The Role of Machine Learning in Improved Functionality of Lower Limb Prostheses

Joaquin Dominguez[1], Richard Kim[1], Robert Slater[2]
[1] Dallas, TX 75275 USA
[2] SMU MSDS Faculty (Adjoint Professor), Southern Methodist University, Dallas, TX 75275 USA
{joaquind, richardk, rslater}@smu.edu

Abstract. Lower-limb amputations can cause a plethora of obstacles that lead to a lower quality of life. Implementing machine learning techniques means advanced prosthetics can contribute to facilitating the lives of those that live with lower-limb amputations. Using the publicly available HuGaDB data set, the current study investigates several classification models (random forest, neural network, and Vowpal Wabbit) to predict the locomotive intentions of individuals using lower-limb prostheses. The results of this study show that the neural network model yielded the highest accuracy, comparable precision, and recall scores to the other models. However, the Vowpal Wabbit model's advantage in speed may allow for other, more practical implementations in practice. These findings provide insight into the advantages of specific classification models over others in predicting the intentions of specific movements during locomotive transitions. These findings present direct comparisons of several machine learning methods, identifying the strengths and weaknesses of each classification model tested.

1 Introduction

Lower-limb Amputations (LLA) are among today's most common amputations associated with high mortality rates and lower quality of life (Robbins et al., 2008). While research has identified risk factors and demonstrated interventions to prevent the need for amputations in the first place (Rathinam et al., 2021), there continues to be a relatively high incidence today, with nearly 185,000 new amputations in the United States each year (Kozak et al., 1995). Consequently, various prostheses have been developed to improve the mobility of individuals living with LLA.

Within the non-activity-specific prostheses, there exist passive and active models. Passive models may include a hydraulic system and an endoskeleton torsion unit to dampen impact forces and bear weight while twisting. Active models, in addition, may contain microprocessor-controlled ankle/foot systems and knee systems that provide active propulsion to lower the amount of effort people may need to use while walking (Stokosa, 2021).

Another element that may considerably influence the quality of life is the fit and quality of prostheses. Less-than-ideal fitting and poorly made prostheses have introduced issues, such as restricted movement, pain with continued use, and a general risk of falling (Bryant, 2019). More recent advances have allowed prostheses that focus on providing solutions to these issues. Medical ultrasonography has been most recently explored as a possible solution to compensate for limb motion in prosthetic sockets (Ranger et al., 2015). Another recent alternative, with considerably less accessibility but high efficacy, is MIT's FitSocket. This tool accounts for leg tissue properties via multiple linear actuators to detect areas of softness and stiffness in a patient's residual limb (Petron, 2016). But even with these advances, further tuning is warranted to improve functionality and quality of life. A dearth of research focuses on user feedback and satisfaction (Eshraghi et al., 2013). Thus, improvements in prostheses should consider qualitative outcomes in conjunction with any quantitative findings.

The current literature on advanced prosthetics functionality has largely explored the use of interfaces to control prostheses—either via neural or muscular implementation. However, the advances in microprocessor technology have facilitated the application of machine learning models to aid active prostheses. Implementing this technology could potentially supplement or even circumvent the invasive elements involved with the use of interfaces.

This study aims to apply machine learning techniques to improve comfort and mobility for individuals using active lower extremity prostheses. The various ways this objective has been explored in the literature are detailed below. The current state of development for active prostheses has been fruitful but remains in its nascent stages by its limited accessibility and functionality.

One primary area of focus is improving action recognition (i.e., classification of actions). Among the most used machine learning methods for this objective are Random Forest (RF), Neural Networks (NN), k-Nearest Neighbors (k-NN), Support Vector Machines (SVM), and Decision Trees (DT). To this end, the most relevant metric to evaluate is the accuracy of action classification, for which the best method is highly contingent on the data set.

This study was limited to publicly available data sets. Among those that provided relevant information about the topic at hand, the Human Gait Database (HuGaDB) contained the most comprehensive documentation and sensory inputs. It uniquely employs inputs from an accelerometer, gyroscope, and EMG-- each placed on the thigh, shin, and foot of each leg, amounting to 6 total positions of the body. The data was collected from 18 participants and had classifications for 12 activities, ranging from 'walking' to [going] 'up by elevator' (Chereshnev et al., 2018).

