is the main idea discussed in this project. In the following paragraphs we briefly present some significant research related to classification in sports along with the results of the classification.

In the specific analysis[7] the researchers were focused on Tennis, and the classification concerns the three main strikes in a tennis match (forehand, backhand, serve)[7] using Wireless Inertial Measuring Units (WIMU). The maximum accuracy achieved using fusion of the sensors was 90%[7] using the Naive Bayes classifier.

The goal of this research[2] was to classify human physical activities of different kinds of intensities using wearable accelerometers. Those activities are low speed walking, high speed walking, sitting, shoulder lifting, squatting, jumping[2]. The Figure 2.1 which shows the results of the classification by comparing the predicted classes with the actual classes is presented below. The results are obtained using the k-Nearest Neighbors classifiers.

		Predicted class					
		Low speed wal-king,%	**High speed wal-king,%**	**Sitting, %**	**Shoulder lifting, %**	**Squatting,%**	**Jumping, %**
Actual class	**Low speed walking**	88.9	0	0	0	11.1	0
	High speed walking	11.1	83.4	0	5.5	0	0
	Sitting	8.33	0	83.3	8.33	0	0
	Shoulder lifting	0	5.4	5.56	66.7	22.2	0
	Squat-ting	0	11.1	0	5.56	77.8	11.1
	Jumping	0	5.56	0	0	5.56	88.9

Figure 2.1: Confusion Matrix showing the correct classifications and misclassifications[2]

This analysis[6] concerns Table Tennis and utilizes Inertial Measuring Units (IMU) for the movement detection. The eight strokes that were classified are the following: Forehand drive, Forehand push, Forehand block, Forehand topspin, Backhand drive, Backhand push, Backhand block, Backhand topspin[6]

The Figure 2.2 shows the accuracy of the predictions in a similar way with the previous case.

Type	FD	FP	FB	FT	BD	BP	BB	BT
FD	**401**	0	2	1	0	0	0	0
FP	0	**183**	0	0	0	0	0	0
FB	12	0	**237**	0	0	0	0	0
FT	4	0	0	**292**	0	0	0	0
BD	0	0	0	0	**261**	0	14	2
BP	0	1	0	0	0	**179**	0	0
BB	0	1	0	0	19	0	**141**	0
BT	0	0	0	0	8	0	0	**213**

Figure 2.2: Confusion Matrix showing the correct classifications and misclassifications[6]

2.2. Examination of previous research on classification in sports

In this chapter we are going to make a review on activity classification studies available in literature, focusing on sensor data. It is of significant importance to mention that the direction for this chapter is to dive through the way of thinking and approaches of different scientific teams that handled problems of classification and activity recognition in sports. For a subject like classification, which is heavily based on research, the consideration of the experience of experts might prove to be crucial for the progress of the analysis. Based on the insight that will be achieved through the project, all the results and observations will be taken into account in order to shed light on the search for the optimal classifier concerning the problem of classifying sport related data from different situations.

2.3. Review on Literature studies on activity classification

2.3.1. Skateboarding

Introduction

The specific study[10] includes an informative and enlightening application of classification in skateboarding. As in soccer it is of significant importance to be able to distinguish and isolate the movements of the subjects through the signals provided by the IMU sensors. In this particular case the researchers where interested in recognition of tricks performed by the subjects using the board. An important observation for the research[10] is that the team considered time windows which included time intervals related to the vital motions related to the activities in examination. "Based on considerations about the length of a trick and the duration of the landing impact, the length of the windows was set to 1 second with an overlap of 0.5 seconds"[10]. The experts computed features as the mean, variance, kurtosis and skewness on all of the data belonging to the time windows mentioned above [10]. The specific features were calculated on the raw data that an IMU produces, meaning the three dimensional data for (x,y,z) axis but also on the norms of the three dimensional sets[10]. For the evaluation, four fundamental classification techniques were used. The Naïve Bayes (NB), Partial Decision Tree (PART), Support Vector Machines (SVM), k-Nearest Neighbors (kNN). These methods were applied in the appropriate time windows that included the data relative to the tricks. Below follows the Figure 2.3 which includes the accuracies, the computational operations and the computational times of the methods[10].

	NB	PART	SVM	kNN
accuracy [%]	97.8	93.4	97.8	96.0
computation	low	low	high	middle
- operations:	360	41	1015	1086
- time [s]:	6.2	10.6	32.7	5.2

Figure 2.3: Accuracies and computational costs[10]

Connection to current project

The experts of [10] utilized norms to create a single representative vector of data that consist the new signal that is going to be the input or training data for the classification algorithms. The use of norms is also considered in the current project along with the time windows to isolate the useful activities from the noise. The use of the mentioned spatial features of [10] will also consist a part of our research. Another similar vital tactic that is going to consist the evaluation of the models of the current project, is the comparison of the predictive precision of our methods and also the comparison of the computational costs. The approach taken by the experts involved in the skateboarding experiment[10] seems to be encouraging, as the steps introduced fit the classification in football idea. In other soccer related classification schemes like[9][17] the way of handling the raw data with the application of features is similar, using measures and norms to cope with the three dimensional nature of acceleration and gyroscope signals.

Conclusion

The highest predictive accuracy was achieved by Naïve Bayes and Support Vector Machines with a percentage of 97.8% as it is inferred by the experts[10]. This appears to be an encouraging observation since our research will also utilize those classifiers in a frequent manner.

2.3.2. Classification of human physical activity

Introduction

This section discusses a study[13] that was conducted with the goal of classifying data that are derived from a bi-axial accelerometer based approach. This study concerns a more general vision on the recognition and detection of human motion. The actions examined are the following: sitting, lying, standing, walking, stair climbing, running, cycling[13]. In this particular case a window segment approach is applied in order to isolate the utilizable data that are relevant to the activities from the noise. There was a 50% overlapping of the sliding windows with 512 samples[13]. The features used were extracted based on that approach. It is clearly implied by the experts that these windows have ”a finite and constant width”[13] and also each window lasted 6.7 seconds with every new window available every 3.35 seconds[13] as also indicated by the researchers. Regarding the features, the Euclidean distance of the signals is considered as one of the dominant ways of refining the data from the signals. In Figure 2.4 the results of the classification are presented by giving the accuracy of every individual classifier trained for the purpose of the study[13]. The classification methods are the Naïve Bayes (NB), Gaussian Mixture Model (GMM), Logistic, Parzen, Support Vector Machines (SVM), Nearest Mean (NM), k-Nearest Neighbors (k-NN), Multilayer Perceptron (ANN), Binary Decision Tree (C4.5).

Classifiers	Classification accuracy, [%]
NB	97.4
GMM	92.2
Logistic	94.0
Parzen	92.7
SVM	97.8
NM	98.5
k-NN	98.3
ANN	96.1
C4.5	93.0

Figure 2.4: Accuracy of classifiers[13]

Connection to current project

We can see that the research [13] has many common characteristics with our own project meaning the use of time windows for the isolation of the useful activities, the application of features on the data and the use of the Euclidean distance to refine the signals.

Conclusion

An obvious but very interesting observation from the Figure 2.4 is that all of the classifiers achieved a very high amount of precision. The SVM classifier achieved one of the most efficient results (97.8%) and it is enlightening to see that the Naive Bayes scored an accuracy of 97.4%. By correlating this study[13] with our project we can infer that the SVM and NB classifiers can prove to be stable and robust candidates.

2.3.3. Tennis Stroke Recognition

Introduction

This study[12] is focused on detection and classification of acceleration data related to the different strikes that can happen in a tennis game. The IMU sensors are attached to the subjects in order to acquire the desired signals that will yield the data set for the analysis[12]. This specific research[12] focused on the Support Vector Machine classifier to derive the desired results. A window based approach was applied in order to isolate and distinguish between the data that concern the strokes, and avoid noisy data irrelevant to the activities of interest[12]. The author mentions that a time interval of 1 second is enough to include the useful information that represent the stroke[12]. The Euclidean norm is used as the main feature in order to handle the multidimensionality of the acceleration signals, as the acceleration is measured in the 3D space. As a result, a vector representative to the striking activities was derived.

Connection to current project

From this research[12] we see the approach of time windows for the extraction of the activities, which is a trait common to our research. The same holds for the use of the Euclidean Norm to derive a single vector from the 3D space signals.

Evaluation of SVM and conclusion

The evaluation of prediction in this study[12] was achieved in three steps. The first step is the consideration of individual SVM classifications for the sensors of the forearm and the upper arm with predictive precisions of 0.69 and 0.70 respectively[12]. Then the methods are not applied to the training data set but to an extended training data set, with increased precisions of 0.75 and 0.79[12] for each of the sensors. The third step combined the expanded data sets of two sensors, and the SVM produced precision of 0.79[12]. These results consist an additional strong indication that the SVM classifier performs very effectively in an environment that includes activities of explosive fast motions as in soccer.

2.3.4. Classification during a soccer match

Introduction

Here we present a research[18] of nature very similar to the current project. This research[18] concerns Classification during a Soccer match using IMU sensor data. The goal was to classify two activities, meaning the shot and the pass[18], but there was an additional activity called "other"[18] that included actions irrelevant to the pass and shot like "tackling"[18], "fast running"[18] and "side steps"[18]. The first goal of the research was to detect the peaks of their data[18]. Then a window segmentation was applied to the data based on the peaks of the previous step[18]. We also mention that the specific research used features like the mean, variance, skewness and kurtosis[18]. The classification methods utilized by the researchers of the project are the Support Vector Machines (SVM), Classification and Regression Tree (CART) and the Naive Bayes (NB)[18].

Connection to current project

Comparing this research[18] with the current project, we see that the window segmentation is a process that was also used in our study. We also see the use of feature extraction which is also applied in our project.

Conclusion

An enlightening observation is the high effectiveness of the three classification methods SVM (99.9%), CART (99.1%), NB (98.5%)[18]. These are the highest rates achieved by those methods. The SVM and NB methods will also be included in our analysis and we receive an important indication of their high level strength on action recognition.

3
Methodology

3.1. Machine Learning

Machine Learning is a term that encompasses a broad variety of algorithms that can be applied to a data set[19][3]. These data that train algorithms make it possible for the algorithms to be applied in new data sets even though they may have entirely different practical applications. Some significant use of the mentioned algorithms is prediction and pattern recognition. The broad term of Machine Learning can be classified into Supervised and Unsupervised learning [3], terms that are discussed in the next paragraph. The main focus of this research is the application of Supervised Learning models but it is useful to mention that Unsupervised learning includes methods that are looking for structure in data sets without having any prior reference or knowledge behind it[1].

3.1.1. Supervised learning

Supervised Learning is the term for the process of building a predictive model that can map a set of input variables to a response. In contrast to Unsupervised learning, in this case we have previous information about the specific class each observation is assigned to.

3.1.2. Introduction to the classification methods

The primary direction of this project from a researchers point of view, is to be able to find and handle the most appropriate supervised learning algorithms that are related to the classification of the soccer based activities. The classification methods that are frequently utilized are the Naïve-Bayes, K-Nearest Neighbors, Multiclass Support Vector Machines, Discriminant Analysis, Decision Trees. The mentioned methods are the fundamental part of supervised learning.

3.1.3. Methodology for classification

K-nearest neighbors

In order to classify a new observation, we assign it to the same category as each closest observation which is called nearest neighbour[1][3]. This method is sensitive to known data which may not be separated very cleanly, also it is very sensitive to outliers. A way to reduce the mentioned weakness of this method is to use many different neighbours. A practical way can be to give weights to the neighbours[1]. Then the closer neighbours have a heavier impact compared to the ones further away, and as a result the noise created by the outliers is limited [8][1]. The K-nearest neighbor method is available in Matlab as the `fitcknn` function[1]. Figure 3.1 shows the details.

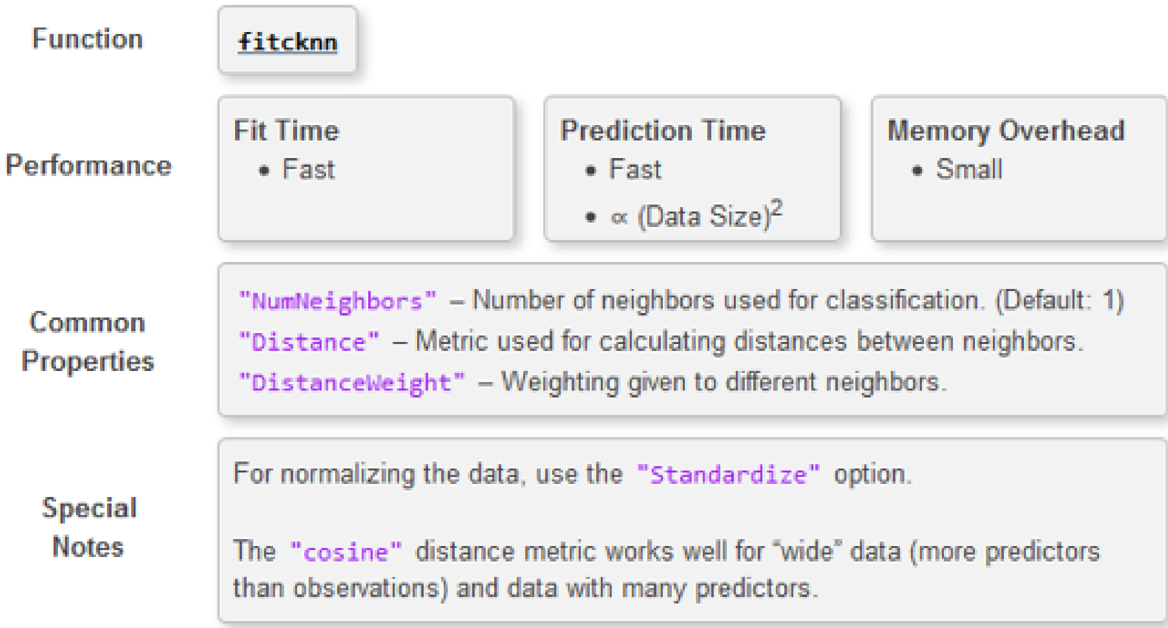

Figure 3.1: Schematic diagram of K-nearest neighbors[1]

Decision Trees

This method considers all the possible splits in each variable and chooses the most efficient way of splitting[1][15]. The process repeats at every next level of the tree, and this continues when all the branches terminate, which happens when no more splits can improve the partition[1]. This method may overfit the data which means that it will be perfectly adapted to the training data, but predictions for new data will be very weak[1]. This implies great classification error. The specific problem can be fixed by pruning the tree and reducing the number of splits, to create a model with higher resubstitution loss (training data misclassification) but better generalization to the new data[15]. The Decision Trees method is available in Matlab as the `fitctree`[1]. Figure 3.2 shows the details.

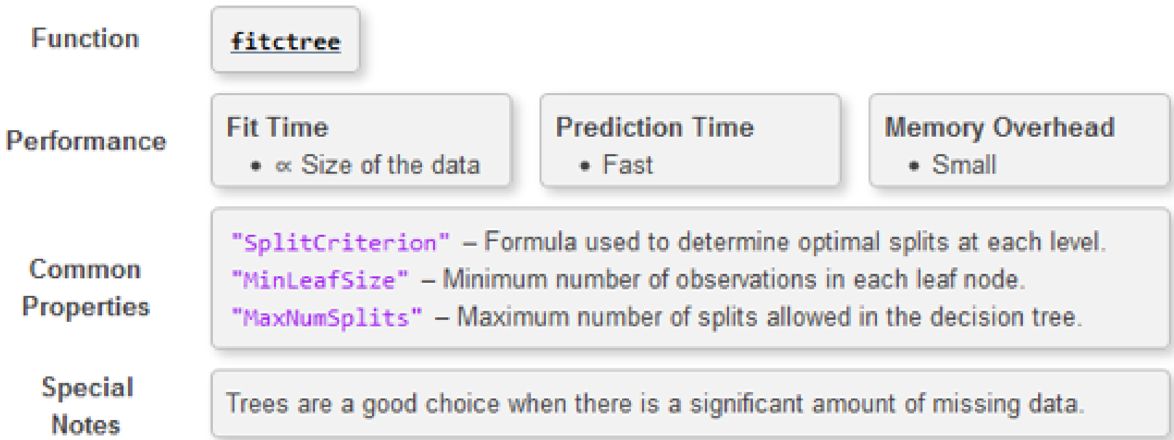

Figure 3.2: Schematic diagram of Decision Trees[1]

Naive Bayes

The observations in each response class are samples from a probability distribution for each class[1][3]. By determining the probability of a new observation belonging to a given class, we manage to assign the observation to the class that it is most likely to belong based on the probabilities[1]. This particular method is based on the hypothesis of independency. The predictions depend on the statistical distribution of all the data and the method is robust to noise from training data [20]. The Naïve Bayes method is available in Matlab as the `fitcnb` function[1]. Figure 3.3 shows the details.

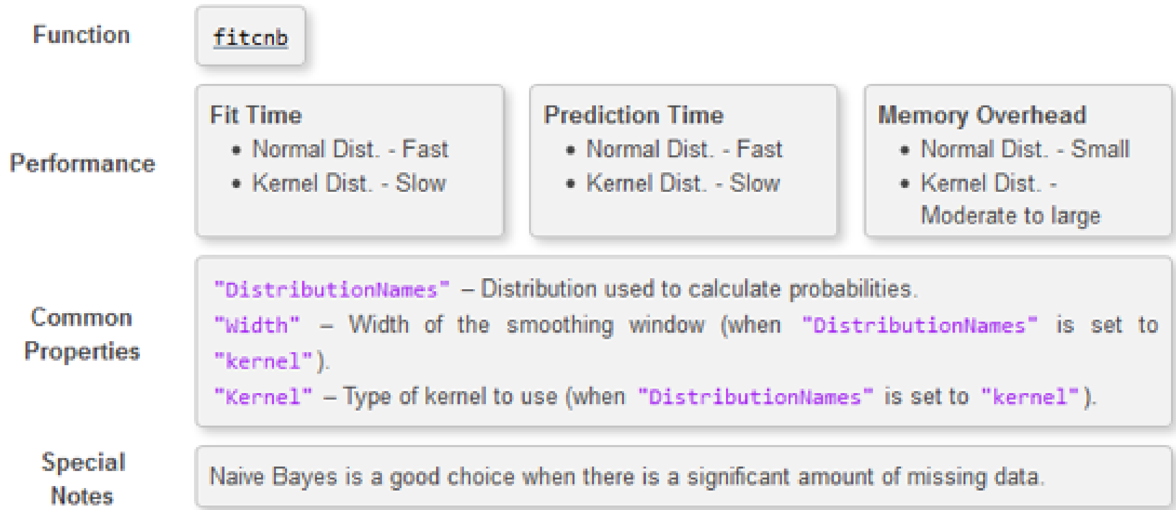

Figure 3.3: Schematic diagram of Naïve Bayes[1]

Multiclass Support Vector Machines

This method is mainly used for the binary class case[1][3]. It classifies the data by setting linear boundaries between them [1][3]. The goal is to maximize the margin between each class, which means that each class will be as cleanly separated from the others as possible, and the penalty for misclassification will be very small[1]. The Support Vector Machines(SVM) method can be used for linear but also nonlinear classification problems by performing a transformation of variables into a space where the classes are linearly separable[1]. Even though the SVM method is utilized for the binary case there are ways to apply it for the multiclass case [4]. There are three different ways to approach SVM for the multiclass case. The One Versus All (OVA) approach, the One versus One (OVO) approach and the Error Correcting Output Code (ECOC) approach[4][22]. The Multiclass Support Vector Machines method is available in Matlab as the `fitcsvm` function[1]. Figure 3.5 shows the details.

OVO

In the One Versus One case, we consider all the possible pairs for all the existing K classes and that is why we end up receiving K(K-1)/2 binary classification problems. We then assign one classifier to each of the pairs. For every pair of classes we use a subset of our data to have a training set for every classifier. After we train our classifiers the data points are labelled as positive or negative for the identification. Each of the K(K-1)/2 classifiers has its own result of a class. In conclusion we select the class that is chosen with the biggest frequency between the classifiers[22].

OVA

In the One Versus All method we have K binary classification problems for a K-class problem. Each time we choose one of the individual classes as one class and all the remaining classes together as the second class, and we train a classifier for each case. In this case we use all of our training data to train every classifier. The data points in the mentioned binary classifications that belong to the individual class are labelled as positive while the other points of the combined classes are labelled as negative. Through this approach we base ourselves to our degree of belief of each of the positive results indicated by the classifiers and we choose the class with the largest confidence. In case of only one positive result, it is selected as the correct class[4].

ECOC

In the Error Correcting Output Codes we consider K classes and N binary classifiers. First of all every classifier will be trained in a binary classification problem by considering one individual class against a class consisted by the combination of all the remaining classes, meaning that the one multiclass problem is converted to K binary classification problems. For every data point predicted to be in a certain class we give a positive mark to the class it is assigned to by the respective classifier. Negative marks will be given to the other classes for the respective classifier case. At the end, for a new data point we assign a positive mark if it belongs to the class marked with positive for the respective classifier and negative otherwise. We measure the distance between the predictions (+ or -) of the classifiers for the specific point and the predictions for the previous data and subsequently we assign the data point to the class with the least number of mismaches[22]. On Figure 3.4 our new data is assigned to class 2.

