INDIAN INSTITUTE OF INFORMATION TECHNOLOGY - ALLAHABAD

MINI PROJECT

Strength Exercise Tracking using Machine Learning on Sensory Data

---

Author:
Mr. Arnav UPPAL - **IIT2021006**

Mr. Tushar SINGH - **IIT2021032**

Mr. Chirag NAIN - **IIT2021079**

Mr. Nitish KUSHWAHA - **IIT2021086**

Mr. Amit KUMAR - **IIT2021088**

Supervisor:
Prof. Pritish VARADWAJ

A paper submitted in fulfillment of the requirements for course project

Information Technology Department

Declaration of Authorship

I, Mr. Arnav UPPAL - **IIT2021006**

Mr. Tushar SINGH - **IIT2021032**

Mr. Chirag NAIN - **IIT2021079**

Mr. Nitish KUSHWAHA - **IIT2021086**

Mr. Amit KUMAR - **IIT2021088** , declare that this thesis titled, "Strength Exercise Tracking using Machine Learning on Sensory Data" and the work presented in it are my own. I confirm that:

- This work was done wholly or mainly while in candidature for a research degree at this University.
- Where any part of this thesis has previously been submitted for a degree or any other qualification at this University or any other institution, this has been clearly stated.
- Where I have consulted the published work of others, this is always clearly attributed.
- Where I have quoted from the work of others, the source is always given. With the exception of such quotations, this thesis is entirely my own work.
- I have acknowledged all main sources of help.
- Where the thesis is based on work done by myself jointly with others, I have made clear exactly what was done by others and what I have contributed myself.

Signed:

Date: 1 December , 2023

Contents

Declaration of Authorship iii

1 Abstract & Introduction 2

2 Problem Statement 4

3 Data Collection 7
- 3.1 Experimental Setup . . . 7
- 3.2 Data Collection Methodology . . . 8
  - 3.2.1 Weights . . . 9
  - 3.2.2 Execution Form . . . 9

4 Data Processing 10
- 4.1 Converting Raw Data . . . 10
- 4.2 Low-pass Filter . . . 11
- 4.3 Principal Component Analysis . . . 11

5 Feature Engineering 13
- 5.1 Aggregated Features . . . 13
- 5.2 Time Domain . . . 13
- 5.3 Frequency Domain: Fourier Transformation . . . 13
- 5.4 New dataset . . . 14
- 5.5 Clustering . . . 14

6 Modeling 17
- 6.1 Classification . . . 17
- 6.2 Classification Results . . . 17
- 6.3 Counting Repetitions . . . 19
- 6.4 Detecting Improper Form . . . 19
- 6.5 Generalization . . . 20

7 Conclusion 21

8 References 22

Acknowledgement

We extend our heartfelt gratitude to ***Prof. Pritish Varadwaj***, whose unwavering support and expert guidance have been the cornerstone of this project. Prof. Varadwaj's insights and encouragement propelled us forward, shaping our understanding of machine learning and its practical applications.

Working on a real-world problem such as classifying five types of strength exercises using accelerometer and gyroscope data, and implementing a mechanism to count repetitions through local maxima analysis, has been nothing short of amazing. The practicality and relevance of this project have provided us with invaluable insights into the intersection of machine learning and real-world applications.

We owe a debt of gratitude to Prof. Pritish Varadwaj for not only guiding us through the intricacies of the project but also for presenting us with the opportunity to tackle a problem with direct applications in the field. His mentorship has been instrumental in bridging the gap between theoretical knowledge and hands-on experience.

To our additional mentors and collaborators, your collective expertise has elevated this project, turning it into a comprehensive exploration of the challenges and opportunities presented by this real-world scenario. Your contributions have added layers of complexity and nuance, making the experience all the more enriching.

This journey has been a collaborative effort, and the support from our classmates has fostered an environment of shared learning and growth. Together, we have navigated the challenges posed by a real-world problem, and for that, we are truly grateful.

Thank you.

Chapter 1

Abstract & Introduction

Abstract

Abstract—Strength training is an essential element of a well-rounded exercise regimen, alongside aerobic exercises. However, the complete investigation of mechanisms governing the automatic monitoring of free weight exercises has yet to occur. The objective of this study is to delve deeper into the potential of context-aware applications in the field of strength training. This is accomplished through the analysis of gyroscope and wearable accelerometer data collected during strength training sessions. The dataset comprised of information pertains to five participants who engaged in a variety of barbell exercises. The purpose is to investigate, develop, and assess models that can monitor exercises, count repetitions, and identify improper form, just like human personal trainers. For classification, the algorithms evaluated in this paper employ supervised learning. A range of machine learning algorithms were trained utilizing the gathered dataset, and their accuracies were contrasted in order to identify and assess the most suitable models.

Introduction

In the past ten years, numerous practical limitations associated with portable sensors such as accelerometers, gyroscopes, and GPS receivers have been successfully resolved. The utilization of wearables, such as smartwatches, for gathering information has facilitated the development of a burgeoning field of research focused on the monitoring and categorization of human behavior. This field mostly revolves around pattern recognition and machine learning techniques. The basis for this lies in the significant commercial potential of context-aware programs and user interfaces. Furthermore, activity recognition can be employed to address significant societal issues such as rehabilitation, sustainability, senior care, and health.