The current benchmark model for this dataset holds a classification accuracy of 98.6% by using a Random Forest algorithm in conjunction with specific data processing and segmentation methods (Badawi et al., 2018). While the accuracy is 98.6%, the nature of the algorithm is limited in scalability for this application. Random Forest requires renewed training of a model to incorporate novel data, essentially creating an entirely new model. An adaptive prosthesis must be able to augment its model without needing the computing resources required to retrain a new/altered data set due to the limited capabilities of an integrated microprocessor. Thus, it is recommended to increase the metric performance score and find an alternative method of modeling that would allow for adaptability for the individualization of functionality. This study examines the possibility of using a neural network framework to tackle this problem, capitalizing on its layering approach, as well as a Vowpal Wabbit (VW) model, which is widely used for its efficiency with extensive data. VW models can be understood as highly optimized linear regressions with the advantage of streaming data.

2 Literature Review

2.1 Quality of Life with Prostheses

Several studies have looked at the quality-of-life aspect of prosthetics. At the same time, much of the literature pertains to functionality, such as accurate detection of motion and recognition of intent. There is a need to address research conducted with a primary interest in the comfort and well-being of the user.

One study administered various questionnaires to lower limb amputees that measured their prostheses' quality of life, functional performance, and body image (Burcak et al., 2021). Results supported higher overall satisfaction, functional performance, and better body image with high-tech prostheses over mechanically controlled prostheses. Although this study specifically used microprocessor-controlled knee prostheses (MPK) and transtibial vacuum-assisted suspension system (VASS) prostheses, these findings still support the advantages of applying more advanced prosthetics and the need to make them more accessible to amputees.

Of course, even before prosthesis usage, the preliminary step of fitting prostheses is a critical step that may determine the user's comfort. Jamaludin and colleagues (2019) investigated this topic by attempting to estimate pressure distribution profiles. Understanding this component can aid in determining a user's level of comfort before fabricating the prosthetic socket. The researchers demonstrated the potential for three-dimensional models using magnetic resonance (MR) images (Jamaludin et al., 2019). These 3D models provided pressure measurement and mapping within the socket, allowing for adjustments tailored to each user.

2.2 Non-Machine Learning Techniques

Most studies investigating functionality in lower-limb prosthetics used some form of machine learning or explored various methods to use alongside machine learning. However, many studies have investigated the feasibility of using non-machine learning techniques. Many machine learning models are run on data collected from sensors, and while this works well in practice, studies have underlined the difficulty of retrieving clean data in the real world (Afzal et al., 2016; Dyson et al., 2020). Consequently, this invariably leads to decreased performance scores. Optimal results would then require retraining the model every several days, if not daily, which would be costly and realistically impractical. Although this problem applies to both machine and non-machine learning techniques, non-machine learning approaches may be computationally and financially less expensive.

With this goal in mind, Dyson and colleagues (2020) developed a paradigm that relies on motor learning instead of a machine learning algorithm. Although this study focused on upper limb prostheses, it offers a novel approach to circumvent a common issue with machine learning algorithms, generalizability. Machine learning models must train on EMG data that reflects people's daily encounters. If a new desired output exists, the model must be retrained to account for this new classification. The study initially looked at how human learning may serve as a viable alternative to machine learning. People, both limb-intact and limb-deficient, using human-machine interfaces, could control myoelectric interfaces using muscles in their arms. By training the muscles directly, this approach allows for more flexibility with its application to the real world.

Stepping beyond applicability advantages alone, another study demonstrated a non-machine learning method with improved performance over machine learning (Afzal et al., 2016). In their study, researchers highlighted the potential advantages of using muscle synergy over standard machine learning techniques, such as neural networks and linear discriminant analysis. The differences in performance varied depending on the type of machine learning technique applied and the same locomotive task. However, the results still support the argument that there are viable alternatives to machine learning.

Unfortunately, the non-machine learning approaches stated concerns regarding the quality of EMG signals, which is also a common concern for machine learning models. Data collected from poor EMG readings were generally excluded or accounted for before analysis. This points to a fundamental problem with the practicality of high-tech prostheses outside a controlled environment. The shortcomings of any advanced model appear to come not from the limitations of the technique employed but rather from the data to which the model is being applied. While this discussion reaches beyond the scope of the current study, it is important to note that the results of any classification technique will be contingent upon the quality of the data.