	Classifier 1	Classifier 2	Classifier 3
Class 1	+	-	-
Class 2	-	+	-
Class 3	-	-	+
New data	-	+	-

Figure 3.4: ECOC case

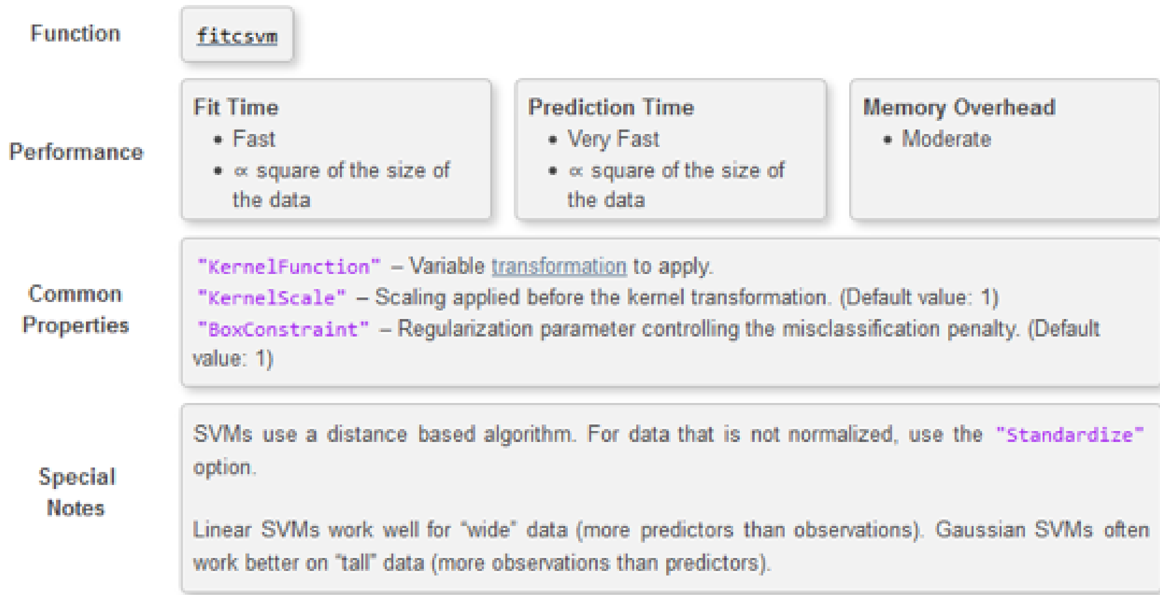

Figure 3.5: Schematic diagram of Support Vector Machines[1]

Discriminant Analysis

This technique uses the training observations to locate the boundary between the response classes[16][1]. The location of the boundary is determined by treating the observations of each class as samples from a multidimensional normal distribution[16]. If we assume the distributions have the same shape, then the boundaries turn out to be linear[16]. In case the covariance matrices are not the same for all classes the boundaries turn out to be quadratic[16]. Quadratic discriminant analysis is still relatively quick although it requires more calculations and memory to evaluate and store the multiple covariance matrices[16]. The Discriminant Analysis method is available in Matlab as the `fitcdiscr` function[1]. Figure 3.6 shows the details.

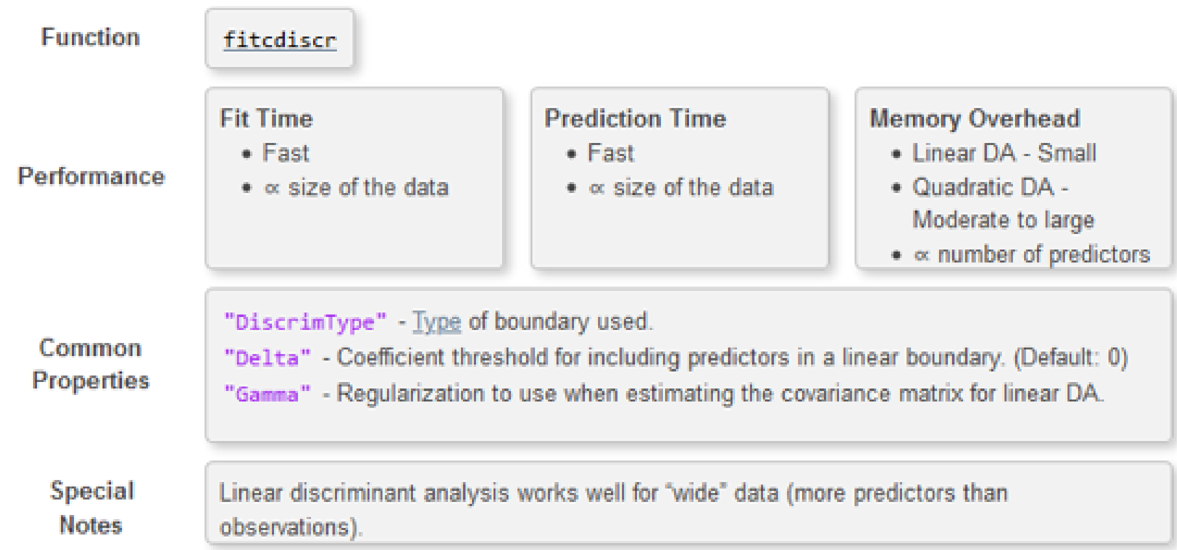

Figure 3.6: Schematic diagram of Discriminant Analysis[1]

3.2. Comparison of methods

One of the standard ways to approach the described methods is to utilize every possible approach and compare based on the minimal percentage of misclassification, but it is also considered useful to develop some prior directional line that indicates which method is preferable for our analysis.

From other studies[10][5], it seems like the Multiclass SVM method is a strong candidate. The reason being that the prediction time is substantially faster. Another strong candidate appears to be the k-nearest neighbour method since the fitting time is low and the prediction time is fast while the memory overhead remains at a low level. The Discriminant analysis could also prove to be a viable method to choose since the memory overhead seems to remain at a reasonable level in our case. Regarding the Decision Trees what seems to be relatively suspicious is that the fitting time is proportional to the size of our data table which is large enough. The Naïve-Bayes appears as a safe choice both in terms of computational speed and memory overhead. The Multiclass SVM method seems to adapt very well in the nature of our data as we have a big number of observations. We address that the default version of the methods were used, without specifying in command properties.

3.3. Goal of the analysis

The goal for the completion of this project, is the creation of a robust classifier that fits well on the training data and also adjusts to new data from well defined exercises as well as new data from field activities. That means that it will be able to classify the observations from soccer related activities with the highest possible accuracy, which means the least amount of misclassification errors.

4

Data overview and preprocessing

4.1. Data overview and experiment

This Chapter focuses on the clarification of the experiment held along with the description of the nature of the data extracted. Later on the Chapter we give the details of the preprocessing procedure held in order to acquire the final form of the data which are utilized as the input information for the models described on Chapter 3. At this point we need to mention that the data described and analyzed were provided by Mister Erik Wilmes[21]

The experiment held included 12 subjects with 5 IMU sensors attached to their pelvis, left thigh, right thigh, left shank and right shank. Each of the 12 subjects completed a number of activities ranging from 77 to 94. These activities are based on soccer related actions. The activities discussed included a variety of movements that simulate the behavior of a soccer player in a real football match. The specific actions completed are the pass, the long pass, the shot, the 90 degree turn, the 180 degree turn, the jump and the running. It is important to mention also for the next sections of our discussion that in the process of completing these actions, the subjects had some running or rest prior to the specific activity and also some running or rest afterwards. The IMU sensors captured all the information from the moment the running or rest started till the player stopped moving completely. This situation naturally created some noise to the signals produced. For a better clarification and understanding of the activities completed by the subjects, we present the definitions of some actions at the next page. We also mention that for the specific project we considered two classification approaches, the results of which are presented and discussed on Chapter 5. The first approach is classification using four activities (pass, shot, 90 degree turn, 180 degree turn). The second approach included all the mentioned actions (pass, long pass, shot, 90 degree turn, 180 degree turn, jump, running). For each of these approaches we considered 40 segmented subsets for each of the actions having a total of 160 actions for the four activity classification and 280 actions for the classification of all seven activities. At this section we also present a picture with the placement of the five sensors (right thigh, left thigh, pelvis, right shank, left shank) on the body of the subjects on Figure 4.1 and also some plots that include the Euclidean Norm of the acceleration signals from the right thigh sensor for all the seven activities on Figure 4.2. The next Section discusses the preprocessing steps for the acquired data obtained by the IMU sensors.

- Pass: The kicking of the ball usually with the side of the foot with the purpose of transferring it to a teammate of close or medium distance
- Long pass: The kicking of the ball with the purpose of transferring it to a teammate of long distance
- Shot: The strong kicking of the ball in order to score a goal through the net of the opponent team
- 90 degree turn: The change of direction of a player by completing a 90 degree maneuver
- 180 degree turn: The change of direction of a player by completing a 180 degree maneuver

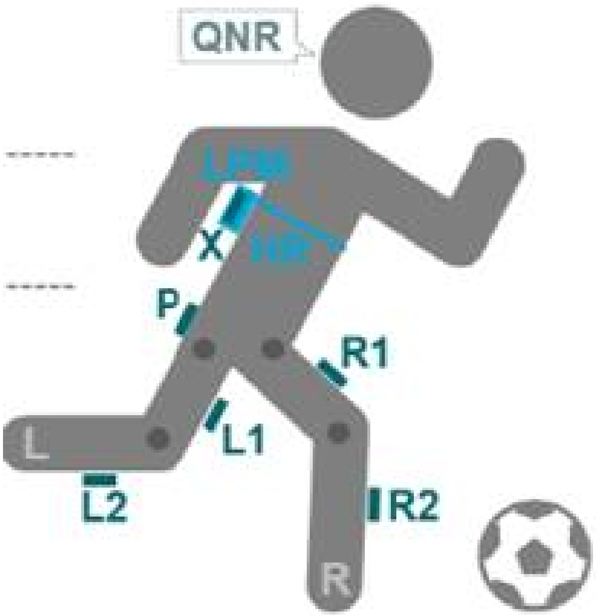

Figure 4.1: Illustration of the IMU sensors on the body of the subjects[17]

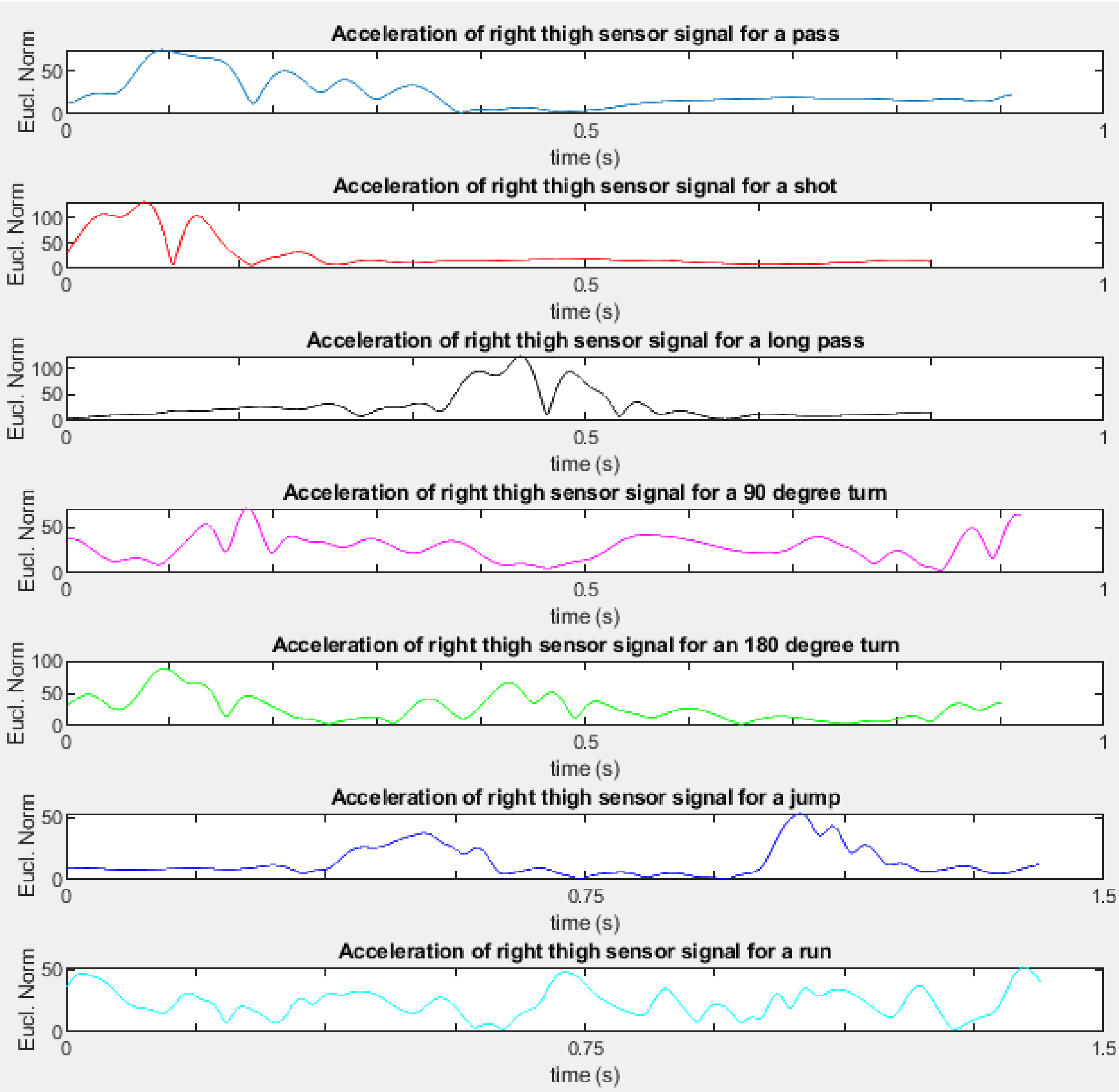

Figure 4.2: Euclidean Norms of the acceleration signals from the right thigh sensor for every action.

4.2. Preprocessing

The preprocessing techniques that are going to follow, served the purpose of creating an informative final dataset, representative of the actions presented previously. An important goal in order to achieve that is to reduce the amount of noise through the time window segmentation procedure. Also the possible calibration error in case of a wrongly placed sensor will be decreased through the use of the Euclidean Norm. The aspiration for the final filtered data set is to consist the basis for the computation of our features in order to establish the fundamentals of a precise and consistent motion recognition and classification.

4.2.1. Time window segmentation

The first step for the filtering of data is to separate the data relevant to the activities discussed from the noise. As mentioned before each of the activities completed by the subjects, included some running or moments of rest before the action started, and also after the finish of the action. The way to achieve the separation of the data relative to the activities from the data relative to the moments of rest or extra irrelevant motions, is to use specific time windows to achieve the extraction of the data that come from the action (pass, long pass, shot, 90 degree turn, 180 degree turn, jump, running). The average time windows used for each of the activities are presented in the Table 4.1. The following numbers also consist the estimation of how long each activity lasted on average. The duration of 90 degree turn is a little longer than the one of the 180 degree turn since the 90 degree turn was performed while running and in the 180 degree case the change of direction was almost instant, having the subject turn without running. The algorithms do not use the window lengths for the classification. The algorithms are applied to the features calculated on the data derived from the described intervals.

Time windows	
Action	Time duration
Pass	1.14s
Long pass	0.92s
Shot	0.71s
90 degree turn	1.01s
180 degree turn	0.92s
Jump	1.29s
Running	1.5s

Table 4.1: Average time windows applied to the data in order to isolate and extract the actions

4.2.2. Euclidean Norm and raw data

The second step after the extraction of the data relevant to the specific activities of interest, is to examine the Euclidean Norm and apply it to the filtered data. The raw data that are produced by the IMU sensors contain three signals for the coordinates of 3D space. The reason to compute the Euclidean Norm of the acquired data is to compensate for any possible small misplaces of the attached sensors on the body of the subjects. By using the Euclidean Norm we obtain a refined signal that compensates for any errors in the calibration of the sensors. Another reason to apply the specific norm is to begin our study and classification of the data using one signal for every sensor. That way we can have a sufficient indication of how the classification methods behave without involving complicated forms of data as in the raw data case.

The next step after the examination and classification based on Euclidean Norm signals, is to apply the data in their raw form. By doing that we manage to utilize all the information included in the signals, because the exact vectors for the three dimensions are used without any modification. The advantage we expect by doing that is to obtain more accurate distinction between the different activities by the methods of Chapter 3. These methods should be able to distinguish between the data patterns of the different classes of actions with higher efficiency. A drawback we expect by applying the data in their pure form, is to have additional computational costs compared to the Euclidean Norm case, since the input of the methods will be much larger. The reason being that we are going to have three signals instead of one in the Euclidean Norm case. Before we proceed to the next step of the preprocessing, we include the definition of the Euclidean Norm which will be computed based on the raw three-dimensional signals (x,y,z).

$$\sqrt{x^2 + y^2 + z^2} \tag{4.1}$$

4.2.3. Features

This section involves the specific features used in this research as the third step of the preprocessing. The features utilized will be computed on the filtered data created by the previous steps and specifically for the Euclidean Norm and the raw data of the signals belonging to the specific activities discussed. In total we have nine features. Seven of them belong to the spatial domain and two belong to the spectral domain. The reason of the specific choice is to include information relative to the distribution of the data and also relative to the domain of frequencies. For spatial features we used seven statistical measures and for the spectral domain we utilized two features based on the coefficients of the Discrete Fourier Transform. Below we present the features followed by some important definitions based on references included in this document.

Spatial domain

- `Mean`
- `Median`
- `Standard deviation`
- `Skewness`
- `Kurtosis`
- `Minimum`
- `Maximum`

Spectral domain

- `Sum of the real parts of coefficients of the Discrete Fourier Transform`
- `Maximum of the real parts of coefficients of the Discrete Fourier Transform`

Skewness

The skewness measures the asymmetry of a distribution and has values in the interval (-∞,∞)[11]. A positive skewness indicates a heavy right tail, showing the existence of a lot of extreme positive values of the distribution[11], while negative values indicate a heavy left tail[11]. Also the skewness consists the third moment of the distribution[11]. Below, the mathematical expression of skewness is presented as appears in [11] for a population of size n and mean $\bar{X}$.

$$s = \frac{\sqrt{n(n-1)}}{n-2} \frac{\frac{1}{n}\sum_i (x_i - \bar{X})^3}{(\frac{1}{n}\sum_i (x_i - \bar{X})^2)^{\frac{3}{2}}} \tag{4.2}$$

Kurtosis

The kurtosis consists the fourth moment of a distribution[11] with range of values at [1,∞)[11]. When the value of the kurtosis is below 3 the distribution "is platykurtic"[11]. When kurtosis is larger than 3 the distribution "is leptokurtic"[11]. Below, the mathematical expression of kurtosis is presented for a population of size n and mean $\bar{X}$ as appears in [11].

$$k = \frac{n(n+1)(n-1)}{(n-2)(n-3)} \frac{\sum_i (x_i - \bar{X})^4}{(\sum_i (x_i - \bar{X})^2)^2} \tag{4.3}$$

Discrete Fourier Transform

The Discrete Fourier Transform for a continuous time non periodic signal x(t)[14] is presented below. The Discrete Fourier Transform synthesis and analysis equations are given in the equations (4.4) and (4.5) respectively as defined in [14]. The synthesis equation decomposes a set of N time samples to N exponentials[14] the magnitude of which is given by the analysis equation[14].

$$x(n) = \frac{1}{N}\sum_{k=0}^{N-1} X(k) e^{j(\frac{2\pi}{N})kn} \tag{4.4}$$

$$X(k) = \sum_{n=0}^{N-1} x(n) e^{-j(\frac{2\pi}{N})kn} \tag{4.5}$$

5
Evaluation of results

5.1. Euclidean Norm case for the classification of four activities

This case involves the classification of four activities which are the short pass, the shot, the 90 degree turn and the 180 degree turn. The processing of data was done using the Euclidean norm of the raw acceleration and gyroscope signals. The next section involves some comparisons and observations regarding the quality and behavior of the methods using the graphs of Section Graphs. The methods taken into account are the `M1:Naive Bayes (NB)`, `M2:K-nearest neighbors (K-NN)`, `M3:Multiclass Support Vector Machines (MSVM)`, `M4:Discriminant Analysis (DA)`, `M5:Decision Trees (DT)`. The classification is done in the following six different settings but we focus on the method with the highest accuracy which is selected for each setting.

- `S1: Euclidean Norm of the acceleration data of the sensors using only spatial domain features`
- `S2: Euclidean Norm of the gyroscope data of the sensors using only spatial domain features`
- `S3: Euclidean Norm of the mixed acceleration and gyroscope data of the sensors using only spatial domain features`
- `S4: Euclidean Norm of the acceleration data of the sensors using spatial and spectral domain features`
- `S5: Euclidean Norm of the gyroscope data of the sensors using spatial and spectral domain features`
- `S6: Euclidean Norm of the mixed acceleration and gyroscope data of the sensors using spatial and spectral domain features`

5.1.1. Evaluation

We considered a data set consisting of 40 subsets of each action (pass, shot, 90 degree turn, 180 degree turn). For the training of the methods we used 70% of the data and 30% for the validation or testing. In order to reduce the error, each classification scheme was evaluated 100 times and only the sums of confusion matrices are shown for each scheme in the Section Graphs. Also, in the confusion matrices of the Section Graphs, we include the percentages of correctly predicted observations for each class in blue color, and the percentage of misclassifications in pink, again for the specific class. The accuracies presented in the descriptions of this section are the highest average accuracies achieved between the models for the test data over the 100 times evaluated schemes. We focus our comments and comparisons mainly on the average test accuracy because this is the precision of prediction on the test data for which we trained the model, as mentioned before. For this section, under every evaluation of a setting, we include a table with additional information including the average percentage of the accuracy for the test data and also for the training data, along with the computational time for every method in seconds. The computational time is given in three different numbers, the total time, the time for the training of the method and the time for the testing or validation. We also clarify that both the test times and train times are based on 100 runs, obtaining an improved prediction for the test accuracy and also for the train accuracy.