In order to encourage better ways of living, previous endeavors have prioritized the monitoring of physical activities and gathering user input through exercise management systems. These systems partially substitute tasks that are now performed by personal trainers. Within the category of aerobic exercises, such as bicycling, swimming, and running, there are various devices available to monitor and track performance. These include accelerometer and GPS-based pedometers to measure running pace and distance, ECG monitors to track exertion levels, and electronic exercise machines like treadmills, elliptical trainers, stair climbers, and stationary bikes. Incorporating weight training alongside aerobic workouts is a crucial element of a well-rounded exercise regimen. Nevertheless, the investigation of monitoring methods for free weight exercises remains incomplete. At present, there exists just a single fitness wearable brand in the market that asserts to autonomously recognize exercises and log the number of repetitions. Nevertheless, the existing body of material regarding context-aware applications in the field of strength training remains scarce, likely because of the potential for commercial success in such approaches.

Through the progress in context-aware applications, the eventual development of entirely digital personal trainers becomes feasible. In order to determine the specifications of a digital trainer, we shall quickly examine the responsibilities of existing personal trainers. Primarily, personal trainers must possess a comprehensive understanding of human anatomy as well as the fundamental principles of exercise science and nutrition. Furthermore, the capacity to create and implement customized workout regimens that cater to the specific requirements and achievable objectives of each individual in a highly efficient and inspiring manner. Finally, a method of monitoring training sessions to guarantee correct technique and gradual increase in intensity is used to assist clients in safely achieving their fitness objectives.

Although there have been significant advancements in digitizing the initial two responsibilities of a personal trainer in recent years, the fourth and crucial responsibility, which pertains to safety and progress, is not yet effectively integrated into current fitness devices and programs. Hence, the objective of this article is to delve deeper into the potential of context-aware applications in the field of strength training. It achieves this by examining data from the wristband's accelerometer and gyroscope that is collected during free weight exercises. The dataset comprises data from 5 persons engaging in a range of barbell workouts with weights of moderate to high intensity. The objective is to investigate, construct, and assess models that possess the ability, akin to real personal trainers, to monitor exercises, tally repetitions, and identify incorrect form. The techniques assessed in this study are derived from the research conducted by Hoogendoorn and Funkuse. These techniques employ supervised learning methods for the purpose of classification. Different machine learning algorithms were trained using the gathered dataset, and the accuracies were evaluated in order to identify the appropriate models.

Chapter 2

Problem Statement

Despite the advancements in activity recognition through wearable sensors, the application of such technology to the domain of strength training, particularly in free weight exercises, presents a significant gap. Existing studies have primarily focused on recognizing general activities or gym exercises, with limited attention given to the specific challenges posed by free weight workouts. These challenges include the need for accurate tracking of exercises, counting repetitions, and detecting improper form—essential aspects of effective strength training.

Current fitness wearables and applications, while successful in recognizing some exercises, lack comprehensive support for free weight exercises. This gap becomes evident when considering critical elements of strength training, such as progressive overload and correct form. Progressive overload, a fundamental principle in strength training, involves gradually increasing stress on the body, either through increased weights or repetitions over time. The existing literature does not adequately address the proportional relationship between weights and repetitions, a crucial factor for effective strength training.

Moreover, the perceived intensity of exercises and the quality of their execution have been overlooked in current studies. For a context-aware application intended for use during workouts, it is essential to capture realistic workout scenarios, including the intensity of exercises and adherence to proper form. The absence of such considerations raises concerns about the applicability of existing technology in promoting safe and effective strength training.

Therefore, the aim of this report is to address these gaps by exploring the possibilities of context-aware applications within the strength training domain. Specifically, the focus is on leveraging wristband accelerometer and gyroscope data during free weight workouts to develop models capable of tracking exercises, counting repetitions, and analyzing the relationship between workout data and nutritional needs to suggest a balanced diet , monitor long-term fitness progress by analyzing trends in exercise performance, repetition counts, and workout consistency . The objective is to bridge the existing disparity between general activity recognition and the nuanced requirements of strength training, ultimately contributing to the development of more effective and comprehensive digital personal trainers for strength-based workouts.

Literature Review

Van Laerhoven published the initial study on activity recognition utilizing wearable sensors in 2000. In their investigation, scientists linked a pair of trousers with accelerometers to a laptop in order to analyze the unprocessed sensor data gathered during everyday tasks. A system was developed that utilized machine learning techniques, including Kohonen maps and probabilistic models, to accurately identify various activities. The identification of daily activities remains one of the most extensively researched issues in the field. The findings of those investigations are currently extensively utilized and incorporated into numerous commercial products such as Fitbit, Apple Watch, and Samsung Gear. Moreover, contemporary smartphones are integrated with functionalities to monitor common physical activities such as walking, jogging, and cycling, which are offered by Google and Apple. Newer gadgets and software also have the capability to identify certain sports activities such as aerobic exercises, elliptical training, and swimming. The SmartTrack feature of Fitbit automatically identifies and logs exercise sessions, while also capturing data such as the length of the activity, calories expended, and heart rate zones. While these products already endorse the concept of a digital personal trainer, they currently do not adequately support free weight exercises.

Several studies have concentrated on the identification of gym exercises using data collected from wearable sensors. Chang et al. integrated a three-axis accelerometer into a workout glove to monitor hand movements and placed an additional accelerometer on a user's waist to monitor body posture. 10 volunteers were surveyed to execute 3 sets of 15 repetitions for nine different exercises using varied weights in order to collect data. Two approaches, namely the Nave Bayes Classifier and Hidden Markov Models, were assessed to identify the workouts. Both systems obtained an overall recognition accuracy of over 90

Koskimki et al. identified a deficiency in earlier research, namely the absence of diverse exercise options. For this investigation, a pair of accelerometers were employed, with one positioned on the left wrist and the other on the torso. The objective was to generate automated activity logs that accurately display the number of repetitions of a certain exercise completed by end users. Data was gathered from an individual for 30 distinct workouts, with each activity including 3 sets of 10 repetitions. This procedure was iterated: the initial group was allocated for training purposes, while the subsequent group was reserved for independent testing. The individualized model demonstrated a 96

Li et al. introduced a novel recognition technique that utilizes a single accelerometer connected to a glove. The technique involves dividing a filtered stream of acceleration data into time series of varying lengths for peak analysis, as opposed to using fixed-length windows commonly used in conventional methods. Using this time series data, Dynamic Time Warping was utilized to identify and classify exercises. The trials were conducted by four volunteers, consisting of three males and one female. Each participant completed every exercise in 3 sets, utilizing varying weights. They were instructed to exert maximum effort in completing 15 repetitions in each set. Their findings demonstrate that the suggested methodology is viable and capable of attaining commendable outcomes.

