CLASSIFICATION OF PHYSICAL ACTIVITIES WIH MACHINE LEARNING

Introduction:

Accelerometry from wearable sensors can be an objective way to classify physical activities (PA). While previous studies have looked at using accelerometry to identify broad activity categories such as walking or sitting, there have been few attempts to classify specific physical tasks using accelerometry. Thus, in order to create a recognition system for human activity, this study focuses on extracting features from a wearable, 3-axis accelerometer. Our dataset, collected from 39 healthy adults, includes wearable sensor data from a wrist accelerometer. We then aim to see if machine learning (ML) models can predict different activities from this wearable wrist accelerator collected from healthy adults. We are also aiming to find out the effect of different windows with different overlaps on the prediction of activities from accelerometer data.

We then hope to expand upon this predictive ability to see if our models are able to differentiate differences in the activities between healthy and unhealthy adults to discover signs of MS. Multiple sclerosis (MS) is a neurological disease affecting the central nervous system, and it is the leading cause of non-traumatic disability in young adults [1]. Traditional diagnostic methods usually rely on clinical laboratory tests and neuroimaging studies, but these clinic visits necessitate frequent monitoring approaches that are often expensive. Wearable accelerometers offer a promising avenue for continuous monitoring of MS in conditions that are not bound by socioeconomic factors, which can enable timely diagnosis, early treatment access, and disease progression management.

The objectives of this paper are to: 1) Identify the most reliable features derived from wearable devices for classifying physical activities; 2) to evaluate the accuracy of different machine learning models for assessment of physical activities. To achieve these goals, we use statistical analysis and machine learning to analyze sensor data collected using a wearable wrist accelerometer.

By leveraging machine learning models, we can unlock valuable insights from wearable accelerometer data, ultimately aiding in the diagnosis and telehealth support for patients with MS.

Methods

Data Collection and Preprocessing

The study utilized accelerometer data to predict various physical activities. Data collection involved using accelerometer sensors to track movements across three axes (x, y, z). The raw data were preprocessed to ensure quality and usability. This preprocessing included normalizing the data for each axis using methods like mean subtraction, standard deviation normalization, skewness adjustment, and kurtosis correction. Specific windowing techniques were then applied to segment the continuous data stream into manageable frames for further analysis. Different window sizes (1, 3, and 5 seconds) and step sizes (creating overlaps of 1/16, 1/8, and 1/4 of the window size) were used to examine the effect of these parameters on the performance of activity prediction.

Feature Extraction

For each segmented window, various statistical features were extracted to characterize the data effectively. These features included the mean, standard deviation, zero-crossing rate, and dominant frequency components of the windowed segments. These metrics were chosen to capture the essence of the motion patterns in the data, providing a robust feature set for training the models.

Model Training

Multiple machine learning models were employed to predict activities from the processed segments, including K-Nearest Neighbors (KNN), Random Forest (RF), Long Short-Term Memory networks (LSTM), LightGBM, and Support Vector Machines (SVM). LSTM was used to leverage the ability to remember patterns over time, which is crucial for time-series data like accelerometer readings. RF and XGBoost are known for their robustness and ability to handle non-linear data distributions. Lastly, SVM and lightGBM explore different algorithmic approaches to handling feature-rich data. Each model was trained on the training dataset comprising 80% of the data, with the remaining 20% used for testing to evaluate model performance.

Model Evaluation

The effectiveness of each model was assessed based on accuracy, precision, recall, and F1-score. These metrics facilitated a comprehensive evaluation of each model’s ability to correctly classify activities and handle imbalanced data. The impact of window size and feature extraction methods on model performance was specifically analyzed to determine optimal strategies for activity prediction.

Results & Discussion

Initial Results
Before compiling all the results into a standardized environment, models were run separately from one other on the aforementioned windows with a 50% overlap and without statistical feature extraction to produce preliminary results. Based on the initial tests of the model, it was observed that most models performed decently on all three axes, with a general trend of worsening accuracy as the window size increased to 5 seconds. Below is a table containing the accuracy of the LSTM, RF, XGB, and SVM models in each window represented as a proportion.

Model	1 second	3 second	5 second
LSTM	.7177	.7011	.6331
RF	.9124	.8974	.8801
XGB	.8087	.8264	.6611
SVM	.9817	.9817	.9808

Figure 1.1

Due to the large number of data points and complexity of the models, running them on unoptimized, personal devices in a timely manner was unfeasible. Instead, the final environment was run on a HAL supercomputer by another teammate to produce standardized results of all the models on the windows and filters. However, due to insufficient memory and processing along with time constraints preventing looking into the issues, no full results were obtained. The results that were obtained were done on a one second window with 1/16 overlap with the mean filter and are shown in the table below.

Model	Accuracy	Precision	Recall	F1-Score
LSTM	.5100	.5056	.5100	.4997
RF	.4803	.4571	.4040	.4195
XGB	.4423	.4221	.4223	.4180
SVM	.9816	1.000	1.000	1.000
LightGBM	.4283	.4091	.4283	.4016

Figure 1.2

Discussion
As seen in Figure 1.1, the SVM model had the highest accuracy, peaking at 98.17% at one second, with RF, XGB, and LSTM following with decreasing accuracy. Additionally, it was noted how SVM trained and predicted much faster than the other models while also having the best accuracy, making it stand out as a strong contender for a good model on these preliminary highly accurate models, it was concluded that the ML classification of an activity based on accelerometer data had a strong foundation and possibility of success.

When looking at the results in Figure 1.2 and comparing them to those seen in Figure 1.1, it can be determined that the mean filter reduces the accuracy of our models significantly, with the exception of the SVM model that remains highly accurate. These results further support our initial findings of the SVM model being the most accurate and efficient. However, no further conclusions could be made about the other filters or windows from these results.

Future work could revolve around resolving issues with the HAL in order to run the intended experiment and the results would provide insights into how different models performed with different windows and features and which features produced the most accurate predictions. Additionally, automatic feature extraction from various models and techniques would be a topic for future research that would build on these results.

Significance:
In this study, we investigated the potential of wearable accelerometers combined with machine learning models to accurately classify various physical activities. By focusing on feature extraction from accelerometer data and employing a range of machine learning algorithms, we aimed to develop a robust activity recognition system. Our initial results showcased promising accuracy rates, with the Support Vector Machine (SVM) model emerging as a standout performer, achieving a peak accuracy of 98.17% on a one-second window with a 1/16 overlap, precision of 100%, recall score of 100% and F-1 score of 100%. These four indicators demonstrated a high classification efficacy of the SVM model.

However, due to overlapped data and computational constraints, comprehensive analysis and optimization of the models were hindered. The final results, obtained on a supercomputer, were limited, and further exploration of different window sizes, overlaps, and feature extraction methods was impeded. Despite this limitation, our findings underscore the potential of SVM and highlight the importance of considering filtering techniques, as observed with the mean filter's impact on model accuracy.

Moving forward, addressing computational challenges, and optimizing model performance remain key priorities. Future research should focus on resolving technical issues to conduct comprehensive experiments to gain a deeper understanding of model performance under different configurations. Moreover, exploring automatic feature extraction methods and expanding the study to include larger datasets, including individuals with medical conditions like multiple sclerosis, could provide valuable insights for clinical applications and telehealth support. Overall, this research lays the foundation for leveraging wearable accelerometers and machine learning for activity classification, with the potential to enhance healthcare monitoring and disease management.

References: [1] Keszler et al. (2022). *International journal of MS care*, *24*(6)