As shown in Table 5.1, for the classification of the acceleration data using only spatial features (setting S1), we observe a test accuracy of 92% which is achieved by the `Discriminant Analysis` method (M4). That means that the different activities are well classified. Also the small error is explained by the fact that there is a big similarity between the two kinds of turns and that can be also seen in Figure 5.1 that includes the correct predictions and the misclassifications, having a relatively large number of wrong predictions for the two turns.

S1					
Method	Test accuracy	Train accuracy	Total time	Test time	Train time
M1 (NB)	85%	90%	33s	26s	7s
M2 (K-NN)	90%	100%	31s	29s	2s
M3 (MSVM)	91%	100%	40s	31s	9s
M4 (DA)	92%	99%	34s	31s	3s
M5 (DT)	78%	96%	32s	30s	2s

Table 5.1: Table of results for the Euclidean Norm case of setting S1

In the classification of the gyroscope data using spatial features (setting S2), we observe from the Table 5.2 that the highest test accuracy of 92% was achieved by the `MSVM` method (M3) which is the best among the six settings taken into account. From Figure 5.2 we observe that the distribution of the small number of misclassified observations is the same as the previous case with the acceleration data. The setting S2 performs more effectively compared to S1 and S3 (Table 5.1 and Table 5.3) because the gyroscope data might contain identifiable patters that for example the mixed data setting S3 may not contain due to the combination of signals. This is still an issue that can be examined more in the future.

S2					
Method	Test accuracy	Train accuracy	Total time	Test time	Train time
M1 (NB)	87%	94%	38s	30s	8s
M2 (K-NN)	90%	100%	30s	28s	2s
M3 (MSVM)	92%	100%	40s	32s	8s
M4 (DA)	87%	99%	33s	30s	3s
M5 (DT)	72%	95%	32s	30s	2s

Table 5.2: Table of results for the Euclidean Norm case of setting S2

Regarding the case of the mixed data using spatial features (setting S3), based on Table 5.3 we see a test accuracy of 89% with an identical distribution of misclassifications as in our previous two cases as observed in Figure 5.3. The specific estimation was made using the `Discriminant Analysis` method (M4).

S3					
Method	Test accuracy	Train accuracy	Total time	Test time	Train time
M1 (NB)	85%	91%	34s	27s	7s
M2 (K-NN)	82%	100%	29s	27s	2s
M3 (MSVM)	85%	100%	41s	32s	9s
M4 (DA)	89%	99%	35s	32s	3s
M5 (DT)	74%	96%	34s	31s	3s

Table 5.3: Table of results for the Euclidean Norm case of setting S3

For this analysis we utilized the acceleration considering spatial and spectral domain features (setting S4). The addition of the spectral features seems to retain the test accuracy at a high level meaning 92% as seen in Table 5.4 by the `Discriminant Analysis` method (M4). Based on Figure 5.4, we also have the same distribution of misclassifications as in the previous cases. We notice that we have a decrease in the performance of the methods when adding spectral features. This is an issue that needs further investigation and it is also addressed in Chapter 6. We also observe a big decline in the accuracy of the method M3 (MSVM) an issue that also needs further research.

S4					
Method	Test accuracy	Train accuracy	Total time	Test time	Train time
M1 (NB)	85%	91%	37s	28s	9s
M2 (K-NN)	67%	100%	33s	31s	2s
M3 (MSVM)	44%	49%	13m	10m	3m
M4 (DA)	92%	100%	31s	29s	2s
M5 (DT)	77%	97%	30s	28s	2s

Table 5.4: Table of results for the Euclidean Norm case of setting S4

In this particular setting of gyroscope data using spatial and spectral features (setting S5), the `Naive Bayes` method (M1) achieved the highest test accuracy of 89%. Based on Table 5.5, we can see a similarity comparing with the previous settings where only spatial features were used, but the addition of the spectral features caused a small differentiation in the distribution of misclassified observations. This can be seen in Figure 5.5 by the fact that we not only have more misclassifications for the two turns, but we also observe a higher number of misclassifications for the ball kicking activities. This phenomenon can be explained by the fact that these are the two most commonly misplaced activities, meaning the ball kicking and the turning, since the movement and behavior of the legs of the player is very similar between a pass and a shot and also between the two different turns.

S5					
Method	Test accuracy	Train accuracy	Total time	Test time	Train time
M1 (NB)	89%	96%	37s	28s	9s
M2 (K-NN)	67%	100%	33s	31s	2s
M3 (MSVM)	79%	100%	5m	4m	1m
M4 (DA)	88%	100%	30s	28s	2s
M5 (DT)	72%	95%	32s	30s	2s

Table 5.5: Table of results for the Euclidean Norm case of setting S5

The last classification scheme is the one including the acceleration and gyroscope data along with all of the features, including the spatial and spectral ones (setting S6). The `Discriminant Analysis` method (M4) was the method that performed the best test accuracy of 88% as inferred by the Table 5.6. The distribution of misclassified observations followed a very similar pattern with the previous case, having the two turns as the main source of confusion as seen by the Figure 5.6. We also see a very low performance of the method M3 (MSVM), an issue that can be clarified through additional future research.

S6					
Method	Test accuracy	Train accuracy	Total time	Test time	Train time
M1 (NB)	85%	91%	39s	30s	9s
M2 (K-NN)	66%	100%	33s	31s	2s
M3 (MSVM)	28%	30%	14m	11m	3m
M4 (DA)	88%	99%	34s	31s	3s
M5 (DT)	73%	96%	34s	31s	3s

Table 5.6: Table of results for the Euclidean Norm case of setting S6

5.1.2. Graphs

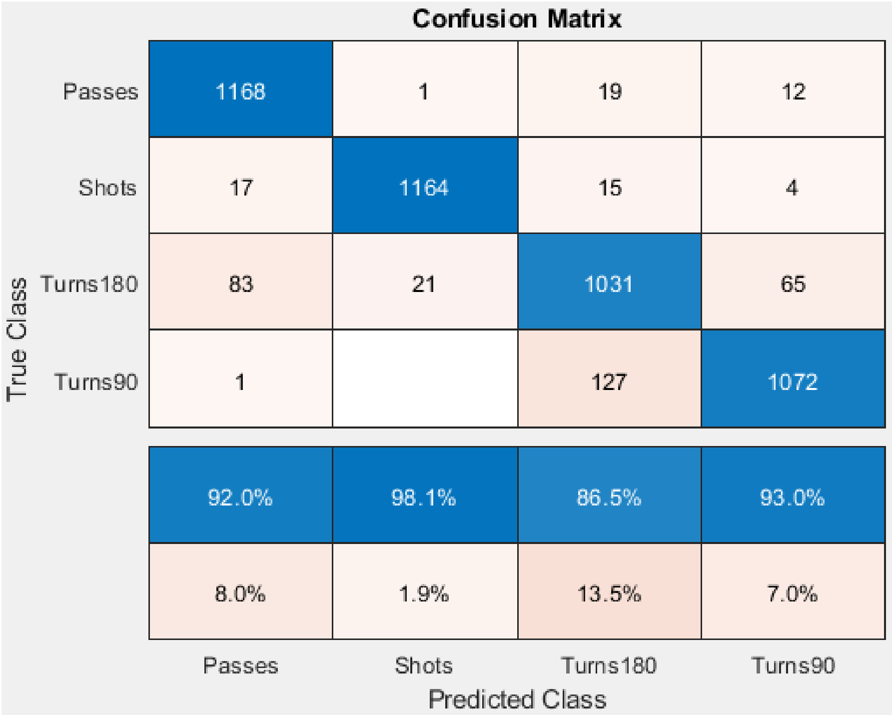

Figure 5.1: Confusion Matrix of M4 (DA) for the setting S1 with predictive accuracy of 92% for the Euclidean Norm case.

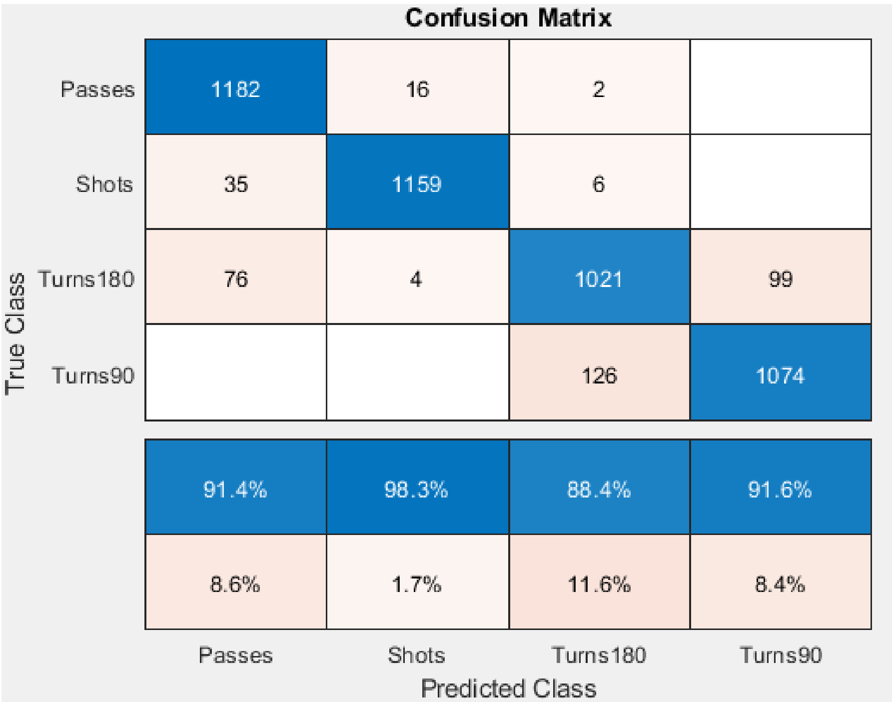

Figure 5.2: Confusion Matrix of M3 (MSVM) for the setting S2 with predictive accuracy of 92% for the Euclidean Norm case.

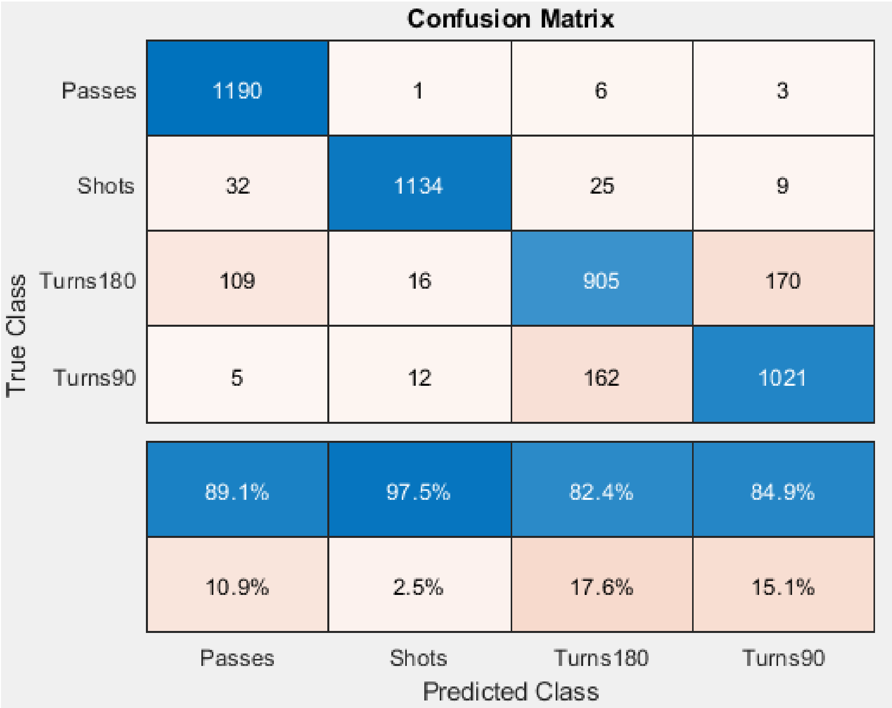

Figure 5.3: Confusion Matrix of M4 (DA) for the setting S3 with predictive accuracy of 89% for the Euclidean Norm case.

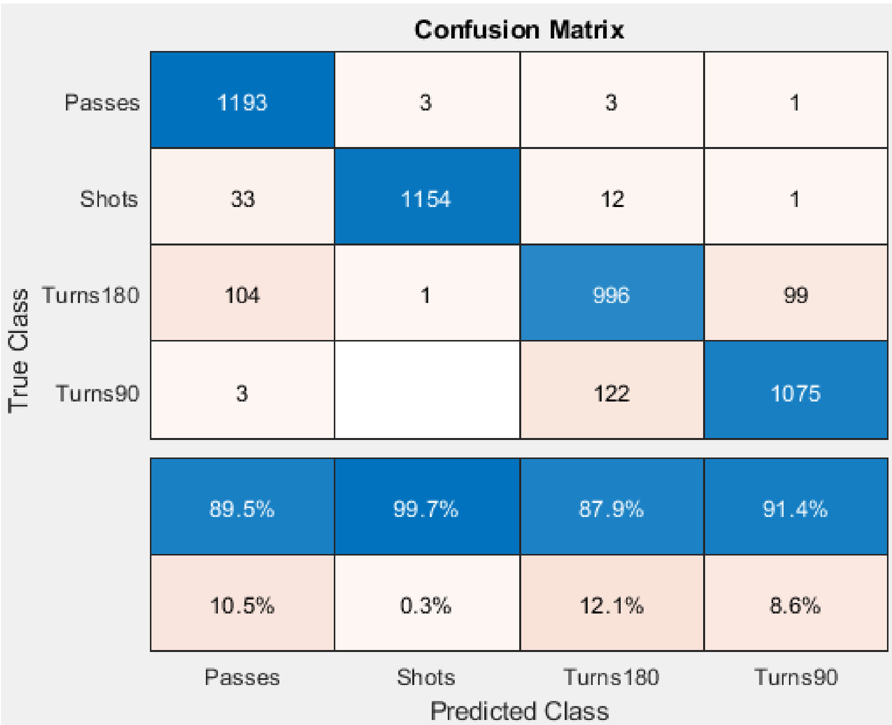

Figure 5.4: Confusion Matrix of M4 (DA) for the setting S4 with predictive accuracy of 92% for the Euclidean Norm case.

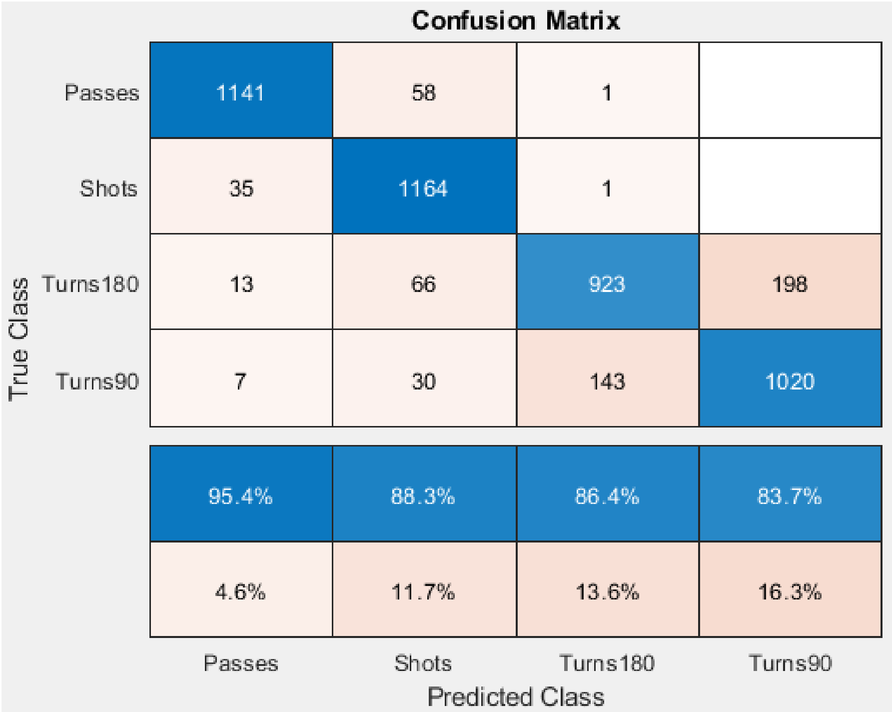

Figure 5.5: Confusion Matrix of M1 (NB) for the setting S5 with predictive accuracy of 89% for the Euclidean Norm case.

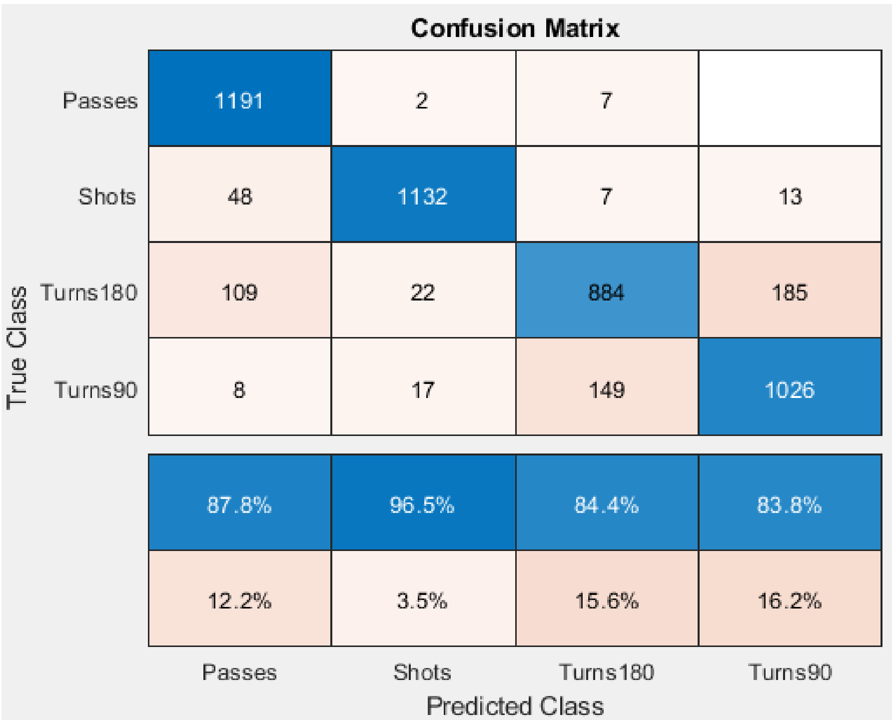

Figure 5.6: Confusion Matrix of M4 (DA) for the setting S6 with predictive accuracy of 88% for the Euclidean Norm case.

5.2. Raw data case for the classification of four activities

This case involves the classification of four activities which are the short pass, the shot, the 90 degree turn and the 180 degree turn. The data were used in their raw form which means that we had the acceleration and gyroscope signals in the form of three dimensional axis (X,Y,Z). The next section involves some comparisons and observations regarding the quality and behavior of the methods using the graphs of Section Graphs. The models that were taken into account are the `M1:Naive Bayes (NB)`, `M2:K-nearest neighbors (K-NN)`, `M3:Multiclass Support Vector Machines (MSVM)`, `M4:Discriminant Analysis (DA)`, `M5:Decision Trees (DT)`. The classification is done in the following six different settings but we focus on the method with the highest accuracy which is selected for each setting.

- `S1: Raw acceleration data of the sensors using only spatial domain features`
- `S2: Raw gyroscope data of the sensors using only spatial domain features`
- `S3: Raw mixed acceleration and gyroscope data of the sensors using only spatial domain features`
- `S4: Raw acceleration data of the sensors using spatial and spectral domain features`
- `S5: Raw gyroscope data of the sensors using spatial and spectral domain features`
- `S6: Raw mixed acceleration and gyroscope data of the sensors using spatial and spectral domain features`

5.2.1. Evaluation

We considered a data set consisting of 40 subsets of each action (pass, shot, 90 degree turn, 180 degree turn). For the training of the methods we used 70% of the data and 30% for the validation or testing. In order to reduce the error, each classification scheme was evaluated 100 times and only the sums of confusion matrices are shown for each scheme in the Section Graphs. Also, in the confusion matrices of the Section Graphs, we include the percentages of correctly predicted observations for each class in blue color, and the percentage of misclassifications in pink, again for the specific class. The accuracies presented in the descriptions of this section are the highest average accuracies achieved between the models for the test data over the 100 times evaluated schemes. For this section, under every evaluation of a setting, we include a table with additional information including the average percentage of the accuracy for the test data and also for the training data, along with the computational time for every method in seconds. The computational time is given in three different numbers, the total time, the time for the training of the method and the time for the testing or validation. Both the test times and train times are based on 100 runs, achieving an improved prediction for the test accuracy and also for the train accuracy.

For the classification of the acceleration data for the spatial features (setting S1) the highest test accuracy achieved is 94% by the `Naive Bayes` method (M1) as can be seen from the Table 5.7. We can observe from Figure 5.7 of Section Graphs that apart from the successfully classified observations, we have a very small number of misclassifications which are well explained by the fact that the two turns are very similar. As an indication, the model can confuse in some simulations the kind of turn that it is performed. The reason why the method M4 (DA) performed so poorly remains a matter that can be discussed more in future research.

S1					
Method	Test accuracy	Train accuracy	Total time	Test time	Train time
M1 (NB)	94%	97%	46s	30s	16s
M2 (K-NN)	91%	100%	40s	38s	2s
M3 (MSVM)	93%	100%	39s	31s	8s
M4 (DA)	57%	100%	33s	30s	3s
M5 (DT)	71%	96%	33s	30s	3s

Table 5.7: Table of results for the Raw Data case of setting S1

For the classification of the gyroscope data using only spatial features (setting S2), the `MSVM` (M3) model achieved the highest test accuracy of 99% as inferred by the Table 5.8. This percentage also consists the highest test accuracy achieved between the settings of the raw data case and the euclidean data case. As we can also see from Figure 5.8 we have a very small number of misclassified observations.