The outcomes attained in these investigations are quite encouraging, demonstrating the feasibility of exercise identification based on accelerometer data. Nevertheless, when considering strength training and bodybuilding only, the works failed to address crucial elements of weight training, specifically increasing overload and proper technique. Progressive overload refers to the progressive and systematic rise in the intensity or difficulty of exercise training, which is considered a crucial principle for achieving success in different types of strength training regimens. Progressive overload can be achieved by gradually raising the intensity of the exercise through either adding more weight or performing more repetitions over a period of time. Prior research discussed the utilization of different weights, but these weights were not directly correlated with the number of repetitions. In all instances, participants were advised to aim for a consistent number of repetitions, regardless of the weight used. All sets were performed with either 10 or 15 repetitions, which is considered high in terms of strength training. Typically, sets in strength training do not surpass 5 repetitions. Therefore, it can be concluded that the weights utilized in these investigations can still be classified as light for the 15 repetitions, and at most, medium for the 10 repetitions. Furthermore, none of these studies make any reference to the perceived intensity of the activities. In order to develop a context-aware application suitable for usage during workouts, it is essential that the training data accurately reflects a genuine workout scenario. Moreover, the studies do not assess the proficiency with which the exercises are performed, a crucial element in free weight training and a fundamental characteristic of a digital personal trainer. Improper technique can result in severe injuries, particularly when handling higher weights.

Chapter 3

Data Collection

3.1 Experimental Setup

Prior studies have demonstrated the viability of employing diverse machine learning algorithms on accelerometer data from free weight exercises, yielding impressive outcomes. Nevertheless, in terms of strength training, the task of gathering high-quality datasets has been overlooked. This study aims to address this issue by establishing a controlled experimental setting that replicates authentic strength training sessions.

Prior studies appear to have selected exercises arbitrarily. The rationale behind the selection of exercises for the dataset is not provided. This study will specifically concentrate on the workouts derived from a single training program, although it is crucial to encompass a diverse range of activities. This guarantees the absence of any interference caused by exercise combinations that would never be executed simultaneously in an actual training situation. The program "Starting Strength," authored by Mark Rippetoe, is widely regarded as the quintessential strength training program that may enhance the strength of both novices and seasoned weightlifters. Starting Strength employs fundamental barbell movements that engage all muscle groups, executed within the most optimal range of motion and gradually raised in intensity, in order to stimulate the essential adaptations for enhanced strength. The program consists of five core barbell exercises: Bench Press, Deadlift, Overhead Press, Row, and Squat. The exercises are depicted in Figure 3.1.

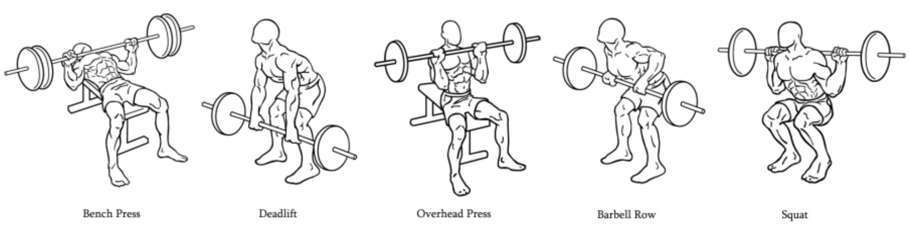

FIGURE 3.1: Basic Strength exercises

Strength training programs are often defined by the concepts of sets and repetitions (reps). A repeat, or rep, refers to a single full movement of an exercise. A set, on the other hand, is a series of successive repetitions performed without interruption, followed by a time of rest. The Starting Strength program involves performing exercises with a substantial load. The objective is to select a weight that allows for the completion of about 5 repetitions. This weight is significantly heavier than the weights often used in other works, and as a result, it will necessitate adjustments in terms of bar route and speed compared to smaller weights. An investigation should be conducted to determine if a trained model can accurately categorize exercises despite these variations. Additionally, it is important to investigate potential methods for verifying the structure of an exercise.

3.2 Data Collection Methodology

In our data collection methodology, we opted for a user-friendly and widely accessible approach by utilizing the *Sensor Data* app, available on the Google Play Store. This app takes advantage of the built-in sensors present in smartphones, such as accelerometers and gyroscopes, to capture the necessary motion and orientation data for our project.

1 : Application Selection

We selected the *Sensor Data* app available on the *Google Play Store* for its capability to record accelerometer and gyroscope data using the sensors already embedded in smartphones.

2 : App Configuration

Upon installing the app, we configured it to record data at the desired sampling rates. The accelerometer was set to capture data at a rate of 25 Hz, and the gyroscope was set at 25 Hz for optimal data resolution.

3 : Participant Instructions

Participants were given instructions on how to use the app, including starting and stopping data recording. Clear guidelines were provided on placing the smartphone in a manner that mimics the sensor placement of a wristband or watch during exercise.

4 : Exercise Performance

Participants performed the designated barbell exercises following the structure of the Starting Strength program. Each participant completed 3 sets of 5 repetitions for one session and 3 sets of 10 repetitions for another session, adhering to the program's guidelines.

5 : Resting Period Data

In between some sets, additional data was collected during 'resting' periods. Participants were not restricted during these resting periods, allowing for a mix of standing, walking, and sitting activities.

6 : Data Recording

Participants initiated data recording using the *Sensor Data* app during each exercise and resting period. The app captured accelerometer and gyroscope readings throughout the specified time frames.

By employing the *Sensor Data* app and leveraging the sensors inherent in smartphones, we aimed to demonstrate the practicality and accessibility of our data collection approach. This methodology aligns with the goal of utilizing commonly available technology for research and encourages broader participation in similar projects.

3.2.1 Weights

In determining the appropriate weight for each participant, the metric of one rep max (1RM) was employed. The 1RM signifies the maximum amount of weight an individual can lift for a single repetition **27**. The calculation of 1RM can be accomplished either through direct maximal testing or indirectly through submaximal estimation, with the latter method being favored for its safety and efficiency. Various formulas exist for estimating 1RM using the submaximal approach, with the Epley and Brzycki formulas being among the most commonly utilized. In this experiment, Epley's formula was employed:

$$1RM = w \times \left(1 + \frac{r}{30}\right)$$

Here, r represents the number of repetitions performed, and w denotes the weight used. Once the 1RM for each exercise is calculated, Epley's formula can further be applied to determine the weight for 5 or 10 reps, corresponding to approximately 85% (5 reps) and 75% (10 reps) of the 1RM. This approach ensures that all participants engage in the exercises with weights tailored to their individual strength levels in the most accurate manner possible.

Table 1. Participants (N=5)

Participant	Gender	Age	Weight (Kg)	Height (cm)	Experience (years)
A	Male	23	95	194	5+
B	Male	24	76	183	5+
C	Male	16	65	181	<1
D	Male	21	85	197	3
E	Female	20	58	165	1

FIGURE 3.2

3.2.2 Execution Form

Next to the main dataset, additional data was collected to examine the quality of execution (i.e., form) of an exercise. Currently, this analysis was specifically conducted for the bench press. A participant was intentionally asked to execute the bench press movement with improper form. The data includes two of the most common errors for this exercise: lowering the bar too high on the chest and not touching the chest at all.

Chapter 4

Data Processing

4.1 Converting Raw Data

The unprocessed dataset consisted of 69,677 records, with each record providing a timestamp for a certain time period and corresponding x, y, and z-values obtained from the sensor. The discrete sensor measurements obtained from the wristband are segregated into distinct files, with each entry being assigned a distinct timestamp. Consequently, the data had to be consolidated. To reduce the amount of information lost during this procedure, a minimal time interval of $\Delta t = 0.20$ s, which corresponds to five instances every second, was selected. The numerical values were consolidated using the arithmetic mean, whereas the categorical features (labels) were consolidated using the statistical mode.

The compiled dataset offered us a favorable initial foundation for creating visual representations. Figure 4.1 displays accelerometer data for a substantial set of each workout. The diagram illustrates that exercise patterns exhibit distinct characteristics, with the repetitions readily identifiable by the peaks. Figure 4.2 illustrates the disparities in y-acceleration between medium and heavy squat sets. The medium weight sets have greater peaks, whereas the heavy sets demonstrate more pronounced dips. This phenomenon can be understood as follows: while performing medium weight repetitions, the speed at which the repetitions increase is higher due to the lower resistance. On the other hand, when performing heavy weight repetitions, the speed at which the repetitions decrease is higher due to the greater load being lifted.

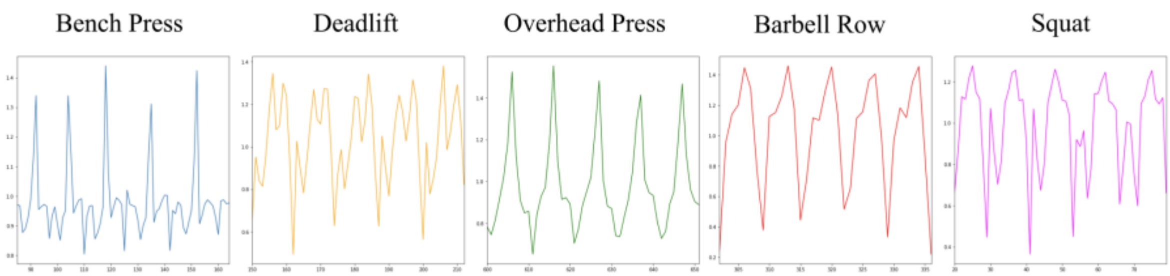

FIGURE 4.1: Accelerometer Data

This section will further elucidate methods for managing unprocessed, unrefined data that contains disturbances. The objective is to manipulate the data in a manner that eliminates minor disturbances and identifies the components of our data that account for the majority of the variability. The discussion will cover two methods:

Low-pass Filtering, which can be used on certain properties, and Principal Component Analysis, which operates on the entire dataset.