S2					
Method	Test accuracy	Train accuracy	Total time	Test time	Train time
M1 (NB)	97%	100%	52s	33s	19s
M2 (K-NN)	94%	100%	34s	32s	2s
M3 (MSVM)	99%	100%	40s	31s	9s
M4 (DA)	62%	100%	38s	34s	4s
M5 (DT)	80%	98%	45s	42s	3s

Table 5.8: Table of results for the Raw Data case of setting S2

For the case of mixed data (setting S3) we have again a very small number of misclassifications for the two turns as we can see from the Figure 5.9. The highest test accuracy level was achieved by the NB method (M1) being 93% as indicated in Table 5.9.

S3					
Method	Test accuracy	Train accuracy	Total time	Test time	Train time
M1 (NB)	93%	95%	53s	35s	18s
M2 (K-NN)	89%	100%	35s	33s	2s
M3 (MSVM)	92%	100%	47s	38s	9s
M4 (DA)	57%	100%	33s	30s	3s
M5 (DT)	72%	96%	38s	35s	3s

Table 5.9: Table of results for the Raw Data case of setting S3

For the classification of the acceleration data using the spatial and spectral domain features (setting S4) we get a similar misclassification rate with the previous cases. The two turns are the main activities that create a very small confusion to the model, and that is the main observation that can be made for this particular scheme based on Figure 5.10. The largest test accuracy was achieved by the NB method (M1) being 94% as can be inferred from Table 5.10. In this case the MSVM (M3) model is not included since the running time was extremely long. The process was stopped after 30 minutes of running time. Moreover we can see that the method M2 (K-NN) provided a very low accuracy and this is a matter of further future discussion.

S4					
Method	Test accuracy	Train accuracy	Total time	Test time	Train time
M1 (NB)	94%	97%	59s	36s	23s
M2 (K-NN)	63%	100%	37s	35s	2s
M3 (MSVM)	-	-	>30m	-	-
M4 (DA)	75%	100%	37s	34s	3s
M5 (DT)	71%	96%	33s	30s	3s

Table 5.10: Table of results for the Raw Data case of setting S4

The current scheme involves the classification of the gyroscope data utilizing the spatial and spectral features (setting S5). In this case we can observe a minimal misclassification rate only for the error of misinterpretation of some passes as a shot, while the two turns have a slightly better prediction rate judging by the Figure 5.11. The performance of NB method (M1) approximates the highest test accuracy at 97% between the models as indicated by the Table 5.11. In this case the MSVM method (M3) is also skipped, since the running time was extremely long. The running of the code was terminated after a period of 30 minutes.

S5					
Method	Test accuracy	Train accuracy	Total time	Test time	Train time
M1 (NB)	97%	100%	54s	32s	22s
M2 (K-NN)	66%	100%	38s	35s	3s
M3 (MSVM)	-	-	>30m	-	-
M4 (DA)	82%	100%	38s	34s	4s
M5 (DT)	83%	98%	37s	34s	3s

Table 5.11: Table of results for the Raw Data case of setting S5

In this case we use the acceleration and gyroscope data with the spatial and spectral features (setting S6). The phenomenon of confusion between the two turns can be observed in this setting as in the previous ones. Additionally a very small amount of misclassification is seen, as some passes were misinterpreted as a shot based on Figure 5.12. The observations made for this scheme agree with the distribution of misclassifications of the previous ones. The highest test accuracy rate between the models reaches 92% for the NB method (M1) as indicated in Table 5.12. For this case, the MSVM method (M3) is also omitted since the computational time was very long. As in the previous cases, the running of the script was stopped approximately after a period of 30 minutes.

S6					
Method	Test accuracy	Train accuracy	Total time	Test time	Train time
M1 (NB)	92%	96%	57s	35s	22s
M2 (K-NN)	59%	100%	37s	35s	2s
M3 (MSVM)	-	-	>30m	-	-
M4 (DA)	79%	100%	38s	34s	4s
M5 (DT)	70%	96%	37s	34s	3s

Table 5.12: Table of results for the Raw Data case of setting S6

5.2.2. Graphs

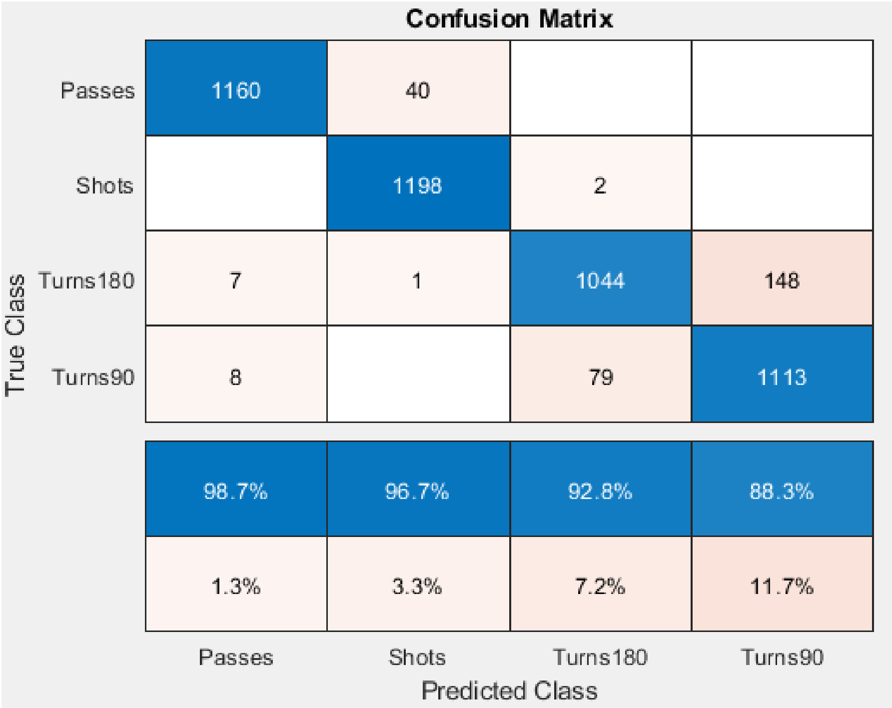

Figure 5.7: Confusion Matrix of M1 for the setting S1 with predictive accuracy of 94% for the Raw Data case.

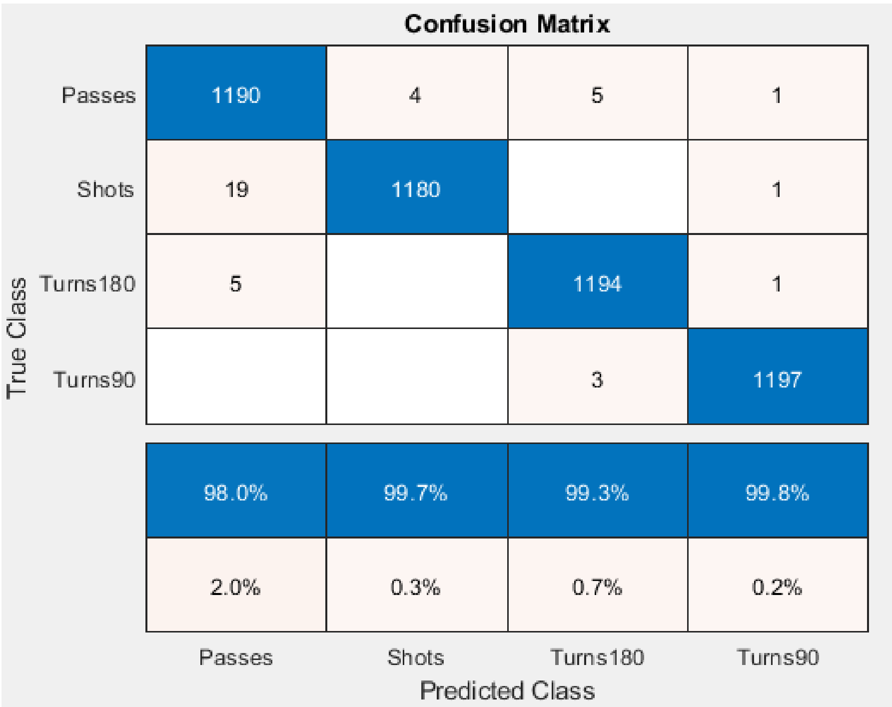

Figure 5.8: Confusion Matrix of M3 for the setting S2 with predictive accuracy of 99% for the Raw Data case.

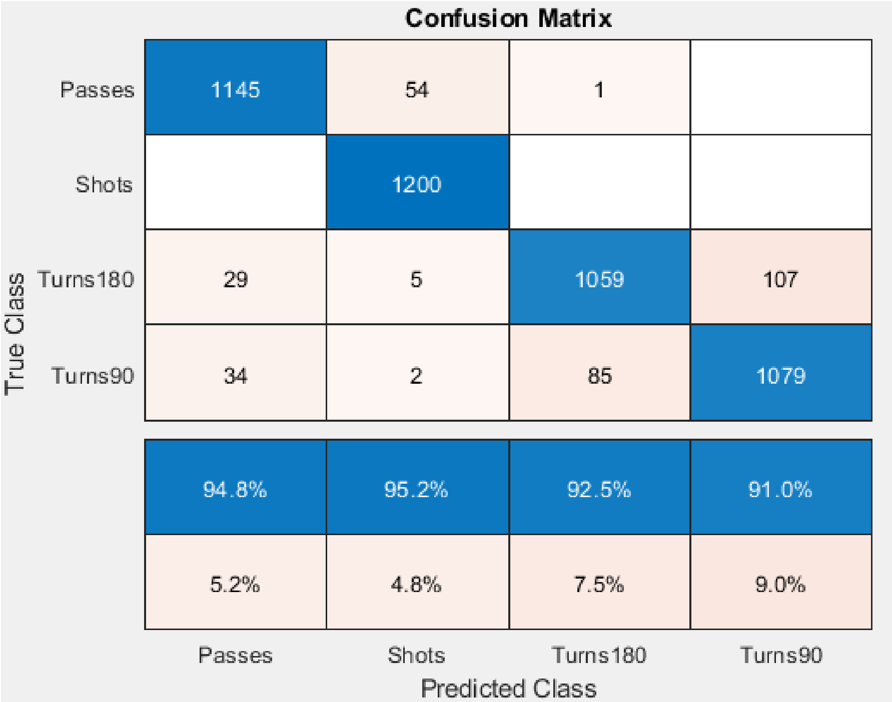

Figure 5.9: Confusion Matrix of M1 for the setting S3 with predictive accuracy of 93% for the Raw Data case.

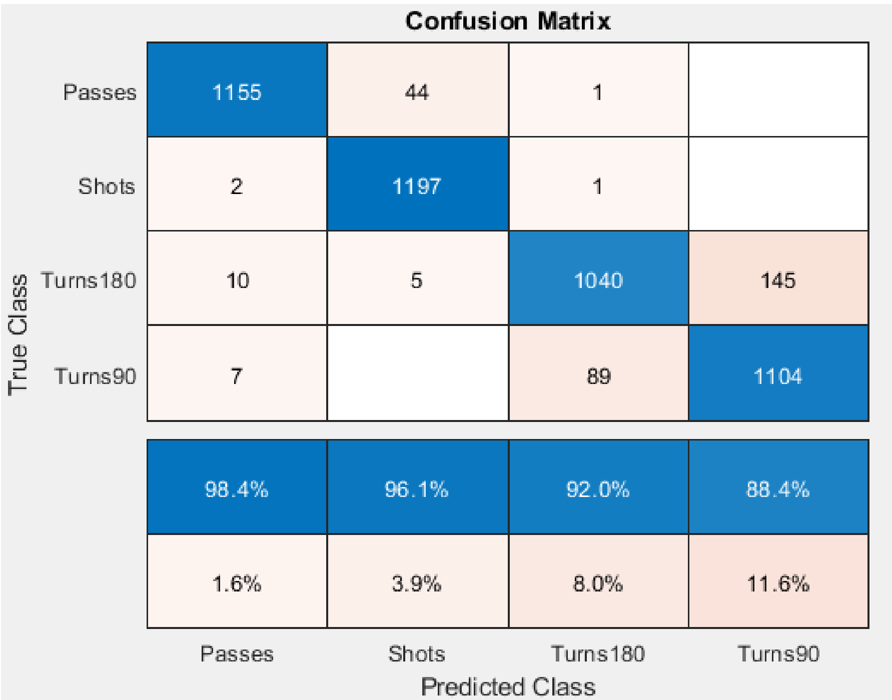

Figure 5.10: Confusion Matrix of M1 for the setting S4 with predictive accuracy of 94% for the Raw Data case.

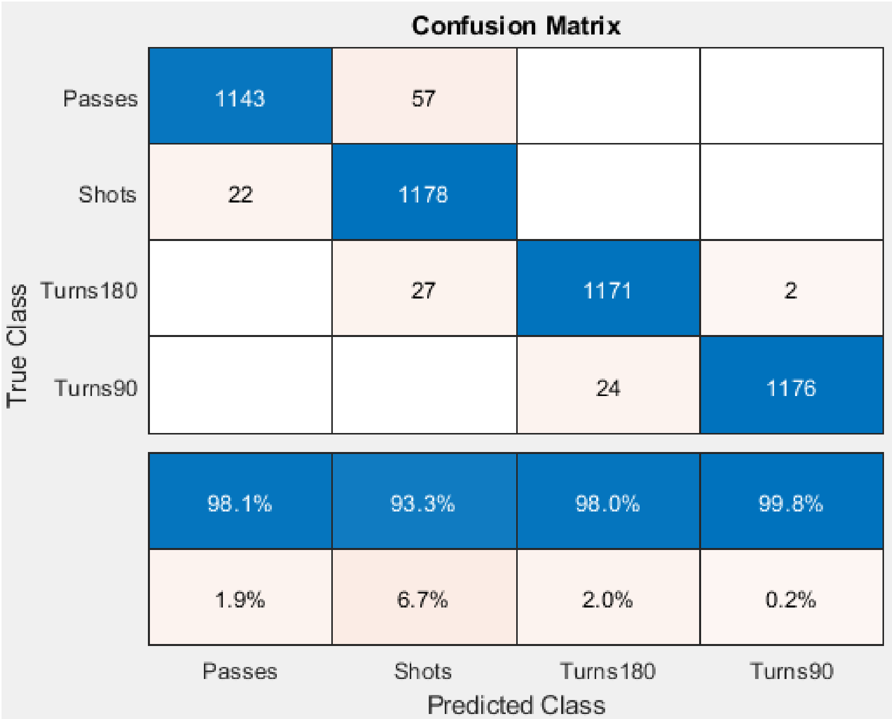

Figure 5.11: Confusion Matrix of M1 for the setting S5 with predictive accuracy of 97% for the Raw Data case.

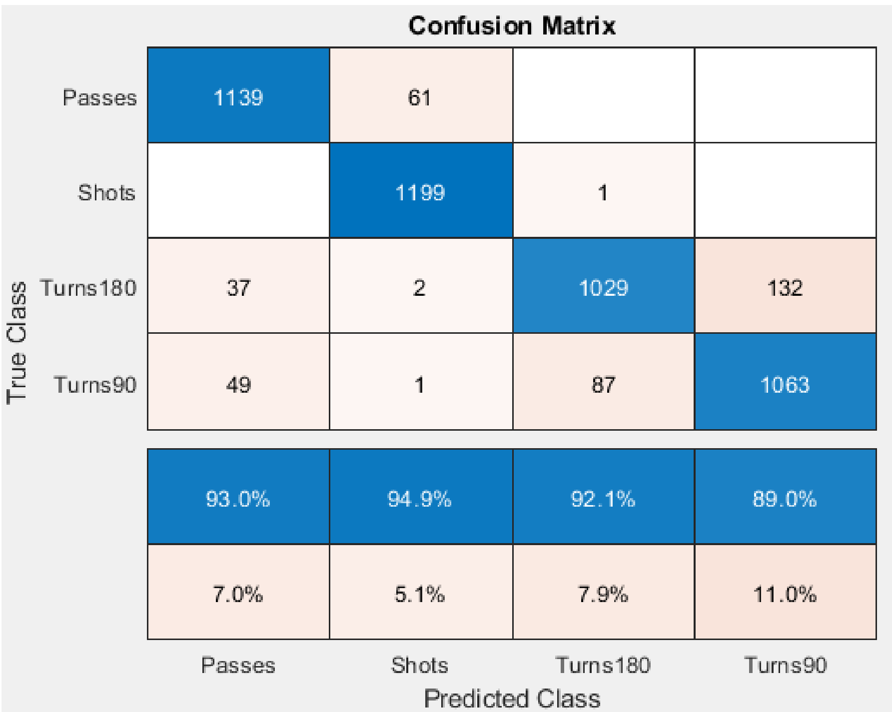

Figure 5.12: Confusion Matrix of M1 for the setting S6 with predictive accuracy of 92% for the Raw Data case.

5.3. Comparison between the Euclidean Norm case and the raw data case

First of all we compare the most important information of the relative tables and figures for the setting S1. We can observe that the maximum test accuracy is achieved by the method M4 (DA) being 92% for the Euclidean Norm case with total computational time of 34s. For the raw data case the highest percentage is produced by the method M1 (NB) and reaches 94% with total computational time of 46s. We observe that the test accuracy is improved by a small fraction of 2% for the raw data case, since the three dimensional form of the acceleration signals can provide increased information and therefore can help the model to capture the patterns of the data for each activity. A similar pattern we observe for the two confusion matrices of Figures 5.1 and 5.7, having a similar distribution of misclassifications. Because of the larger amount of processing for the raw data, the model is 12s slower.

In this case we have the comparison for the setting S2. The largest test accuracy is achieved by the method M3 (MSVM) and reaches 92% with total computational time of 40s for the Euclidean Norm case. For the raw data case the perfect test accuracy is produced again by the method M3 (MSVM) reaching 99% with a total computational time of 40s. We observe that the model behaves more precisely for the raw data providing an extra test accuracy of 7% having almost no misclassifications, while the Euclidean Norm case reaches a lower but still very high percentage of accuracy as inferred by Figures 5.2 and 5.8. This can be explained again by the fact that the raw data provide some additional information for the identification of the classes by the model.

Now we compare the differences between the two cases for the setting S3. For the Euclidean Norm case we obtain the largest test accuracy by the method M4 (DA) with a percentage of 89% and total computational time of 35s. The raw data case scores the maximum test accuracy at 93% with the method M1 (NB) with total computational time of 53s. We have an additional accuracy of 4% for the raw data case. A similar distribution of misclassifications can be observed on the two relative confusion matrices of Figures 5.3 and 5.9, as the resemblance in the performance of the two turns is the source of small confusion in the identification of classes by the models. Also the normal increase in time duration (18s) for the model of the raw data is observed, due to the increased processing time.

Here follows the comparison for the setting S4. The highest test accuracy of 92% is achieved by the model M4 (DA) for the Euclidean data case with total computational time of 31s. The maximum test accuracy for the raw data case is 94% and it is obtained by the model M1 (NB) with total computational time of 59s. The extra accuracy for the raw data is 2% for this case. We observe a similar distribution of misclassifications because of the similarity between the two turns based on the relative confusion matrices of Figures 5.4 and 5.10. We also see the difference in computational times, having the raw data case slower by 28s due to the nature of the raw data.

The next comparison concerns the setting S5. The biggest test accuracy obtained is 89% by the method M1 (NB) for the Euclidean data case with total computational time of 37s. The relative test accuracy of the raw data case is 97% and is achieved by the method M1 (NB) with total computational time of 54s. The added accuracy of 8% is observed in favor of the raw data. We can explain the small prediction error in the Figures 5.5 and 5.11, since most of the confusion happened between the two similar turns and also between the two similar ball kicking activities. As far as the raw data case, the prediction is almost perfect having mainly a small number of misinterpretations between the two ball kicking activities. The success of the raw data case can be attributed to the increased amount of information by the three dimensional form of the data. Moreover the difference in computational time is 17s more for the raw data case since it deals with more descriptive data.

Regarding the setting S6 we have the largest test accuracy of 88% by the method M4 (DA) with a total computational time of 34s for the Euclidean data case. For the raw data case we obtain the highest test accuracy with the method M1 (NB) having 92% with 57s of total running time. The improvement of prediction is 4% for the raw data case. We can infer in this case also, that the more detailed information by the raw data contributed to a better precision but with the cost of 23s of time duration. Furthermore, based on the Figures 5.6 and 5.12 we see the same pattern of misclassifications compared to the previous cases, meaning that the two turns are more likely to be confused by the relative model.

5.4. General discussion of settings for all cases

We can observe that in the Euclidean Norm case the method M4 (DA) holds the highest percentages for almost all the settings (S1,S3,S4,S5,S6) with the exception of setting S2 for which the method M3 (MSVM) achieves the highest test accuracy. The differentiation of the features with the addition of the spectral ones seem to not make a significant difference in the highest test accuracies for the settings of the Euclidean Norm case. For the raw data case the method M1 (NB) obtained the largest test accuracies for five settings (S1,S3,S4,S5,S6) while the most precise method for the setting S2 was the M3 (MSVM) which consists a strong common trait of the two cases discussed. For the raw data case the test accuracies are significantly higher compared to the ones for the Euclidean Norm case because of the additional information that the true three dimensional form of data possess. This additional accuracy comes with a cost of higher computational time because of the larger processing time by the methods as extensively explained in Section 5.3. The differentiation of the accuracy because of the addition of the spectral features is still very small similar to the Euclidean data case. Another observation that can be made for the Euclidean data case, is that with the addition of spectral features to the spatial ones (S4,S5,S6), the test accuracy of method M4 (DA) reaches the highest percentage for settings S4 and S6 while the M1 method (NB) follows with a similar rate while it reaches the maximal accuracy for the setting S5. This fact consists a characteristic shared by the raw data case where we have the largest accuracies for the settings with the mixed features (S4,S5,S6) with the method M1 (NB). The method M1 (NB) seems to have a significant strength of prediction when it comes to cases that included spectral domain features, while the method M3 (MSVM) seem to establish the most fitting option when it comes to classifying data that include spatial features like the ones utilized in this research. As mentioned before the overall accuracy achieved by the raw data with spatial and spectral features (S4,S5,S6) is higher compared to the same settings of the Euclidean data case. The most effective settings to use are the S1 and S2 for the Euclidean Norm case and the S2 and S5 for the raw data case, based on the experience from the previous results.