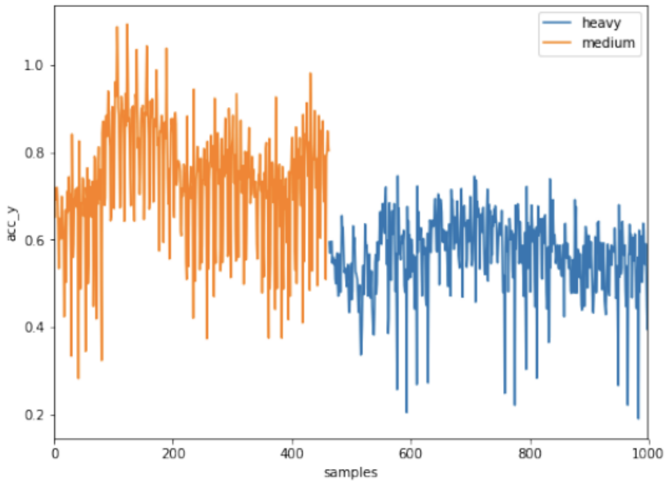

FIGURE 4.2: Medium and heavy weight squats

4.2 Low-pass Filter

In processing temporal data with a perceived form of periodicity, a low-pass filter can be applied to mitigate high-frequency noise that might disturb the learning process. Specifically, the Butterworth low-pass filter was utilized, applied to all but the target features. Visualizations indicated that the movements have a frequency of around 2 seconds per repetition. After further visual inspection and trial-and-error results, as suggested in the work of van den Bogert, the cut-off point was set at 1.3 Hz. Figure 4.3 y-acceleration for a heavy bench press set before and after smoothing.

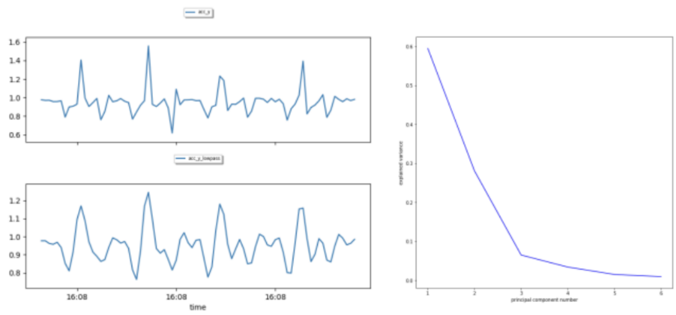

FIGURE 4.3: Low-pass Filter and Principal Component Number

4.3 Principal Component Analysis

A principal component analysis (PCA) was performed to identify the characteristics that could account for the majority of the variance. Principal Component Analysis (PCA) was utilized on all features, with the exception of the target columns. The results are depicted in Figure 4.3, illustrating a significant decline in the explained variance beyond 3 components. Thus, three components are chosen and their values are added to the dataset.

Chapter 5

Feature Engineering

This section we will discuss how additional features were derived from the original dataset including *aggregated features*, *time features*, *frequency features*, and *clusters.*

5.1 Aggregated Features

To further exploit the data, the scalar magnitudes r of the accelerometer and gyroscope were calculated. r is the scalar magnitude of the three combined data points: x, y, and z. The advantage of using r versus any particular data direction is that it is impartial to device orientation and can handle dynamic re-orientations. r is calculated by:

$$r_{\text{magnitude}} = \sqrt{x^2 + y^2 + z^2}$$

5.2 Time Domain

In order to leverage the time-based characteristics of the data, numerical data points were combined by calculating the standard deviation (*sd*) and mean of all features, excluding the target columns, across different window widths. The standard deviation should accurately represent the fluctuations in the data across time. For instance, the standard deviation is anticipated to be greater during exercise compared to resting periods. The temporal mean provides a more accurate representation of the overall data levels, as it is less affected by sudden fluctuations or noise. When choosing an optimal window size, it is important to strike a balance between the susceptibility of the data to noise and the ability to retain forecast accuracy. Figure 4.4 displays the outcomes obtained using window sizes of 2, 4, and 6 seconds. A window duration of 4 seconds was selected and incorporated into the dataset.

5.3 Frequency Domain: Fourier Transformation

Besides the time domain, the frequency domain will also be explored. The idea of a Fourier transformation is that any sequence of measurements can be represented by a combination of sinusoid functions with different frequencies.The same window size of 4 seconds was used to compute the frequency features with a Fourier Transformation including the maximum frequency, the frequency signal weighted average, and the power spectral entropy