5.5. Euclidean Norm case for the classification of all seven activities

This analysis involves the classification of seven activities which are the short pass, the shot, the long pass, the 90 degree turn, the 180 degree turn, the jump and the running. The processing of data was done using the Euclidean Norm of the raw acceleration and gyroscope signals. The next section involves some comparisons and observations regarding the quality and behavior of the methods using the graphs of Section Graphs. The models that were taken into account are the `M1:Naive Bayes (NB)`, `M2:K-nearest neighbors (K-NN)`, `M3:Multiclass Support Vector Machines (MSVM)`, `M4:Discriminant Analysis (DA)`, `M5:Decision Trees (DT)`. Similar to the previous cases the classification is done in the following six different settings but focusing on the method with the highest accuracy which is selected for each setting.

- `S1: Euclidean Norm of the acceleration data of the sensors using only spatial domain features`
- `S2: Euclidean Norm of the gyroscope data of the sensors using only spatial domain features`
- `S3: Euclidean Norm of the mixed acceleration and gyroscope data of the sensors using only spatial domain features`
- `S4: Euclidean Norm of the acceleration data of the sensors using spatial and spectral domain features`
- `S5: Euclidean Norm of the gyroscope data of the sensors using spatial and spectral domain features`
- `S6: Euclidean Norm of the mixed acceleration and gyroscope data of the sensors using spatial and spectral domain features`

5.5.1. Evaluation

For this case we considered a data set consisting of 40 subsets of each action (pass, shot, long pass, 90 degree turn, 180 degree turn, jump, running). The same procedure of the previous case with the four classes was followed meaning that for the training of the methods we used 70% of the data and 30% for the validation or testing. In order to reduce the error, each classification scheme was evaluated 100 times and only the sums of confusion matrices are shown for each scheme in the Section Graphs. Also, in the confusion matrices of the Section Graphs, we include the percentages of correctly predicted observations for each class in blue color, and the percentage of misclassifications in pink, again for the specific class. The accuracies presented in the descriptions of this section are the highest average accuracies achieved between the models for the test data over the 100 times evaluated schemes. We focus our comments and comparisons mainly on the average test accuracy because this is the precision of prediction on the test data for which we trained the model, as mentioned before. For this section also, under every evaluation of a setting, we include a table with additional information including the average percentage of the accuracy for the test data and also for the training data, along with the computational time for every method in seconds. The computational time is given in three different numbers, the total time, the time for the training of the method and the time for the testing or validation. We mention that both the test times and train times are based on 100 runs, scoring an improved rate for the test accuracy and also for the train accuracy.

Based on the Table 5.13 we infer that the highest test accuracy for the setting S1 is achieved by the method M4 (DA) being 81%. Also by observing the Figure 5.13 we see a small amount of misclassification between the two turns due to the similarity of the two activities and also some confusion in the prediction of the three ball kicking activities, meaning the shot, the long pass and the regular pass. The error for the last three activities is also a natural observation since the movements completed for the ball kicking are also very similar.

S1					
Method	Test accuracy	Train accuracy	Total time	Test time	Train time
M1 (NB)	74%	81%	38s	26s	12s
M2 (K-NN)	77%	100%	30s	28s	2s
M3 (MSVM)	80%	100%	51s	31s	20s
M4 (DA)	81%	93%	31s	28s	3s
M5 (DT)	62%	93%	33s	30s	3s

Table 5.13: Table of results for the Euclidean Norm case of setting S1

The largest test accuracy for the setting S2 is 85% as seen by the Table 5.14 and it is obtained by the method M3 (MSVM). Comparing the Figure 5.14 with the Figure 5.13 we can see a very similar distribution of misclassifications with the only important difference that in the S2 case we have a smaller amount of wrongly predicted activities. The fact that the two turns and the ball kicking activities can cause the existing error can be observed in this case also. The reason why the setting S2 performs more accurately compared to S1 and S3 (Table 5.13 and Table 5.15) could be that the gyroscope data contain patterns more recognizable that for example the mixed data setting S3 do not contain due to the combination of signals. This is still an issue that requires further analysis.

S2					
Method	Test accuracy	Train accuracy	Total time	Test time	Train time
M1 (NB)	78%	87%	39s	27s	12s
M2 (K-NN)	80%	100%	30s	28s	2s
M3 (MSVM)	85%	99%	68s	37s	31s
M4 (DA)	80%	95%	31s	28s	3s
M5 (DT)	64%	92%	30s	28s	2s

Table 5.14: Table of results for the Euclidean Norm case of setting S2

For the setting S3 we obtain the largest test accuracy of 80% by the method M4 (DA) as inferred by the Table 5.15. From Figure 5.15 it is clear that the main source of confusion for the predictions of the model are the movements that share the same characteristics. The trait that is common between our current figure and the previous ones is the misinterpretation of the two turns and the ball kicking movements. Generally the model behaves in an efficiently accurate way, since apart from the small amount of misclassified observations, we see a high percentage of prediction for each individual activity on average.

S3					
Method	Test accuracy	Train accuracy	Total time	Test time	Train time
M1 (NB)	73%	81%	41s	28s	13s
M2 (K-NN)	73%	100%	30s	28s	2s
M3 (MSVM)	78%	100%	55s	31s	24s
M4 (DA)	80%	93%	31s	28s	3s
M5 (DT)	64%	93%	36s	33s	3s

Table 5.15: Table of results for the Euclidean Norm case of setting S3

For the setting S4 we obtain the highest test accuracy of 86% which is also the largest among all the settings for the Euclidean data case if we observe the Table 5.16. This percentage is achieved by the method M4 (DA). Again we can see from Figure 5.16 a small amount of misinterpretations regarding the shots, the passes and the long passes, but we also see a reduction of misclassifications between the two turns. This can be caused due to the addition of spectral features to the specific setting. The method M3 is not filled in the table due to the extreme amount of running time. The process was terminated after 40 minutes of running. An additional observation that needs further examination is that the performance of method M2 (K-NN) decreases when the spectral domain features are added.

S4					
Method	Test accuracy	Train accuracy	Total time	Test time	Train time
M1 (NB)	76%	84%	42s	28s	14s
M2 (K-NN)	63%	100%	30s	28s	2s
M3 (MSVM)	-	-	>40m	-	-
M4 (DA)	86%	96%	30s	27s	3s
M5 (DT)	67%	93%	31s	28s	3s

Table 5.16: Table of results for the Euclidean Norm case of setting S4

The setting S5 has the maximal test accuracy of 83% produced by the method M4 (DA) of the Table 5.17. The distribution of wrong predictions are shared especially between the ball kicking activities, having also a small amount of error between the two turns as inferred by the Figure 5.17. The previous observation is a pattern that can be seen in the previous figures of the other settings. The results of method M3 are missing due to the very long amount of running time. The computations were stopped after the approximate time of 40 minutes.

S5					
Method	Test accuracy	Train accuracy	Total time	Test time	Train time
M1 (NB)	79%	88%	42s	28s	14s
M2 (K-NN)	59%	100%	31s	29s	2s
M3 (MSVM)	-	-	>40m	-	-
M4 (DA)	83%	96%	33s	29s	4s
M5 (DT)	66%	93%	31s	28s	3s

Table 5.17: Table of results for the Euclidean Norm case of setting S5

For the setting S6 we receive a test accuracy reaching the 85% by the method M4 (DA), based on Table 5.18. On Figure 5.18 we see the amount of misinterpreted activities spread out mainly to the ball kicking activities and the turns, a pattern observed multiple times in our previous descriptions and figures. We can infer that our models for the six settings behaved satisfyingly for the distinction of completely different activities like the jump and the ball kicking activities, but had some small difficulties in the recognition of the classes of similar nature. The small decrease in the error of predictions for the last three settings can be attributed to the addition of the spectral features in our analysis. In this last case the method M3 is also omitted because of the large amount of computational time. Similar to the previous cases the running of the code was terminated when exceeded the duration of 40 minutes.

S6					
Method	Test accuracy	Train accuracy	Total time	Test time	Train time
M1 (NB)	76%	83%	43s	29s	14s
M2 (K-NN)	58%	100%	30s	28s	2s
M3 (MSVM)	-	-	>40m	-	-
M4 (DA)	85%	96%	31s	28s	3s
M5 (DT)	64%	93%	31s	28s	3s

Table 5.18: Table of results for the Euclidean Norm case of setting S6

5.5.2. Graphs

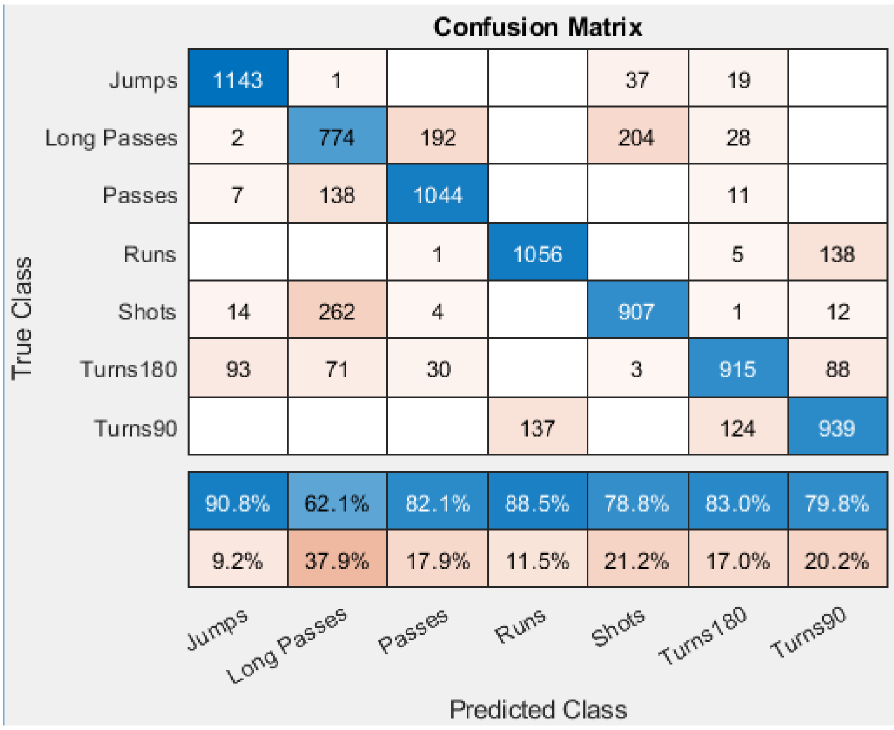

Figure 5.13: Confusion Matrix of M4 (DA) for the setting S1 with predictive accuracy of 81% for the Euclidean Norm case.

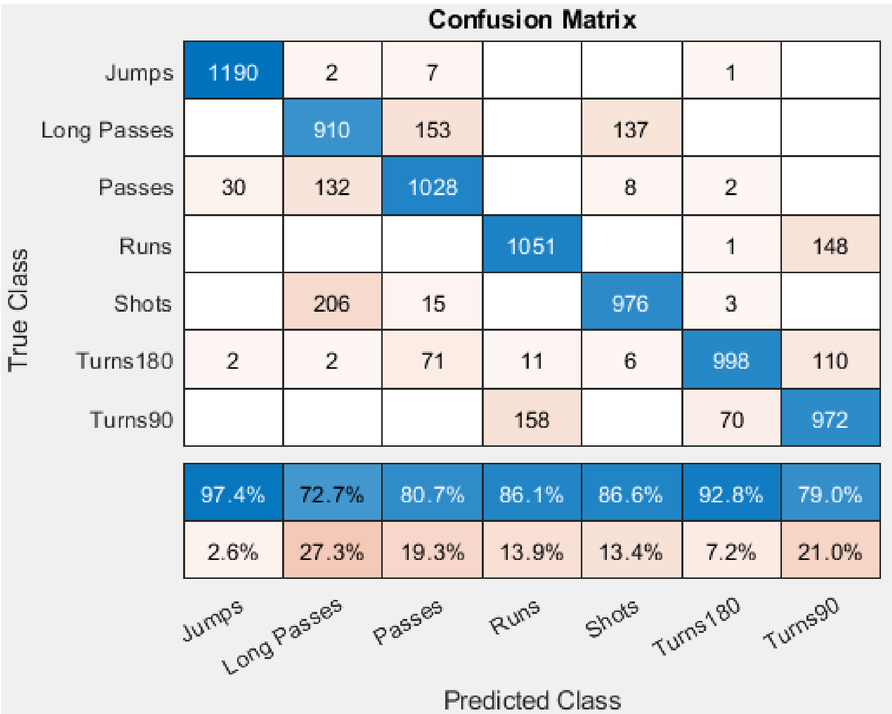

Figure 5.14: Confusion Matrix of M3 (MSVM) for the setting S2 with predictive accuracy of 85% for the Euclidean Norm case.

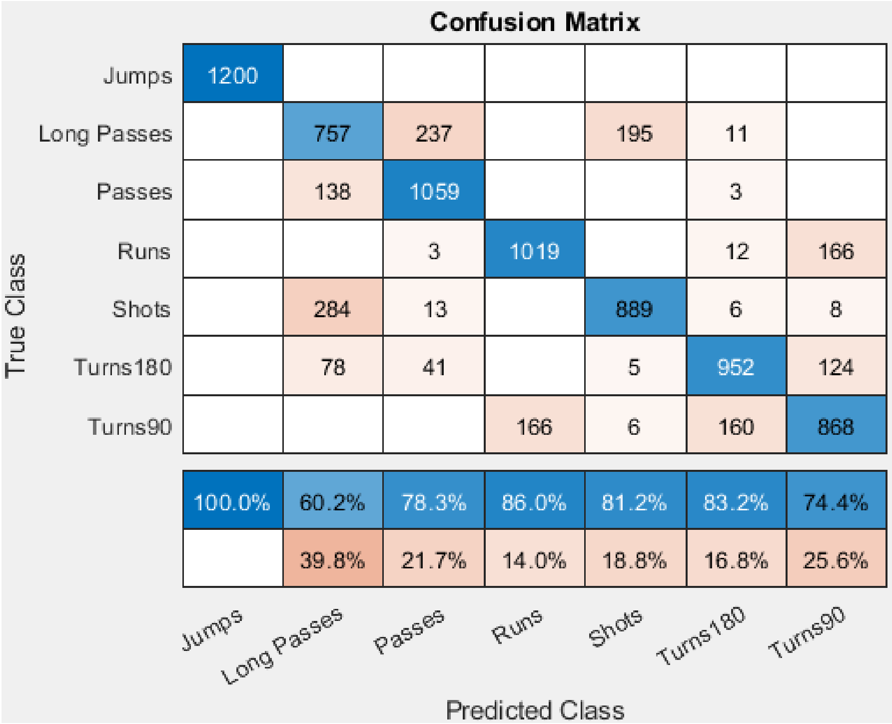

Figure 5.15: Confusion Matrix of M4 (DA) for the setting S3 with predictive accuracy of 80% for the Euclidean Norm case.

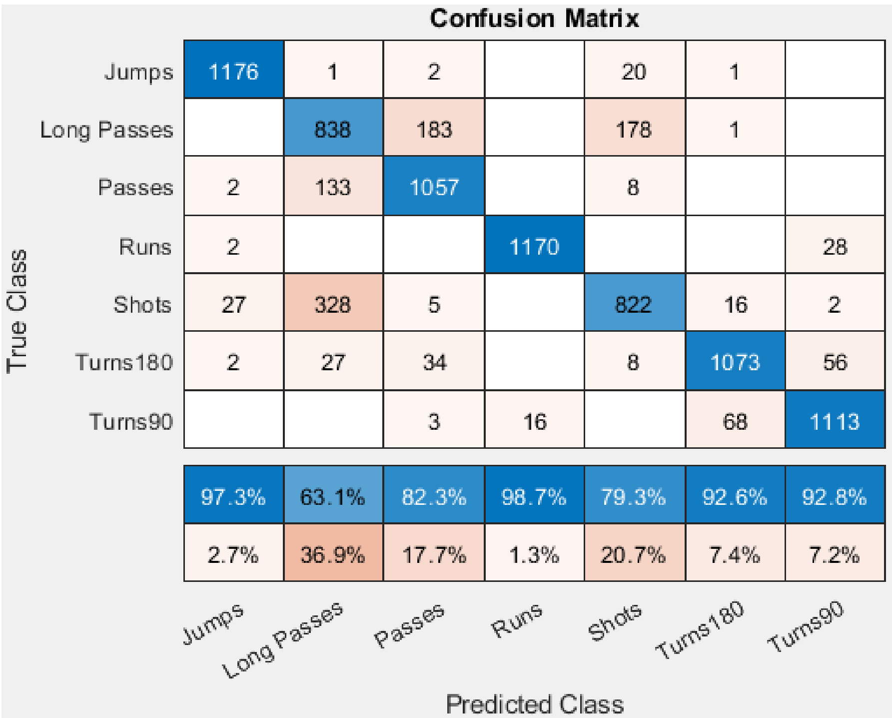

Figure 5.16: Confusion Matrix of M4 (DA) for the setting S4 with predictive accuracy of 86% for the Euclidean Norm case.

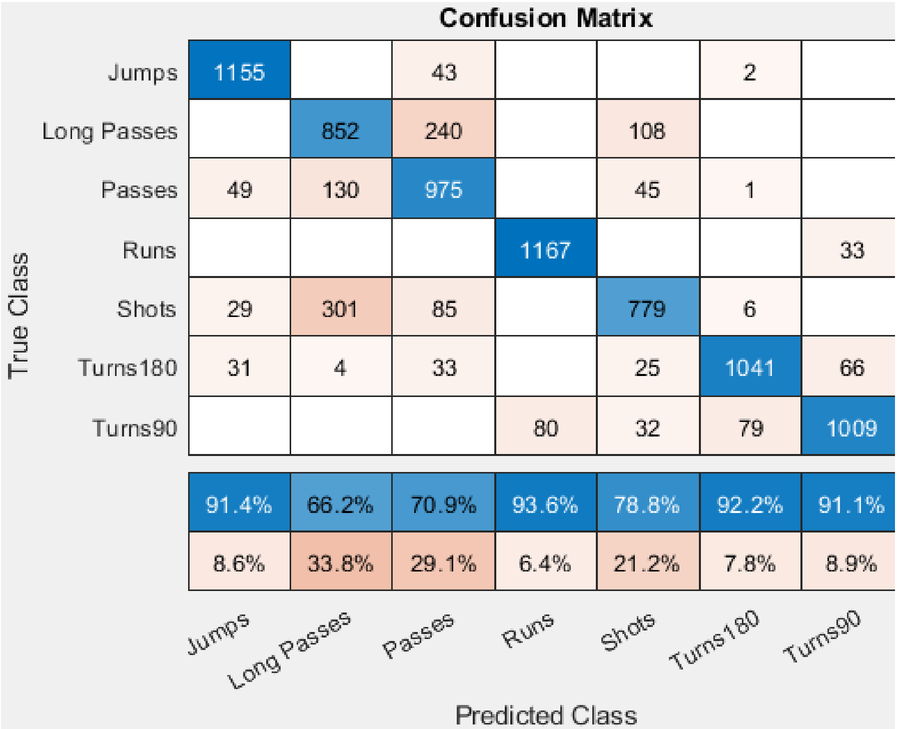

Figure 5.17: Confusion Matrix of M4 (DA) for the setting S5 with predictive accuracy of 83% for the Euclidean Norm case.

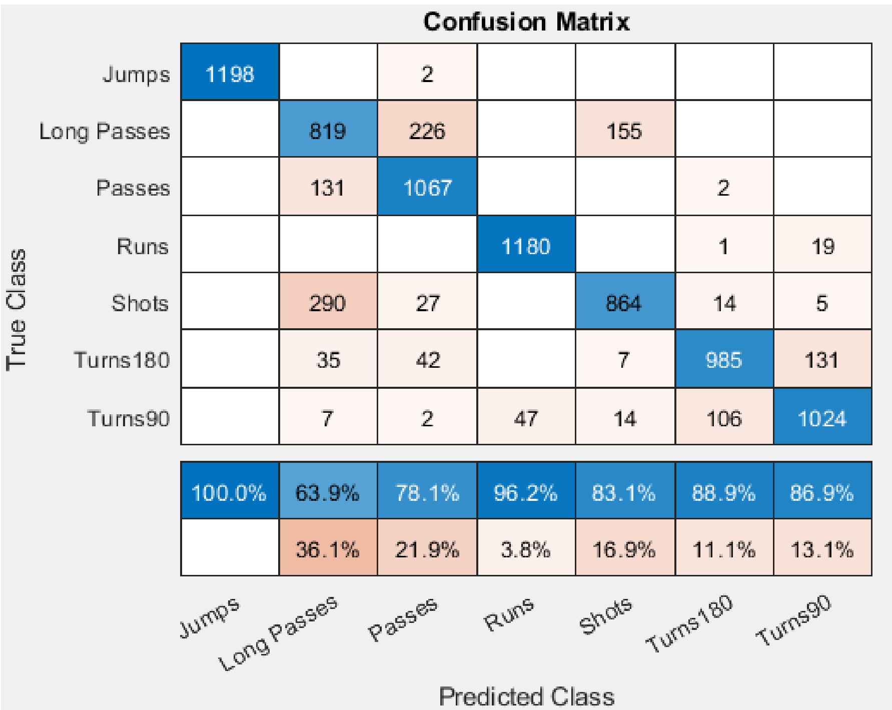

Figure 5.18: Confusion Matrix of M4 (DA) for the setting S6 with predictive accuracy of 85% for the Euclidean Norm case.

5.6. Raw data case for the classification of all seven activities

This case involves the classification of seven activities which are the short pass, the shot, the long pass, the 90 degree turn, the 180 degree turn, the jump and the running. The data were used in their raw form which means that we had the acceleration and gyroscope signals in the form of three dimensional axis (X,Y,Z). The next section involves some comparisons and observations regarding the quality and behavior of the methods using the graphs of Section Graphs. The models that were taken into account are the `M1:Naive Bayes (NB)`, `M2:K-nearest neighbors (K-NN)`, `M3:Multiclass Support Vector Machines (MSVM)`, `M4:Discriminant Analysis (DA)`, `M5:Decision Trees (DT)`. The classification is done in the following six different settings but we focus on the method with the highest accuracy which is selected for each setting.

- `S1: Raw acceleration data of the sensors using only spatial domain features`
- `S2: Raw gyroscope data of the sensors using only spatial domain features`
- `S3: Raw mixed acceleration and gyroscope data of the sensors using only spatial domain features`
- `S4: Raw acceleration data of the sensors using spatial and spectral domain features`
- `S5: Raw gyroscope data of the sensors using spatial and spectral domain features`
- `S6: Raw mixed acceleration and gyroscope data of the sensors using spatial and spectral domain features`

5.6.1. Evaluation

For this case we considered a data set consisting of 40 subsets of each action (pass, shot, long pass, 90 degree turn, 180 degree turn, jump, running). The same procedure of the previous case with the four classes was followed meaning that for the training of the methods we used 70% of the data and 30% for the validation or testing. In order to reduce the error, each classification scheme was evaluated 100 times and only the sums of confusion matrices are shown for each scheme in the Section Graphs. Also, in the confusion matrices of the Section Graphs, we include the percentages of correctly predicted observations for each class in blue color, and the percentage of misclassifications in pink, again for the specific class. The accuracies presented in the descriptions of this section are the highest average accuracies achieved between the models for the test data over the 100 times evaluated schemes. We focus our comments and comparisons mainly on the average test accuracy because this is the precision of prediction on the test data for which we trained the model, as mentioned before. For this section also, under every evaluation of a setting, we include a table with additional information including the average percentage of the accuracy for the test data and also for the training data, along with the computational time for every method in seconds. The computational time is given in three different numbers, the total time, the time for the training of the method and the time for the testing or validation. We clarify that both the test times and train times are based on 100 runs, obtaining an improved precision for the test accuracy and also for the train accuracy.

In the setting S1 we observe the highest test accuracy between the methods at 87% obtained by the model M3 (MSVM) as seen by the Table 5.19. In Figure 5.19 we can see a higher rate of misclassifications for the activities of similar nature, meaning the two turns and the three ball kicking activities. Apart from that we can see a significantly smaller amount of misclassifications between the other classes that are not similar, as the long pass and the 180 degrees turn for example.

S1					
Method	Test accuracy	Train accuracy	Total time	Test time	Train time
M1 (NB)	86%	92%	58s	31s	27s
M2 (K-NN)	84%	100%	32s	30s	2s
M3 (MSVM)	87%	100%	53s	33s	20s
M4 (DA)	83%	100%	34s	30s	4s
M5 (DT)	63%	93%	33s	29s	4s

Table 5.19: Table of results for the Raw Data case of setting S1

Regarding the setting S2 we see the highest test accuracy achieved by the method M3 (MSVM) reaching 92% as inferred by the Table 5.20. This is the maximal percentage reached between all the six settings of our current case of raw data analysis for all the activities. Based on Figure 5.20 we see a very high prediction rate between non similar activities but we see an increased number of misinterpretations for the ball kicking activities. A trait that is not observed in this setting but shared by most of the cases is the large amount of confusion between the two turns. The error of prediction for the non similar activities remains at a very low level, a characteristic that is shown in our previous setting, as well as in most of the comparisons of this project. We observe that the setting S2 achieves higher precision compared to S1 and S3 (Table 5.19 and Table 5.21) probably because the gyroscope data contain highly identifiable patterns that the mixed data setting S3 lacks due to the combination of signals. Further examination of the issue is required to achieve better clarification.

S2					
Method	Test accuracy	Train accuracy	Total time	Test time	Train time
M1 (NB)	87%	94%	58s	31s	27s
M2 (K-NN)	86%	100%	32s	30s	2s
M3 (MSVM)	92%	100%	54s	34s	20s
M4 (DA)	86%	100%	39s	30s	34s
M5 (DT)	76%	96%	34s	30s	4s

Table 5.20: Table of results for the Raw Data case of setting S2

The setting S3 presents the maximal test accuracy between the methods as 87% achieved by the method M1 (NB) from Table 5.21. The Figure 5.21 shows a distribution of misclassifications very similar to our previous settings, having the largest part of the error explained by the very high similarity between the two turns and between the ball kicking activities, since the legs of the players behave almost identically when performing the relative moves.

S3					
Method	Test accuracy	Train accuracy	Total time	Test time	Train time
M1 (NB)	87%	93%	58s	31s	27s
M2 (K-NN)	83%	100%	32s	30s	2s
M3 (MSVM)	84%	100%	55s	33s	22s
M4 (DA)	85%	99%	34s	30s	4s
M5 (DT)	65%	94%	34s	30s	4s

Table 5.21: Table of results for the Raw Data case of setting S3

In the setting S4 the optimal test accuracy is obtained by the method M1 (NB) in the Table 5.22, being 85%. Figure 5.22 provides information of a pattern similar to the previous cases, having the largest amount of error attributed to the ball kicking activities and the turns. A lighter distribution of misclassifications can be seen between the activities of different kind of performance by the players. The results of method M3 are skipped due to the extremely long running time. The running process of the code was stopped after 40 minutes duration. At this point we see again that the method M2 (K-NN) gives a low accuracy when spectral features are added, a matter that can be discussed further in future research.

S4					
Method	Test accuracy	Train accuracy	Total time	Test time	Train time
M1 (NB)	85%	93%	67s	33s	34s
M2 (K-NN)	56%	100%	34s	31s	3s
M3 (MSVM)	-	-	>40m	-	-
M4 (DA)	81%	100%	36s	32s	4s
M5 (DT)	64%	93%	35s	31s	4s

Table 5.22: Table of results for the Raw Data case of setting S4

For the setting S5 the largest test accuracy is achieved by the model M1 (NB) reaching 89% as seen by the Table 5.23. In Figure 5.23 we have a stronger confusion on the distinction of the ball kicking activities and an almost perfect prediction on the identification of the two turns. Furthermore, a very low number of misclassified observations can be seen for the classes of different nature. The model M3 is omitted in this case because of the extreme computational time. The running of the code was terminated after exceeded the duration of 40 minutes.

S5					
Method	Test accuracy	Train accuracy	Total time	Test time	Train time
M1 (NB)	89%	95%	66s	32s	34s
M2 (K-NN)	57%	100%	33s	31s	2s
M3 (MSVM)	-	-	>40m	-	-
M4 (DA)	82%	100%	35s	31s	4s
M5 (DT)	75%	96%	35s	31s	4s

Table 5.23: Table of results for the Raw Data case of setting S5

The largest test accuracy of 85% is achieved by the method M1 (NB) for the setting S6 based on Table 5.24. We can infer from Figure 5.24 a distribution of correctly and wrongly classified observations similar to most of the previous cases, having a high error for the two turns and the ball kicking activities and an improved prediction between the other classes. The model M3 is missing because of the extreme running time. As in the previous cases, the running of the code was stopped after 40 minutes duration.

S6					
Method	Test accuracy	Train accuracy	Total time	Test time	Train time
M1 (NB)	85%	93%	75s	39s	36s
M2 (K-NN)	52%	100%	33s	31s	2s
M3 (MSVM)	-	-	>40m	-	-
M4 (DA)	81%	100%	36s	31s	5s
M5 (DT)	65%	93%	37s	32s	5s

Table 5.24: Table of results for the Raw Data case of setting S6

5.6.2. Graphs

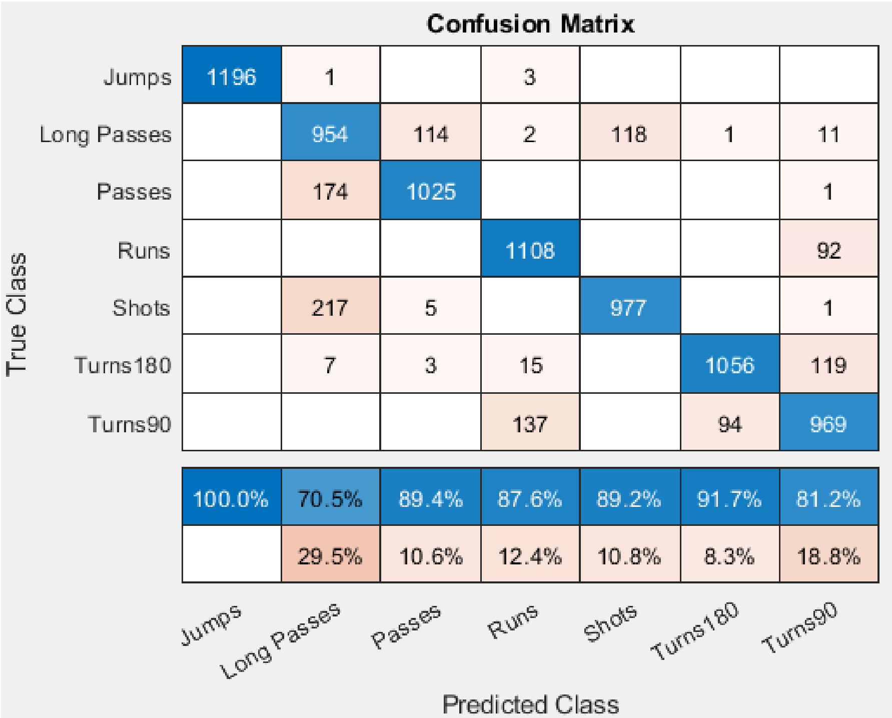

Figure 5.19: Confusion Matrix of M3 (MSVM) for the setting S1 with predictive accuracy of 87% for the Raw Data case.

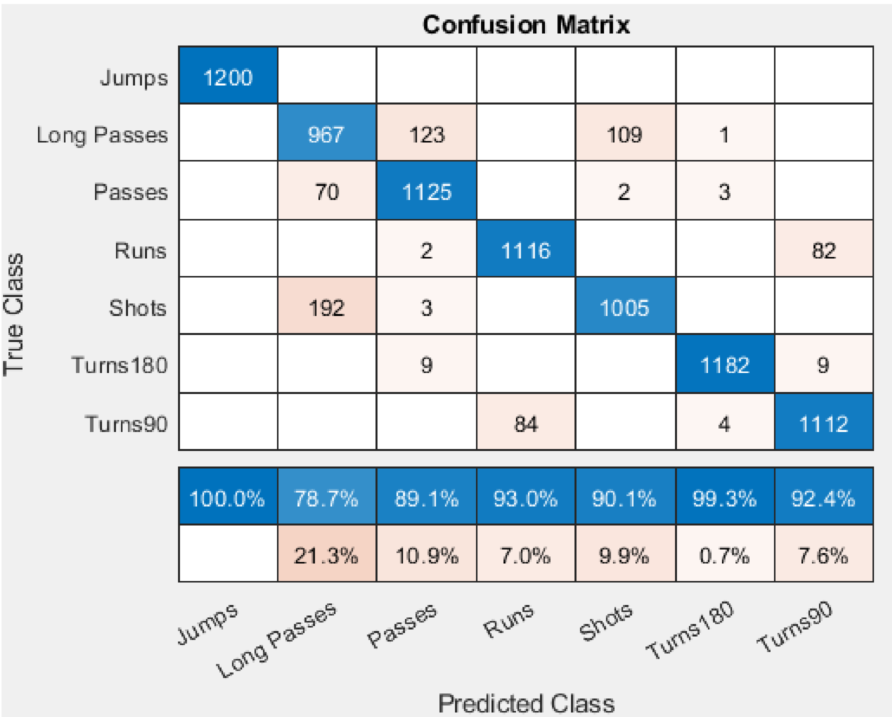

Figure 5.20: Confusion Matrix of M3 (MSVM) for the setting S2 with predictive accuracy of 92% for the Raw Data case.

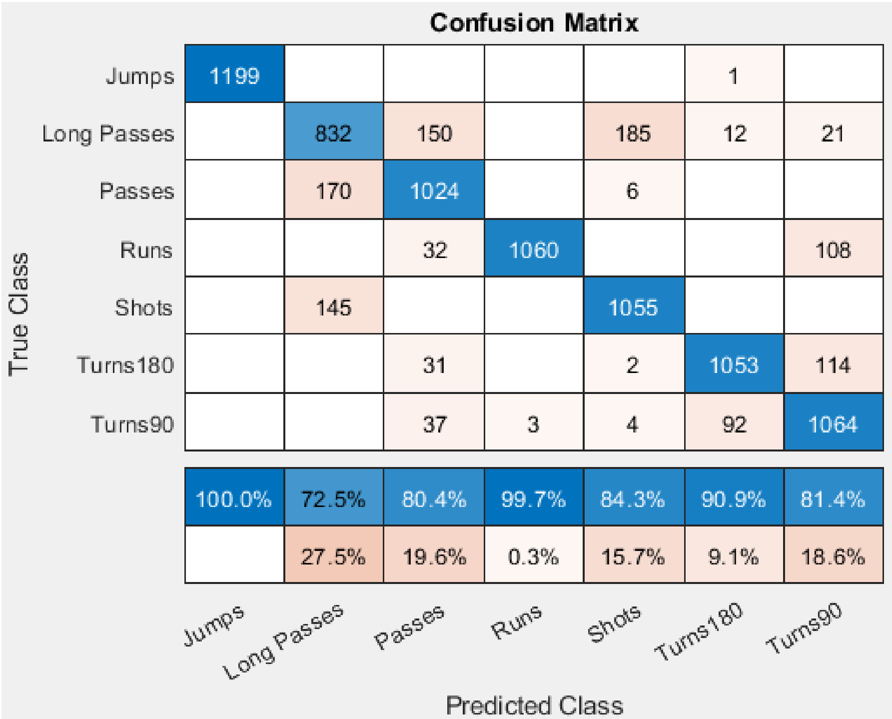

Figure 5.21: Confusion Matrix of M1 (NB) for the setting S3 with predictive accuracy of 87% for the Raw Data case.

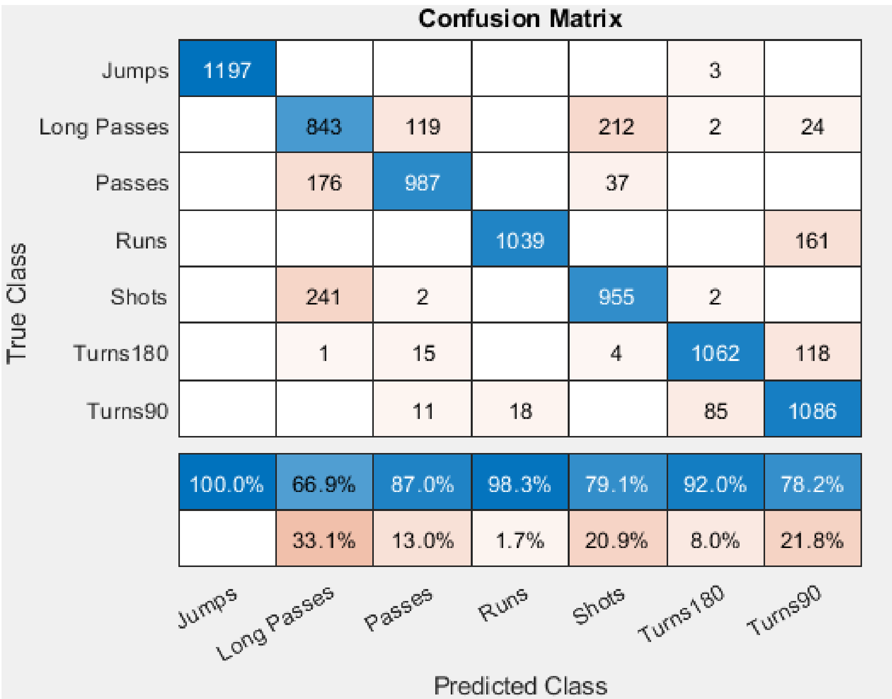

Figure 5.22: Confusion Matrix of M1 (NB) for the setting S4 with predictive accuracy of 85% for the Raw Data case.

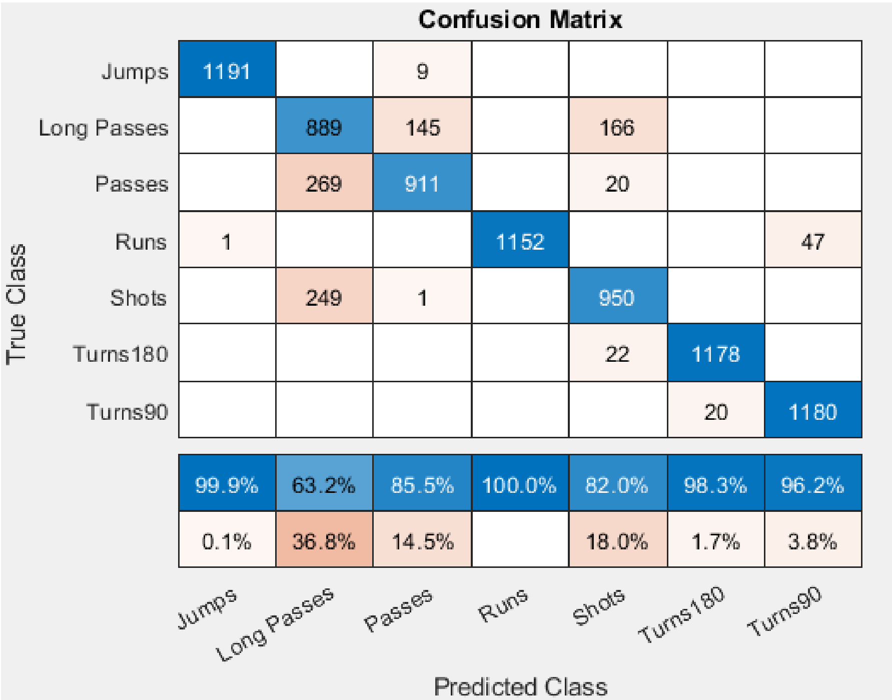

Figure 5.23: Confusion Matrix of M1 (NB) for the setting S5 with predictive accuracy of 89% for the Raw Data case.

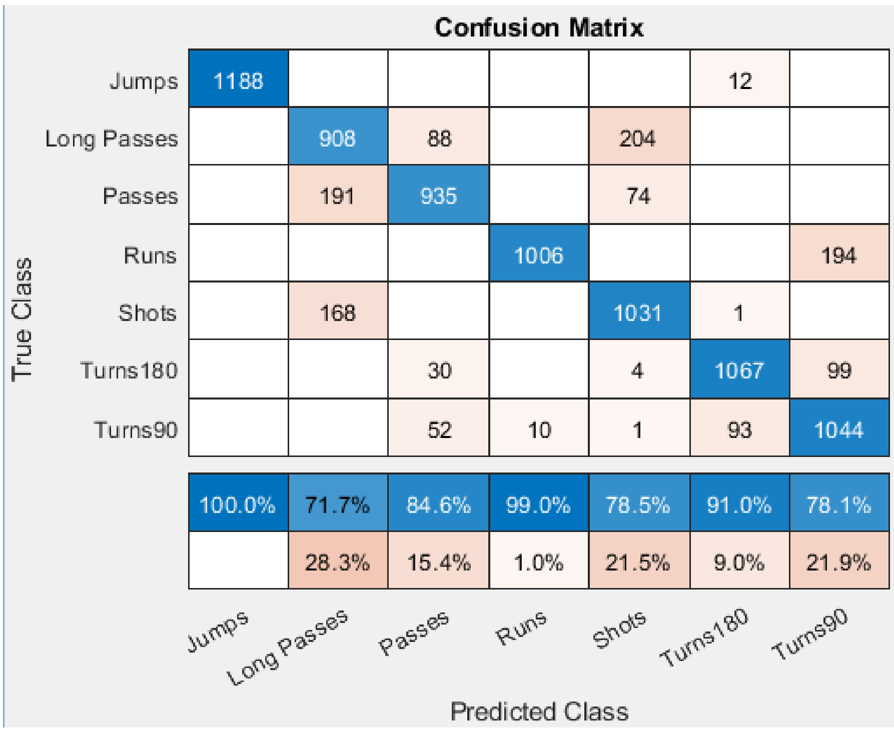

Figure 5.24: Confusion Matrix of M1 (NB) for the setting S6 with predictive accuracy of 85% for the Raw Data case.

5.7. Comparison between the Euclidean Norm case and the raw data case

In this section we discuss the most important information we obtained for each setting and we compare the Euclidean data case with the raw data case of the same setting. We begin with the setting S1 for which we observe the maximal test accuracy of 81%, obtained by model M4 (DA) with total computational time of 31s. For the raw data case we see an increased maximal test accuracy of 87% by the model M3 (MSVM) with total running time of 53s. The increase in maximal accuracy is 6% in favor of the raw data case. We can infer that the increased test accuracy for the raw data case is explained due to the detailed information provided by the true three dimensional form of our data. The difference of 22s in total computational time can be explained by the fact that the model of the raw data case has way more information to process. The distribution of misclassifications is very similar between the Figures 5.13 and 5.19, having an increased error of prediction between the similar activities, meaning the ones related to the ball kicking and the two turns.