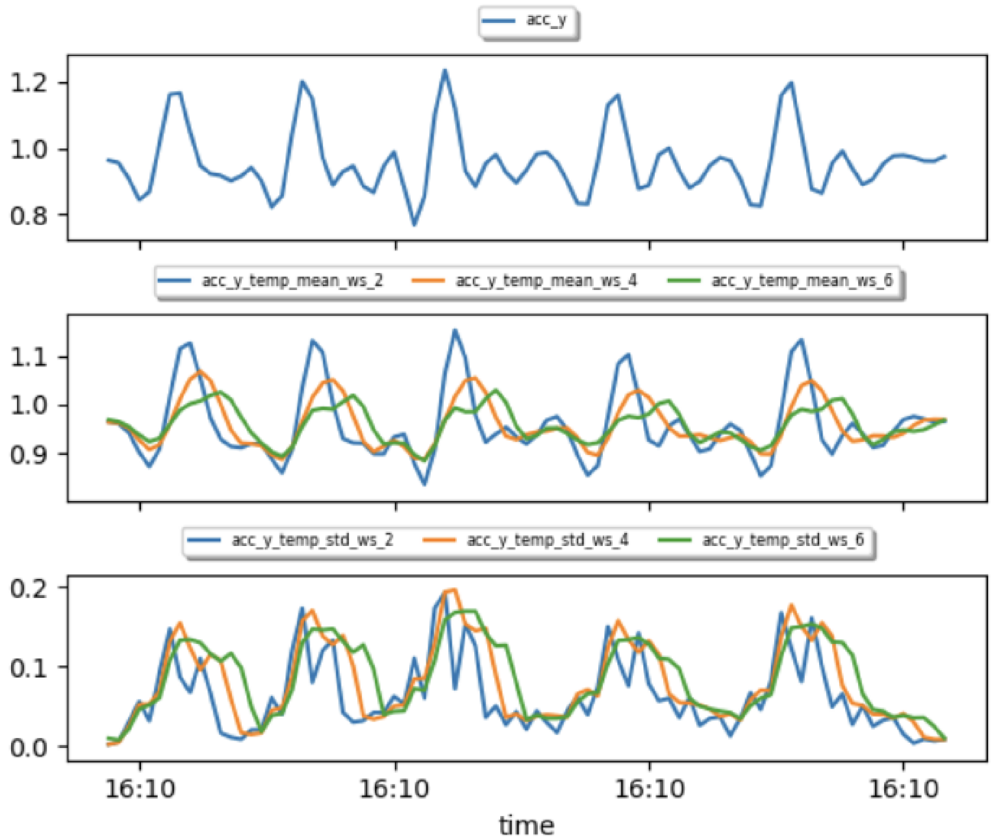

FIGURE 5.1: Numerical temporal aggregation with window sizes of 2, 4, and 6 seconds

5.4 New dataset

The dataset now includes a diverse range of additional attributes. Due to the presence of overlapping time periods, the resulting qualities exhibit a strong correlation. Incorporating all examples may not yield new information, as just one point in the window differs between neighboring instances. To mitigate overfitting, a limit was imposed on the amount of overlap allowed for the windows. Consequently, any occurrences that did not meet this condition were eliminated, with a 50% overlap permitted. The dataset was reduced to 4505 instances. Although some information is lost in this procedure, it has been demonstrated to be beneficial since it reduces the number of highly similar occurrences in the dataset, which can lead to overfitting.

5.5 Clustering

A membership to a specific cluster may have predictive value for a label. The primary objective will be to cluster the acceleration data, as the findings indicated that the gyroscope data was not deemed valuable. Following the execution of many tests, the k-means clustering algorithm with a value of $k = 4$ was determined to be the most favorable approach. This conclusion was based on the observation that it yielded the highest silhouette score of 0.6478, surpassing alternative methods such as k-means and agglomerative clustering. Four clusters were selected since they best represented the various labels, even if they had a slightly higher silhouette score for $k = 2$. Figure 6 displays the results.

The distribution of the measures and labels is presented in Table 2. Cluster 1 encompasses the majority of the bench press and overhead press data. This is logical as both pressing movements are highly similar. The squat exercise is classified within cluster 2, whereas the deadlift and row exercises are classified within cluster 3 with a flawless score. Cluster 4 includes a portion of the remaining data, although it does not accurately represent it.

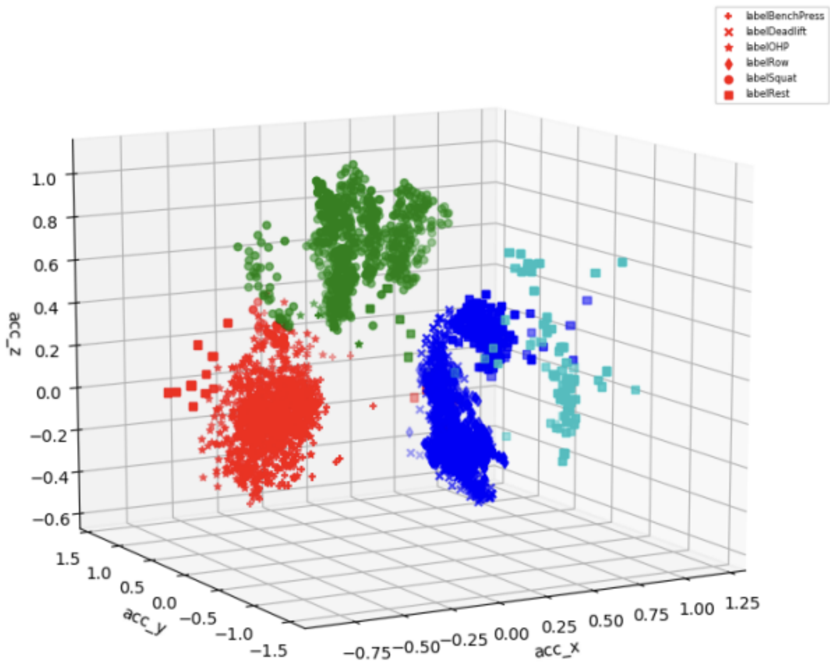

FIGURE 5.2: Clusters

Table 2. Cluster Coverage

Label	Cluster 1	Cluster 2	Cluster 3	Cluster 4
BenchPress	99.88 %	0.12 %	0.00 %	0.00 %
Deadlift	0.00 %	0.00 %	100.00 %	0.00 %
OHP	99.28 %	0.72 %	0.00 %	0.00 %
Row	0.00 %	0.00 %	100.00 %	0.00 %
Squat	2.98 %	97.02 %	0.00 %	0.00 %
Rest	4.14 %	3.78 %	50.45 %	41.62 %

FIGURE 5.3

Chapter 6

Modeling

The dataset is now processed and ready for training. It contains the 6 basic features, 2 scalar magnitude features, 3 PCA features, 16 time features, 12 frequency features, and 1 cluster feature. This section will explain how the models for classification, repetition counting, and form detection were built and evaluated.

6.1 Classification

Because of the temporal nature of our dataset, the training and test set were split up based on the exercise sets. The training data contains the first two sets for each exercise, weight, and participant combination, and the test set contains the remaining sets. This ensures that we have valid test data of sets the models have not seen before.

Feature selection Forward feature selection was used to investigate which features contribute the most to performance, as useless features could impact the performance of the algorithms. Using a simple decision tree and gradually adding the best features, the results showed us that after 15 features, the performance no longer significantly improved. The 5 features with the most predictive power are: ***pca_1, acc_y, pca_3, gyr_x_temp_std_ws_4, acc_r_pse.***

Regularization In order to penalize more complex models, a regularizer was added to the objective functions. Figure 7 shows the impact of adding a regularization parameter on the performance for the training set. It shows that the accuracy of the test set slightly improves with a higher regularization parameter, but only until a certain point. After that, the accuracy decreases on both the training and the test set.

Models First, an initial test run was done to determine the performance of a selection of models and features. This test included the following models: Neural network, Random Forest, Support Vector Machine, K-nearest Neighbours, Decision Tree, Naive Bayes. Grid search was performed on all of the models.

6.2 Classification Results

The findings for each model are displayed on the left side of Figure 6.2. The colors correspond to the distinct feature sets, as elucidated. The chapter 5 feature set encompasses all of the available features. Using these data, a Random Forest model was further refined by selecting the top 15 characteristics that performed the best.

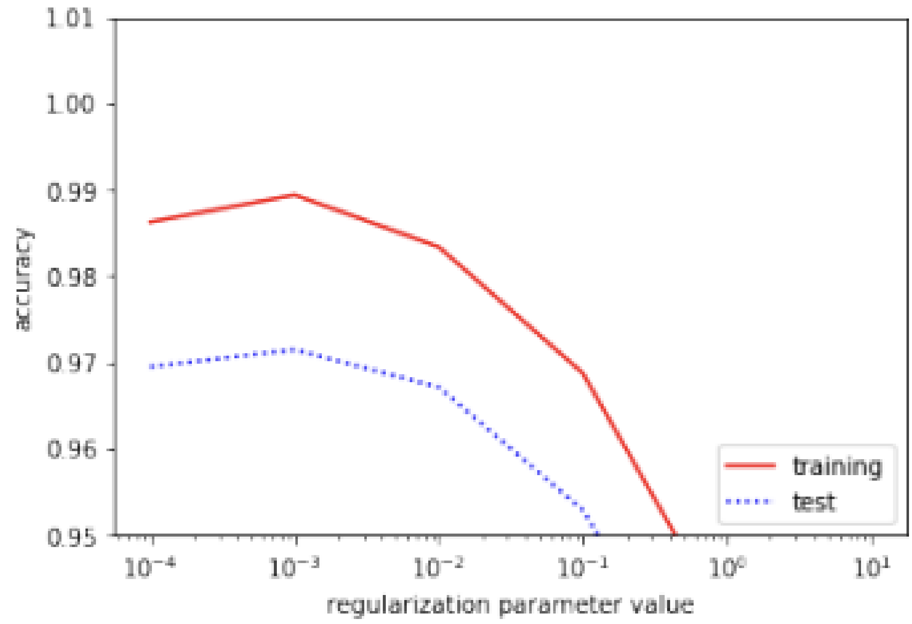

FIGURE 6.1: Regularization

By doing a 5-fold cross-validation and conducting grid search parameter adjustment, the Random Forest algorithm yielded the following ideal parameters: The minimum number of samples required to be at a leaf node is set to 2. The number of estimators, or decision trees, in the ensemble is set to 100. The criterion used for splitting the nodes is the Gini index. Figure 8 displays a confusion matrix of the ultimate predictions. The model attained a comprehensive accuracy of 98.51% on the test set, exhibiting comparable precision, recall, and f1-score.

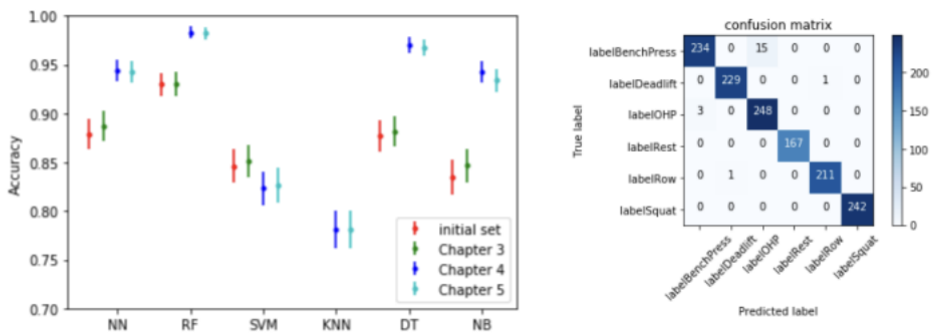

FIGURE 6.2: Model Performance and RF Classification Confusion Matrix

6.3 Counting Repetitions

A straightforward peak counting approach was utilized to tally the occurrences of repetitions in the scalar magnitude acceleration data. In order to disregard minor local peaks, a robust low-pass filter with a 0.4 Hz cut-off frequency was first implemented. It has been discovered that to achieve optimal performance, the method of counting repetitions needs to be customized for each unique activity. Counting the minimum values yielded superior outcomes for the deadlift and overhead press exercises. The aggregate error rate for tallying duplications was approximately 5% for the gathered dataset. Figure 6.3 illustrates a demonstration of performing 10 repetitions of the deadlift exercise.