The setting S2 has the largest test accuracy at 85%, achieved by the model M3 (MSVM) with total computational time of 68s for the Euclidean data case. For the raw data we observe the largest test accuracy at 92% by the model M3 (MSVM) with total computational time of 54s. The difference in maximal accuracy is therefore 7% for the raw data. As in the previous setting we see an increased test accuracy by the raw data case that can be well explained by the more descriptive form of data that consist the input of our model. The difference in total running time is 14s more for the Euclidean Norm case. The Figures 5.14 and 5.20 show a similar spreading of the misclassified observations mainly between the ball kicking activities. One important difference between the two Figures is that in the raw data case we see a significantly lower amount of misclassifications between the two turns.

Regarding the setting S3 we have the largest test accuracy at 80% by the method M4 (DA) with total computational time of 31s for the Euclidean data case. The maximal test accuracy for the raw data case is 87% by the method M1 (NB) with total computational time of 58s. The improvement for the test accuracy is 7% for the raw data case. We also observe for this setting the increase in accuracy for the raw data which can be explained by the difference in the handling of data, presenting them in their natural three dimensional structure. The additional running time of 27s for the raw data model is attributed to the larger amount of input data for processing. The information of Figures 5.15 and 5.21 is very close in comparison, having the largest amount of error shared between the predictions of the ball kicking activities and also between the two turns, seeing a lighter distribution of misinterpretations for the other classes.

The maximal test accuracy achieved for the setting S4 of the Euclidean data case is 86% by the method M4 (DA) with total computational time of 30s. The highest test accuracy for the raw data case is 85% achieved by the method M1 (NB) with a total computational time of 67s. The accuracy is reduced by 1% for the raw data. The test accuracy of the two cases is almost identical and the increase in total computational time for the raw data case (37s) is due to the larger amount of input information for the running of the model. Based on the Figures 5.16 and 5.22 we see a very close rate of prediction between the two settings, retaining a higher error for the activities of similar nature.

The setting S5 shows a maximal test accuracy at 83% by the model M4 (DA) with total computational time of 33s for the Euclidean data case. The largest test accuracy of 89% is obtained by the model M1 (NB) for the raw data case with total computational time of 66s. We observe an increase in accuracy for the raw data case by 6% and also an addition of 33s to the total running time of the model as observed in most of the previous descriptions. The main observation with respect to the Figure 5.23 is the very high distinction between the two turns as opposed to the Figure 5.17. Again the higher misclassification rate for the ball kicking activities of both figures, still holds.

The largest test accuracy for the setting S6 is 85% and it is obtained by the method M4 (DA) for the Euclidean data case with total computational time of 31s. For the raw data case we have the maximal percentage also at 85% achieved by the model M1 (NB) with total computational time of 75s. Here we observe an identical prediction rate, having again a larger total computational time by 44s for the raw data case. The Figures 5.18 and 5.24 show a distribution of misinterpreted observations very similar to the previous cases, indicating the lower rate of prediction of the models for activities that contain movements and performance of similar nature.

5.8. General discussion for all settings

In this paragraph we present the most important characteristics and observations concerning a general discussion between all of the settings of the Euclidean and raw data cases. The classification of all seven activities using the Euclidean Norm receives most of the maximal percentages of accuracy by the method M4 (DA) for the settings S1,S3,S4,S5,S6. The only exemption is the setting S2 which reaches the largest test accuracy with the method M3 (MSVM). The addition of spectral features to the analysis seems to raise the highest prediction rates for the settings S4,S6 compared to the respective cases with only spatial features (S1,S3), while the percentage of S5 has a slight decrease compared to the respective setting S2 containing only spatial features. Regarding the raw data case we obtain the highest test accuracies for the settings S3,S4,S5,S6 from the method M1 (NB) while the settings S1 and S2 achieve their maximal accuracy through the method M3 (MSVM). This is a characteristic similar to the Euclidean data case, since the M3 (MSVM) also achieves the highest test accuracy for the setting S2 and also for the setting S1 the M3 (MSVM) is almost identical to the largest percentage of M4 (DA). We can observe a significant increase in the overall accuracies of the raw data case compared to the Euclidean Norm case, since the input information of our models is a lot more extensive and representative of the data pattern of each class or activity. The additional precision comes with a raised computational cost as extensively discussed in Section 5.7. For the raw data case, the addition of spectral features to the analysis seems to cause a small decrease in accuracy for the settings S4,S5,S6 compared to the respective settings S1,S2,S3 without the spectral features. A trait shared by the Euclidean Norm case and the raw data case is that the method M3 (MSVM) scores a sufficient rate of prediction for the spatial feature cases S1,S2. The difference between the two cases is that for the Euclidean data the method M4 (DA) has the highest prediction rates for the mixed feature settings (S4,S5,S6), while the raw data case achieves the best predictions through the method M1 (NB) for the same mixed feature settings. Based on all the previous observations we can infer that the method M3 (MSVM) consist a very strong and robust classifier when it comes to recognition of data that contain spatial features, and this observation seems to exist for both the Euclidean and raw data cases. Another important observation is the strong predictive accuracy of the method M4 (DA) and M1 (NB) for the classification of mixed feature based data of the Euclidean Norm and raw data cases respectively. Based on our experience, the setting S2 is recommended as the most effective choice for the extended raw data case, while for the Euclidean Norm case the setting S4 is preferred.

5.9. Connection to literature

In this Section we present the comparisons and observations between the results of our classification and the reviews of Chapter 2. The possible similar patterns and behavior of the models discussed, can prove to clarify and shed light on the way of handling problems of classification in sports. Furthermore, the comparison can help future researchers create a fundamental basis that will assist them in the achievement of the desired high accuracy of their motion recognition and prediction.

5.9.1. Comparison of current research and literature

In this paragraph we present a brief comparison of our study and the projects of Section 2.1. The research [7] achieved a high accuracy of 90% using the Naive Bayes classifier, while we achieved the second highest prediction rate for the same classifier at 97% for the setting S2 for the four activity case of the raw data. Regarding the research [2] we see the highest prediction rate at 88.9% using the k-Nearest Neighbors. In our project the K-Nearest Neighbors achieved the optimal accuracy at 94% for the setting S2 for the four activity classification of the raw data case. The study [6] achieved the largest accuracy by the SVM classifier at 96.7%, as mentioned by the researchers of the project[6]. In our case the MSVM achieved the optimal percentage at 99% for the setting S2 of the four activity classification of the raw data. In the next paragraphs we provide some detailed comparisons between our work and studies of Section 2.3.

Classification in football and Skateboarding

The first significant similarity between our research and the skateboarding classification[10] is the use of time windows in order to isolate the specific activities and movements relative to the experiment, to obtain data that provide information with the least amount of noise. Additionally we utilized some spatial features used by the researchers of the skateboarding analysis[10]. Moreover the skateboarding analysis[10] utilized the raw data as well as norms of the described features, a plan that was also followed in this research. The most significant common trait of our analysis and the skateboarding classification[10] is the use of the Naive Bayes and Support Vector Machines classifiers. Both methods scored a very high prediction rate of 97.8% for [10] as seen in Chapter 2. These methods also achieved a very high accuracy in our research having the maximal percentage of 92% (setting S2) between all the settings(S1,S2,S3,S4,S5,S6) obtained by method M3 (MSVM) for the Euclidean Norm case for the four activities (pass,shot,90 degree turn,180 degree turn). The same method achieved 85% accuracy (setting S2) as the maximal percentage between all settings(S1,S2,S3,S4,S5,S6) for the classification using the Euclidean Norm for the case of all seven activities (pass,shot,long pass,90 degree turn,180 degree turn,jump,running). Regarding the classification of four activities using raw data we obtained the highest result of the whole analysis at 99%, achieved by the method M3 (MSVM) for the setting S2. For the classification of all seven activities using raw data we receive the highest percentage (92%) between all the settings, achieved by the method M3 (MSVM) for the setting S2. It is important to mention that for the case of four activities using raw data the method M1 (NB) achieved 97% accuracy for the setting S5 which also is the highest score observed from the specific classifier in this project. We can infer that the Support Vector Machines and the Naive Bayes methods consist very strong candidates when it comes to classification in sports judging by the results of our analysis and also from [10].

Classification in football and human physical activity

Comparing the research [13] with our analysis we observe the common approach of window segmentation to utilize data relevant to the specific activities and to reduce noise and prediction error. The Euclidean Norm of the signals is also utilized as a way of refining the data. Based on the results of Chapter 2 for the relevant research [13], we see almost identical prediction rates for the Naive Bayes and Support Vector Machines having 97.4% and 97.8% respectively. It is obvious that these results are very similar to the results of the research [10]. We mention here also that for our research the most effective methods was the M3 (MSVM) and M1 (NB). The highest percentage of M3 (MSVM) was 99% for the setting S2 for the case of classification of four activities using the raw data. The largest accuracy achieved by method M1 (NB) was 97% for the setting S5 for the approach of four activities using raw data.

Classification in football and Tennis

Regarding the research [12] we observe that the window based approach was applied by the researcher in order to isolate and distinguish the useful data that concern the activities from the noise. The Euclidean Norm is also used to handle the multidimensional form of the signals (X,Y,Z) for the 3D space. The main classification method utilized for the research [12] is the Support Vector Machines, a method that is also applied very often in our research while achieving very high prediction rates. We mention that the author of [12] achieved 79% as the highest percentage of accuracy with the Support Vector Machines while the same method of classification achieved the highest prediction accuracy for our research, meaning 99% for the setting S2 for the case of classification of four activities using the raw data. Regarding the classification of all seven activities the method M3 (MSVM) scored the highest predictive precision of 92% for the setting S2 using the raw form of our data.

At this point we also need to include the most effective results of the method M3 (MSVM) for our Euclidean Norm cases, since this consists the main feature utilized by the author of [12]. For the classification of four activities, we achieved the percentage of 92% for setting S2, while we obtained the accuracy of 85% for the recognition of all seven activities for the setting S2.

Classification in football and Classification during a soccer match

Comparing the research [18] with our project we see the common use of the window segmentation to obtain refined data relevant to the specific activities. Moreover, our analysis included some spatial domain features utilized also in [18] meaning mean, variance, skewness and kurtosis. Some of the most significant results of [18] are the ones obtained by SVM (99.9%) and NB (98.5%). In our case, we have the method M3 (MSVM) and M1 (NB) which are very commonly used in our project, achieving also the highest prediction accuracies of 99% (S2) for the four activity case of the raw data and 97% (S5) for the same case. We mention that these are the maximal accuracies achieved in this research. For the seven activity case the MSVM produced the largest accuracy of 92% (S2) for the raw data, while the NB followed with 89% (S5) for the same case.

5.10. Summary of classification of four and seven activities

At this section we present some summaries for the most important results of the analysis based on the classification of four activities and also on the classification of seven activities. The descriptions and Figures that follow are focused on the test accuracies for every method and for every setting. For the Euclidean Norm case we present the percentages of the accuracies in bold when they exceed 79%. For the raw data case we give the percentages that are above 85% in bold for a better distinction.

5.10.1. Summary of classification of four activities

Euclidean Norm

In Figure 5.25 we see that for the settings S1,S2,S3 the accuracies are larger on average compared to the settings that include also spectral domain features (S4,S5,S6). The reason for that can be that the added spectral features introduce information less representative of each activity by giving numbers that are similar to all the actions, reducing the identifiable patterns of each action. We can also see that the highest percentages were achieved mainly for the settings S1 and S2, were we used the acceleration and gyroscope data separately. This can be explained by the fact that each of the separate signals contain a distribution that can be well identified by the methods, while by merging the data, the information of the new signal present a pattern different and less unique for each activity of a specific class. We mention that the most effective classifications were obtained by the method M4 (DA) (92%) for the setting S1 and by the M3 (MSVM) for the setting S2 (92%). We also present the computational times for the best three results, having 34s for the M4 (DA) of S1 (92%), 40s for the M3 (MSVM) of S2 (92%) and 31s for the M4 (DA) of S4 (92%). An observation that can be made is that the method M3 (MSVM), while it has a high accuracy for the spatial feature settings, it gives significantly lower percentages when spectral features are added.

	S1 Ac / Spa	S2 Gy / Spa	S3 Ac+gy / Spa	S4 Ac / Spa+spe	S5 Gy /Spa+spe	S6 Ac+gy /Spa+spe
M1 (NB)	**85%**	**87%**	**85%**	**85%**	**89%**	**85%**
M2 (KNN)	**90%**	**90%**	**82%**	67%	67%	66%
M3 (MSVM)	**91%**	**92%**	**85%**	44%	79%	28%
M4 (DA)	**92%**	**87%**	**89%**	**92%**	**88%**	**88%**
M5 (DT)	78%	72%	74%	77%	72%	73%

Figure 5.25: Summary of results for the Euclidean Norm approach of four activities

Raw data

For the Figure 5.26 we observe that the methods that preserved a very high accuracy are the M1 (NB), M2 (KNN) and M3 (MSVM). We can see that most of the mentioned methods scored a high prediction, mainly for the settings that include only spatial features (S1, S2, S3) with the exception of M1 (NB) which scored a large precision for the settings with spectral features also. The drop in accuracy for classifiers like the M2 (KNN), can be explained by the addition of the spectral features that can confuse the method due to giving observations that are not representative of each action and making the data of each activity similar with each other. The observation that the settings S1 and S2 establish an input of a high rate of identification can be made in this case also, having 99% accuracy for the M3 (MSVM) (S2) and 94% for the M1 (NB) (S1). We also provide the computational costs for the best three results, having 40s for the M3 (MSVM) of S2 (99%), 52s for the M1 (NB) of S2 (97%) and 54s for the M1 (NB) of S5 (97%). A remark similar with the one of the previous paragraph is the reduced performance of the method M2 (KNN) when spectral features are added to the analysis. An issue that can be solved through additional future discussion.

	S1 Ac / Spa	S2 Gy / Spa	S3 Ac+gy / Spa	S4 Ac / Spa+spe	S5 Gy /Spa+spe	S6 Ac+gy /Spa+spe
M1 (NB)	**94%**	**97%**	**93%**	**94%**	**97%**	**92%**
M2 (KNN)	**91%**	**94%**	**89%**	63%	66%	59%
M3 (MSVM)	**93%**	**99%**	**92%**	-	-	-
M4 (DA)	57%	62%	57%	75%	82%	79%
M5 (DT)	71%	80%	72%	71%	83%	70%

Figure 5.26: Summary of results for the raw data approach of four activities

5.10.2. Summary of classification of seven activities

Euclidean Norm

For the classification of seven activities using the Euclidean Norm we see that the highest prediction rates are obtained by M2 (KNN), M3 (MSVM) and M4 (DA). It is obvious that we have a lower accuracy in general, due to the increased number of activities and the confusion induced to the methods. We infer that the method M4 (DA) had the most effective predictions for the specific case, having most of the percentages exceeding 80% and reaching the peak at 86% for setting S4 as inferred by Figure 5.27. The methods M2 (KNN) and M3 (MSVM) behaved reasonably for the settings S1 and S2 which consist a strong input for the classification in this approach also. Additionally we present the computational costs for the best three cases, having 30s for the M4 (DA) of S4 (86%), 31s for the M4 (DA) of S6 (85%) and 68s for the M3 (MSVM) of S2 (85%).

	S1 Ac / Spa	S2 Gy / Spa	S3 Ac+gy / Spa	S4 Ac / Spa+spe	S5 Gy /Spa+spe	S6 Ac+gy /Spa+spe
M1 (NB)	74%	78%	73%	76%	79%	76%
M2 (KNN)	77%	**80%**	73%	63%	59%	58%
M3 (MSVM)	**80%**	**85%**	78%	-	-	-
M4 (DA)	**81%**	**80%**	**80%**	**86%**	**83%**	**85%**
M5 (DT)	62%	64%	64%	67%	66%	64%

Figure 5.27: Summary of results for the Euclidean Norm approach of seven activities

Raw data

For the raw data case we can see that the methods behave with increased precision when it comes to settings with spatial features (S1,S2,S3), having a small decline with the addition of spectral features as observed in Figure 5.28. Furthermore, the largest percentage is obtained by the method M3 (MSVM) for setting S2 (92%). We also observe that the method M1 (NB) proves to be very effective when it comes to the addition of spectral features. Moreover, we indicate the computational costs for the best three results, having 54s for the M3 (MSVM) of S2 (92%), 66s for the M1 (NB) of S5 (89%) and 53s for the M3 (MSVM) of S1 (87%).

	S1 Ac / Spa	S2 Gy / Spa	S3 Ac+gy / Spa	S4 Ac / Spa+spe	S5 Gy /Spa+spe	S6 Ac+gy /Spa+spe
M1 (NB)	86%	87%	87%	85%	89%	85%
M2 (KNN)	84%	86%	83%	56%	57%	52%
M3 (MSVM)	87%	92%	84%	-	-	-
M4 (DA)	83%	86%	85%	81%	82%	81%
M5 (DT)	63%	76%	65%	64%	75%	65%

Figure 5.28: Summary of results for the raw data approach of seven activities

5.11. Classification for the three sensor case

5.11.1. Evaluation

In this section we present some classification settings (S1,S2,S4,S5) for the case of only three sensors (right thigh, left thigh, pelvis) for all the seven activities for the Euclidean Norm case. We give the results and relative discussions as supplemental information, to be able to obtain an idea of how the efficiency of prediction is retained if we decrease the amount of information from the movements that the subjects perform.

For the setting S1 we achieve the highest test accuracy by the method M4 (DA) being 75% as seen in the Table 5.25. An important observation is that the same method scored the highest percentage also for the S1 of the five sensor case but with a higher accuracy due to the increased information utilized by the additional right shank and left shank sensors. The five sensor case (81%) had 6% better precision compared to our current three sensor case (75%). Also the running time for the S1 of the five sensor case is 31s while we have the same total time for our three sensor case. By observing the relative confusion matrix of Figure 5.29 of Section Graphs, we see a distribution of misclassifications very similar to the five sensor case, having a lot of wrong predictions between similar actions, meaning the shot, pass, long pass and also between the two turns (180 degree and 90 degree).

S1					
Method	Test accuracy	Train accuracy	Total time	Test time	Train time
M1 (NB)	70%	76%	45s	35s	10s
M2 (K-NN)	72%	100%	30s	28s	2s
M3 (MSVM)	72%	96%	57s	32s	25s
M4 (DA)	75%	85%	31s	28s	3s
M5 (DT)	62%	85%	31s	28s	3s

Table 5.25: Table of results for the Euclidean Norm case of setting S1

The setting S2 has the largest test accuracy by the method M3 (MSVM) being 69% as seen in the Table 5.26. The specific method gave us the highest test accuracy also for the S2 of the five sensor case but with a larger test accuracy due to the increased input information acquired by the right and left shank sensors. The five sensor case (85%) had 16% better precision compared to our current three sensor case (69%). The computational time for the S2 for the five sensor case is 68s while we have a total time smaller by 17s for the three sensor case (51s). From Figure 5.30 we see a distribution of misclassifications very comparable to the five sensor case, as for the setting S1.

S2					
Method	Test accuracy	Train accuracy	Total time	Test time	Train time
M1 (NB)	64%	73%	37s	28s	9s
M2 (K-NN)	67%	100%	30s	28s	2s
M3 (MSVM)	69%	87%	51s	31s	20s
M4 (DA)	65%	80%	31s	28s	3s
M5 (DT)	57%	90%	30s	28s	2s

Table 5.26: Table of results for the Euclidean Norm case of setting S2

We obtained the highest test accuracy for the setting S4 by the method M4 (DA) at 80% as taken from Table 5.27. This method provided us with the maximal test accuracy (86%) also for the S4 of the five sensor case with an increase in test accuracy by 6% because of the extra input data. The computational time for the S4 for the five sensor case is 30s while we have an almost identical total time for the three sensor case (31s). From Figure 5.31 we see a a confusion matrix very similar to the previous cases. The results of method M3 are omitted due to the extremely long running time. The running process of the code was stopped after 30 minutes duration.

S4					
Method	Test accuracy	Train accuracy	Total time	Test time	Train time
M1 (NB)	74%	82%	37s	26s	11s
M2 (K-NN)	57%	100%	31s	29s	2s
M3 (MSVM)	-	-	>30m	-	-
M4 (DA)	80%	90%	31s	28s	3s
M5 (DT)	65%	92%	30s	28s	2s

Table 5.27: Table of results for the Euclidean Norm case of setting S4

The largest percentage for the setting S5 is 75% by the method M4 (DA) as seen in Table 5.28. This method has the optimal test accuracy (83%) also for the five sensor case for the setting S5. The difference is 8% in favor of the five sensor case due the addition of the two lower leg sensors. The computational time for the S5 for the five sensor case is 33s, being almost the same with the three sensor case. The Figure 5.32 indicates a pattern very similar to the previous settings showing a weakness in the recognition of similar activities especially in this case with a reduced amount of input. The results of method M3 are skipped because of the extended running time. The running process of the code was stopped after 40 minutes duration.

S5					
Method	Test accuracy	Train accuracy	Total time	Test time	Train time
M1 (NB)	74%	81%	40s	29s	11s
M2 (K-NN)	52%	100%	31s	29s	2s
M3 (MSVM)	-	-	>40m	-	-
M4 (DA)	75%	100%	31s	28s	3s
M5 (DT)	65%	88%	30s	28s	2s

Table 5.28: Table of results for the Euclidean Norm case of setting S5

5.11.2. Graphs

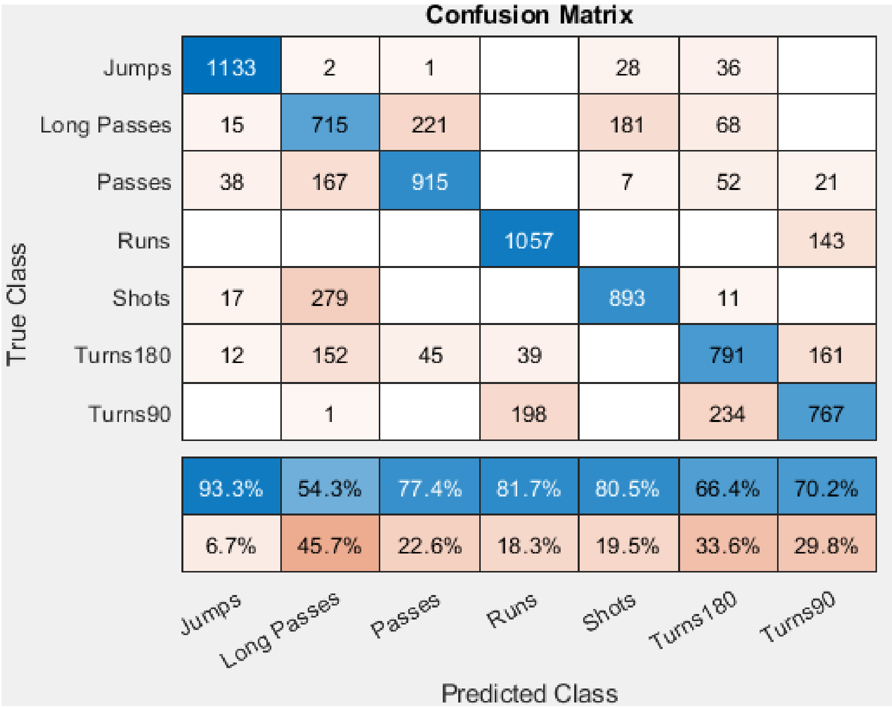

Figure 5.29: Confusion Matrix of M4 (DA) for the setting S1 with predictive accuracy of 75% for the Euclidean Norm case.

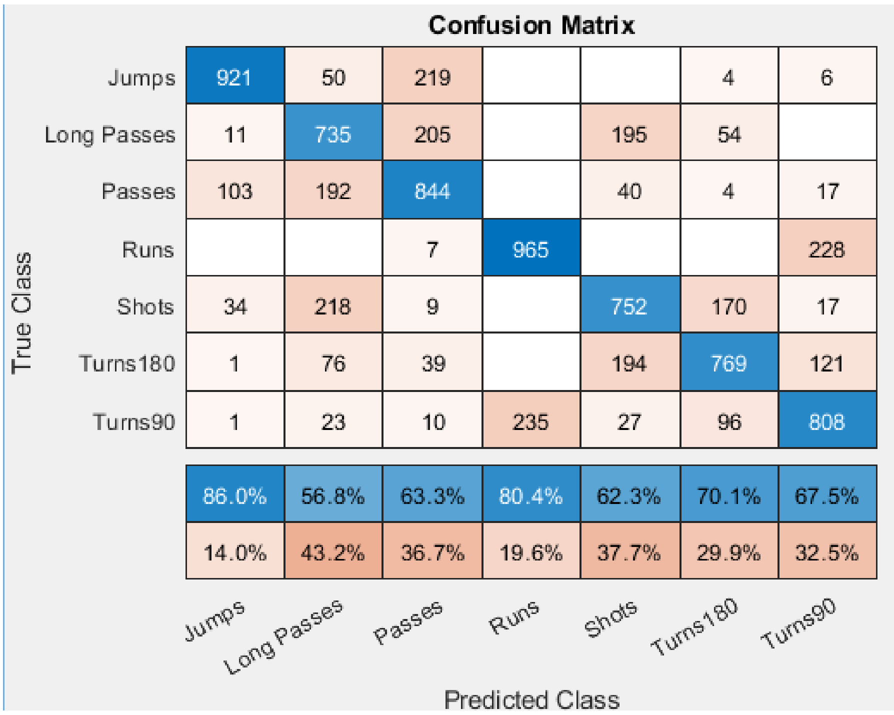

Figure 5.30: Confusion Matrix of M3 (MSVM) for the setting S2 with predictive accuracy of 69% for the Euclidean Norm case.

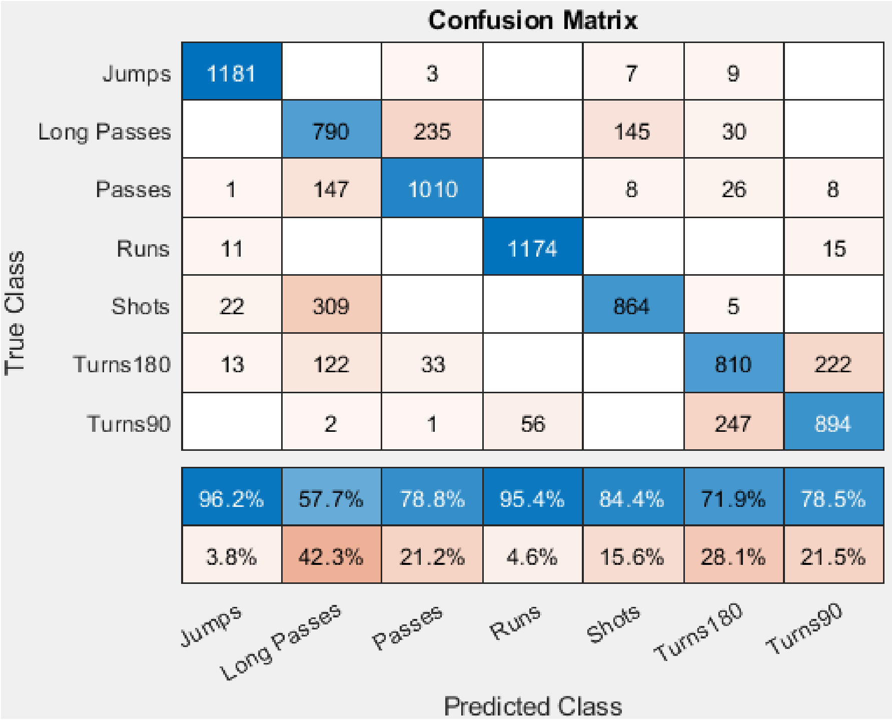

Figure 5.31: Confusion Matrix of M4 (DA) for the setting S4 with predictive accuracy of 80% for the Euclidean Norm case.

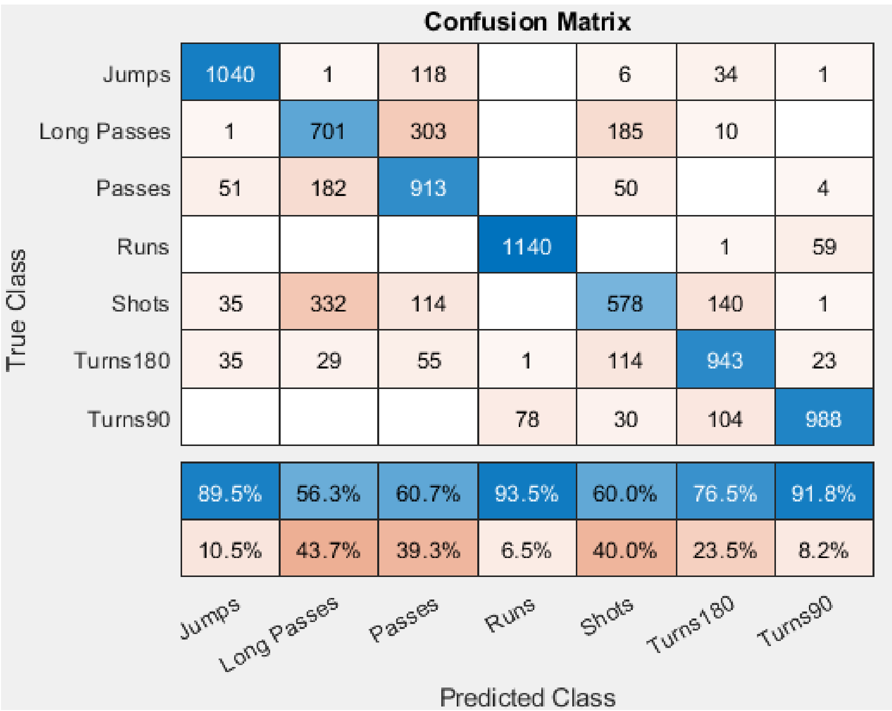

Figure 5.32: Confusion Matrix of M4 (DA) for the setting S5 with predictive accuracy of 75% for the Euclidean Norm case.

5.11.3. General discussion

From the previous analysis we infer that the most effective classifier is the M4 (DA) for the setting S4 with an accuracy of 80% and computational time of 31s. Regarding the setting S1 that includes only spatial features, we reach the maximal test accuracy at 75% by M4 (DA) with 31s of computational time, while the relative setting with added spectral features (S4) achieves the largest prediction rate at 80% again by M4 (DA) with an identical time cost. The setting S2 reaches the highest precision at 69% by M3 (MSVM) with a running time of 51s and the relative setting equipped with spectral features (S5) achieves the percentage of 75% by M4 (DA) with 31s. We notice an increase in accuracy when spectral features are added and a slight raise in computational time. As a final observation we can see that the method M4 (DA) seems to work well for the three sensor case.

6

Conclusion and future improvements

Here we present the answering of the research questions which were formulated in Section 1.4 of Chapter 1. These research questions indicate the main points of interest that led the whole research and finally shed light on the most vital findings that were derived from the whole effort. Initially we restate the questions and the clarifications follow immediately as conclusive remarks. In the last section of this report we provide ideas for possible future improvements and recommendations based on the experience from the whole procedure and evaluation.

1. Should we focus on the analysis of raw three dimensional data or should we consider the Euclidean Norm also? Do we save a significant amount of time by analysing only the norm of the data?

2. Is it more effective to use all of the five sensors or three sensors (pelvis, right thigh, left thigh) are enough? How much loss in accuracy there is for the simplified version of the sensor shorts?

3. Do we need the settings S2 (gyroscope and spatial domain features) and S5 (gyroscope with spatial and spectral domain features) for the raw data case or the settings S1 (acceleration with spatial features) and S4 (acceleration with spatial and spectral features) also?

4. Should we utilize only spatial domain features or mixed features (Spatial together with spectral domain features)?

5. Which of the six settings works best and which one gives the lowest accuracy?

6. Does the classification accuracy increase or decrease if seven instead of four actions are considered and why?

6.1. Conclusive remarks

1. In case we need to make a choice between the Euclidean Norm and the raw data we can choose the raw data. The reason being that for the four activities the prediction rates for the Euclidean Norm are high but we have an average test accuracy of 4.5% points more for the raw data. Also for the raw data case the additional average computational time compared to the Euclidean data is only 20s which is a very low cost for the added accuracy. For the seven activity case we have an average accuracy increase by 5% points for the raw data and 25s additional time cost. For this case also we have a satisfying trade-off. We mention that for the four activity case the most effective classifier for the raw data was the M3 (MSVM) with an accuracy of 99% for the setting S2 (Gyroscope data with spatial features). For the raw data of the seven activity case the best classifier was again the M3 (MSVM) with a precision of 92% for S2.

2. The five sensor case is the most effective approach compared to the three sensor case, since the increased information from the extra sensors provide enhanced classification accuracy. The classification of the three sensors was based on the Euclidean Norm of the seven activities for the settings S1,S2,S4,S5. The relative five sensor case provide an additional average accuracy of 9% points when comparing to the relative settings (S1,S2,S4,S5) of the three sensor case. The computational time is almost identical to both cases. The most effective classifier for the five sensor case is the M4 (DA) (86%) for the S4 (Acceleration data with spatial and spectral features). For the three sensor case the highest accuracy is achieved again by M4 (DA) but with 80% accuracy for the setting S4.

3. It is preferable to use the raw gyroscopic data compared to the acceleration data. If we examine the four activity case we observe an average accuracy increase for the setting S2 (Gyroscope data with spatial features) by 5% points compared to setting S1 (Acceleration data with spatial features). Also for the settings relative to the spatial and spectral features (S4,S5) we observe an increase in average accuracy also by 5% points for the S5 which includes gyroscope data. For the seven activity case we have an increase by 5% points in average accuracy in favor of S2 if we compare it with the S1. The increase is at 3% points in favor of S5 if we compare it with S4.

4. It could be more preferable if we focus on analysis that includes only spatial features compared to an analysis with spatial and spectral features. It can be observed through the tables discussed in Chapter 5 and especially in Section 5.10 that the settings relative to only spatial features (S1,S2,S3) have higher prediction rates on average compared to the settings with mixed features (S4,S5,S6). The highest percentage for the raw data between the gyroscope settings (S2,S5) for the four activity case is 99% by M3 (MSVM) (S2). The highest score between the acceleration settings (S1,S4) is 94% by the M1 (NB) (S1,S4). The largest percentage for the raw data between the gyroscope settings (S2,S5) for the seven activity case is 92% by the M3 (MSVM)(S2). The highest score between the acceleration settings (S1,S4) is 87% by the M3 (MSVM) (S1). We mention one more time that for the four activity classication, the highest percentage is achieved by M3 (MSVM) at 99%, and this rate belongs to the setting S2 that does not include spectral features. The same holds for the maximal rate of the seven activities, achieved by M3 (MSVM) at 92% for the setting S2. The above observations can be an insightful indication that we may omit the spectral domain features.

5. The setting that achieved the best results is the S2. If we compare the relative tables of Section 5.10 we clearly infer that S2 has the highest prediction rates on average. We mention that the best scores of the classification were achieved for S2 having 92% for the Euclidean Norm case of the four activities by M3 (MSVM) and 99% for the raw data case again by M3 (MSVM). Regarding the seven activities, the S2 had the second highest percentage of 85% by the M3 (MSVM) for the Euclidean Norm case. The largest percentage is achieved by the setting S4 for the method M4 (DA)(86%) in the same case. For the raw data case the S2 achieved once more the maximal rate at 92% (MSVM). We also get an important indication of the effectiveness of setting S2, due to the easily recognizable pattern from the gyroscope signal. We can observe that the setting S6 has the worst efficiency on average compared to the other settings. The reason being that S6 contains the mixed signals (Acceleration and Gyroscope). It might be the case that the separate signals include a unique distribution that can be effectively distinguished by the methods, while by fusing the data, the information of the new signal represent a pattern different and many times similar to the activities of the different classes. Another issue can be that the addition of spectral features may create extra similarities in the observations relative to each class. This is still a matter that may need further investigation in the future and we also incorporate that in our recommendations of Section 6.2.

6. We can clearly see that the classification accuracy decreases when we have seven instead of four classes. The reason being that the three additional classes (long pass, jump, running) cause increased confusion to the methods. As mentioned many times before in Chapter 5 there is a big similarity between the two kinds of turns (90 degree turn, 180 degree turn) and also between the two ball kicking activities (pass, shot). When we add another ball kicking activity (long pass) we receive increased false predictions since now we have three actions of similar data pattern, as we see in Figure 4.2 of Section 4.1. Also the running can also be confused with the two turns because by definition when the subject performs the turns, running is included from beginning to end. Regarding the jump, we clearly observe by the confusion matrices presented on Chapter 5 that it retains the highest distinction between the methods, so the specific activity does not contribute to the additional misclassification errors.

6.2. Future improvements and recommendations

Now that the interpretation of all the details of the project reached the end, we provide some ideas for future modifications and applications since there is always room for improvement. One plan that can prove to affect the classification in a positive way, is to increase the percentage of the training data of our methods so that the models can be adapted even more properly. Through this change the methods will raise their prediction accuracy on the test data, since the enhanced information they receive, will enable them to identify the patterns of the data that are derived from the activities more effectively. In order to establish this idea we may need to enlarge our data set so that we avoid having a decrease of the test data size by increasing the training set, which brings us to the next proposition. An approach of similar nature is the augmentation of the size of our input data, by including even more actions for the same amount of classes. The adjustment of the methods will assist the improved distinction of classes by strongly establishing the characteristics of the classes through the enlarged number of activities. A phenomenon that can cover further research is the very slow convergence of the method M3 (MSVM), as observed in the tables of results of Chapter 5, mainly for settings that include spectral features (S4,S5,S6). The recommendations that follow can provide additional assistance in order to shed light on the matter that arose in Chapter 5 regarding cases of low performance of classifiers like the M2 (K-NN), M3 (MSVM) and M4 (DA) when spectral domain features are involved. A supplemental study can be accomplished in order to identify which features are the most discriminative, especially between the spectral domain features. A testing of different spectral features can be held in order to shed light on their links with physical phenomena, as the shaking of muscles during the performance of the actions by the subjects. The dominance of frequency can be examined further to comprehend the productive utilization of spectral features.

References

[1] Matlab and statistics toolbox release 2019b, the mathworks, inc., natick, massachusetts, united states, 2019.

[2] R. Adaskevicius. Method for recognition of the physical activity of human being using a wearable accelerometer, article, vol.20, no.5, 2014. Elektronika ir Elektrotechnika.

[3] Charu C. Aggarwal. Managing and mining sensor data, ibm t. j. watson research center, yorktown heights, ny, usa, 2013.

[4] Yashima Ahuja and Sumit Kumar Yadav. Multiclass classification and support vector machine, global journal of computer science and technology interdisciplinary volume 12 issue 11 version 1.0 year 2012, lovely professional university, jalandhar (punjab) india.

[5] A. Bagheri-Khaligh, R. Raziperchikolaei, and Mohsen Ebrahimi Moghaddam. A new method for shot classification in soccer sports video based on svm classifier, article, department of computer engineering sharif university of technology tehran, iran, m. ebrahimi moghaddam department of electrical computer engineering shahid beheshti university tehran, iran, 2012 ieee southwest symposium on image analysis and interpretation, image analysis and interpretation (ssiai), 2012 ieee southwest symposium on 2012, pp. 109-112. Proceedings of the IEEE Southwest Symposium on Image Analysis and Interpretation.

[6] Peter Blank, Julian Hobach, Dominik Schuldhaus, and Bjoern M. Eskofier. Sensor-based stroke detection and stroke type classification in table tennis, 2015 doi: 10.1145/2802083.2802087.

[7] Connaghan, D., Kelly, P., O' Connor, N. E., Gaffney, M., Walsh, M., and O' Mathuna C. Multisensor classification of tennis strokes. article, 2011 ieee sensors proceedings, sensors, 2011 ieee, pp.1437-1440, 28-31. limerick.

[8] Padraig Cunningham and Sarah Jane Delany. k-nearest neighbour classifiers, article, technical report ucd-csi-2007-4 march 27 2007, university college dublin, dublin institute of technology. Mult Classif Syst.

[9] Marieke de Vries. The use of data science for sports analysis purposes, master thesis report, 2019, tu delft.

[10] Benjamin H. Groh, Thomas Kautz, Dominik Schuldhaus, and Bjoern M. Eskofier. Imu-based trick classification in skateboarding, friedrich-alexander university erlangen-nürnberg (fau), germany, digital sports group, pattern recognition lab, kdd workshop on large-scale sports analytics 2015, sydney, australia.

[11] Andrew D. Ho and Carol C. Yu. Descriptive statistics for modern test score distributions: Skewness, kurtosis, discreteness, and ceiling effects, article, harvard graduate school of education, educational and psychological measurement vol: 75, issue 3, 2015, pp. 365-388.

[12] Xinyu Liu. Tennis stroke recognition, stroke classification using inertial measuring unit and machine learning algorithm in tennis, master thesis report, 2020, tu delft.

[13] Andrea Mannini and Angelo Maria Sabatini. Machine learning methods for classifying human physical activity from on-body accelerometers, article, open access sensors issn 1424-8220, 2010, arts lab, scuola superiore santanna, piazza martiri della libertà, 3356124 pisa, italy.

[14] R.N. Mutagi. Understanding the discrete fourier transform, article, 2004, indus university. R. F. Design.

[15] Bhaskar N. Patel, Satish G. Prajapati, and Dr. Kamaljit I. Lakhtaria. Efficient classification of data using decision tree, vol. 2, no. 1, march 2012. Bonfring International Journal of Data Mining.

[16] T. Ramayah, Noor Hazlina Ahmad, Hasliza Abdul Halim, Siti Rohaida, Mohamed Zainal siti, and May-Chiun Lo. Discriminant analysis: An illustrated example, vol. 4(9), pp. 1654-1667, 4 august, 2010, school of management, universiti sains malaysia, minden, 11800 penang, malaysia. faculty of economics and business, universiti malaysia sarawak, 94300 kota, samarahan, sarawak, malaysia. African Journal of Business Management.

[17] Rozemarijn Schotel. Monitoring local muscle load in football, master thesis report, 2019, tu delft.

[18] Dominik Schuldhaus, Constantin Zwick, Harald Körger, Eva Dorschky, Robert Kirk, and Bjoern M. Eskofier. Inertial sensor-based approach for shot/pass classification during a soccer match, kdd workshop on large-scale sports analytics 2015, sydney, australia, digital sports group, pattern recognition lab, friedrich-alexander-universität erlangen-nürnberg (fau), germany.

[19] Alex Smola and S.V.N. Vishwanathan. Introduction to machine learning, cambridge university press, 2008, yahoo! labs santa clara and departments of statistics and computer science purdue university and college of engineering and computer science australian national university.

[20] Sona Taheri and Musa Mammadov. Learning the naïve bayes classifier with optimization models, int. j. appl. math. comput. sci., 2013, vol. 23, no. 4, 787-795, centre for informatics and applied optimization school of science, information technology and engineering, university of ballarat, victoria 3353, australia, victoria research laboratory national ict australia, victoria 3010, australia.

[21] Erik Wilmes. Measuring changes in hamstring contractile strength and lower body sprinting kinematics during a simulated soccer match, master thesis report, 2019, tu delft.

[22] Guangyong Zheng, Ziliang Qian, Qing Yang, Chaochun Wei, Lu Xie, Yangyong Zhu, and Yixue Li. Bmc bioinformatics, the combination approach of svm and ecoc for powerful identification and classification of transcription factor, article, 16 june 2008, biomed central.