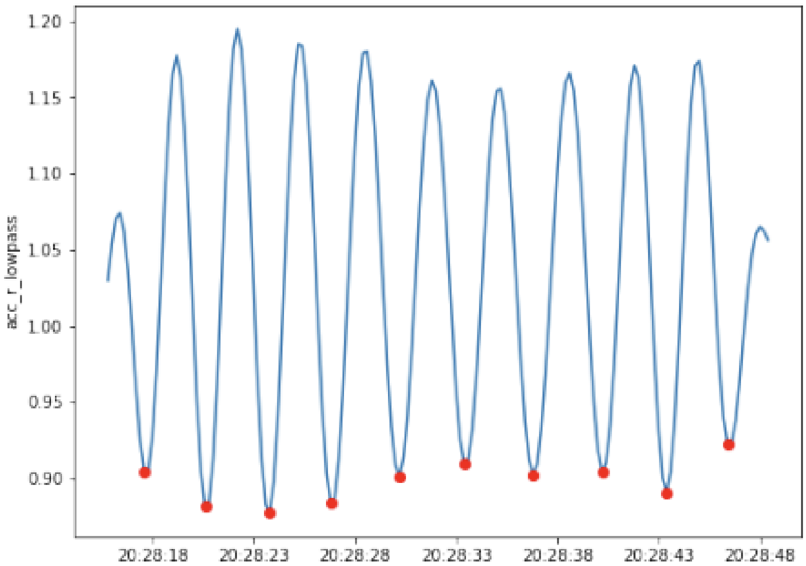

FIGURE 6.3: Counting deadlift repetitions using the minimum values after applying a lowpass filter

6.4 Detecting Improper Form

Additional data was gathered during the studies from a person who performed the bench press with incorrect form. The subject executed several sets in which the bar was intentionally positioned too high on the chest or did not make contact with the chest at all. Subsequently, this data was utilized to train a Random Forest, akin to the one elucidated in this part, for the purpose of classifying the form. The dataset had three distinct labels: accurate form, too high, and absence of contact. The dataset, which contained 1098 cases, underwent the same training and testing methodology as previously described. The model attained a precision of 98.53% on the test dataset.

6.5 Generalization

Assessing the generalizability of the models to novel data is crucial. Firstly, let us examine the various weight categories. As demonstrated in Figure 3 of Section 4, there are already apparent disparities in pace between the two classes. In order to analyze these disparities, the data was divided into a training set and a testing set, categorized according to the weight class. When the model is trained on large datasets, it struggles to effectively apply its knowledge to smaller datasets. Here, the precision decreases to 79.97%. The inverse yields a comparable accuracy of 79.51%. Now, let's examine the various contestants. The leave-one-out strategy was utilized for training and testing, yielding an average accuracy of 85.43% across all participants.

Chapter 7

Conclusion

This paper aimed to explore the possibilities of context-aware applications within the strength training domain. The research was motivated by the insufficient attention given to strength programs in the existing literature and the absence of support from current activity trackers. Data from wristband accelerometers and gyroscopes were gathered during actual strength training sessions involving five individuals. The participants performed the five fundamental barbell exercises using both medium and high weights. The 1RM metric was employed to guarantee that participants executed the lifts using weights proportionate to their strength. Upon applying machine learning techniques to the quantified self cycle data [2], it was determined that a Random Forest model yielded the most accurate workout classification results. The model attained a comprehensive accuracy of 98.51% in accurately categorizing occurrences that were not previously encountered. In order to address the potential accuracy paradox, an assessment was conducted on the precision, recall, and f1-score, which confirmed the reliability of the obtained data. The model's performance was not flawless since it erroneously identified certain occurrences of bench press as overhead press, and vice versa. The same principle applies to both the deadlift and row exercises. One potential rationale for this is that both the bench press and overhead press are vertical movements that involve pressing. Although the torso is in a different posture, the alignment of the wrist remains similar. Similarly, the deadlift and row exercises can be classified as vertical pulling actions. The data was subjected to a straightforward peak counting technique, along with a robust low-pass filter, to determine the number of repetitions. It has been discovered that in order to reach optimal outcomes, rep counting models should be adjusted for each exercise. Although the procedure is simplistic, it yielded satisfactory outcomes with a miscount rate of 5%. A separate Random Forest model was trained specifically to classify the form of bench press exercises. Data was intentionally acquired from a single participant who deliberately executed the bench press exercise with improper technique. The Random Forest achieved a classification accuracy of 98.53% by correctly distinguishing between the appropriate and inappropriate cases. Despite the restricted availability of form detection data, next research should explore alternative workouts. Nevertheless, the presented method can function as a proof of concept. Although these results demonstrate significant potential for potential commercial applications, certain tests revealed that the models struggle to generalize to unfamiliar data. Hence, to actualize such systems, a substantial amount of training data from many sources is necessary. Considering this, we are in the process of developing entirely digital personal trainers that can be accessed from our wrists. These trainers will assist us in achieving our health, fitness, and strength training objectives.

Chapter 8

References

1. Sensor Data, `https://play.google.com/store/apps/Sensor+Data`
2. Samsung Gear, `https://www.samsung.com/us/mobile/wearables/`
3. Apple Watch, `https://www.apple.com/lae/apple-watch-series-4/`
4. Location API, `https://developers.google.com/location-context/`
5. MotionActivity, `https://developer.apple.com/documentation/coremotion`
6. Fitbit Smarttrack, `https://www.fitbit.com/nl/smarttrack`
7. Hoogendoorn, M., & Funk, B. Machine Learning for the Quantified Self.
8. Van Den Bogert, M. (1996). Practical Guide to Data Smoothing and Filtering.
9. Atlas Wristband 2, `https://atlaswearables.com/`
10. Koskimki, H., & Siirtola, P. (2014, December). Recognizing gym exercises using acceleration data from wearable sensors. In CIDM (pp. 321-328).
11. Leutheuser, H., Schuldhaus, D., & Eskofier, B. M. (2013). Hierarchical, multi-sensor based classification of daily life activities: comparison with state-of-the-art algorithms using a benchmark dataset. PloS one, 8(10), e75196.
12. Banos, O., Damas, M., Pomares, H., Prieto, A., & Rojas, I. (2012). Daily living activity recognition based on statistical feature quality group selection. Expert Systems with Applications, 39(9), 8013-8021.