2.3 Machine Learning Techniques

In most studies employing machine learning techniques, classification appears to be the predominant objective. The schematic process generally consists of identifying movements of the person using the prosthesis or identifying surroundings via sensors–presumably, with the intent of having the respective classification be followed by a corresponding assisted movement. To this end, Griffiths et al. (2021) used data from an accelerometer placed at the shank level to classify and predict postures using a Random Forest algorithm with an accuracy of 93%. In another study, Lee et al. (2021) attempted to predict the gait phase by taking the angular positions and velocities of the thigh, torso, and heel force as input. As a caveat, however, it is important to consider how a model can be built while minimizing the additional hardware imposed on amputees. In the case of Lee, although highly effective concerning classification metrics, it appears unfeasible to apply multiple Inertial Measurement Units (IMUs) outside of the prosthesis as part of a quotidian application.

Apart from using sensory inputs, another approach explored using a Brain-Computer Interface (i.e., neural inputs) to classify and interact with prostheses (Dillen et al., 2022). Dillen and colleagues decoded EEG data with an accuracy of 84% to discriminate different lower extremity movements. With further development and more elaborate machine learning modeling, this approach can be effective in isolation or supplemental to sensory inputs. It must be noted that adding a BCI can prove invasive or financially unfeasible. A practical use case for this approach can be found in patients with Spinal Cord Injury (SPI), whereby the condition leaves them unable to ambulate. Wang et al. (2012) similarly used EEG and ML methods to decode said EEG input for eight patients with SPI–of note. However, in contrast to Dillen and colleagues (2022), Wang and colleagues built a Virtual Reality Environment after a 10-minute training to allow control over the ambulation of an avatar in real-time, essentially replicating the real-world output if a prosthetic responded immediately to neural inputs. This begets the possibility of integrating an autonomous element into prostheses' active-assistance movement.

The use of neural networks (NN) algorithms, specifically in the application of active prostheses, is relatively sparse in the literature, partly due to its recent popularity in the data science field and complexity in structure and interpretability. Still, as Nayak & Das (2020) emphasized, its application may have "a huge impact in achieving independent mobility and enhances the quality of life in Persons with Disabilities" (p. 17). There have been some notable publications, however, such as Karlik et al. (2003) 's seminal work applying a NN architecture to classify myoelectric signals. Since then, the technology (both in computing and robotics/prosthetics) has allowed for the refinement and application of their methods, as proposed in their original paper. Another study outlines the use of NN to determine EMG sensor requirements from gait analysis data. Of particular interest is the nature of the data used for modeling, which is highly applicable to LLA cases, and their objective to minimize the use of inputs to have a possible use case for integration in active prostheses (Keres, 2017).

Outside of these cases, studies have used other ML techniques, such as Random Forest (RF), XGBoost, and Logistic Regression, to identify relationships between amputations and medical comorbidities. Bolourani et al. (2020) combined Logistic Regression with sampling algorithms to identify patients with traumatic arterial injuries at elevated risk for amputation with an 88% accuracy. Alternatively, Anderson et al. (2021) used a neighbors-based approach to predict an individual patient's future mobility after a lower-leg amputation. Though not directly related to the prosthesis, this approach can still be implemented as a supplemental overview of the patient's case and may provide utility for deriving a patient's best course of action.

2.4 Supplemental Methods for Machine Learning

A body of literature demonstrates methods that can work in conjunction with machine learning techniques. While some of these methods will not be directly applicable to our course of research, it is useful to contextualize how ML methodology has been used to study different facets of this field. For those studies that used machine learning models to decode EMG signals, it must be noted that said models and systems are prone to suffer from issues with robustness and stability of accuracy over time (Schulte et al., 2022). This effect, called concept drift, is primarily due to a changing input signal and was found to affect pattern recognition by 20%-30% throughout the day or between days. The feasibility of this technology for long-term use is contingent on frequent retraining of myoelectric pattern recognition systems, which makes its application in prostheses impractical, at least in its current state. Another example of varying applications within an ML framework can be seen with Lee et al. (2021), wherein a different combination of sensors–concerning body placement–was tested and compared. In the first set (S1), sensors were able to pick up the velocity and angular position of the thigh and torso, and in the second set (S2), the heel force sensor was added to the previous two. S1 allows for greater ease of use and thus facilitates training and testing, compared to S2; however, as previously mentioned, the use of additional sensors outside of the prosthesis limits its practical application outside of a controlled study.

To address the issue of external sensors, Griffiths et al. (2021) created an ML model using a sensor at the shank-level, with the implication that said sensor could be integrated into the active-prosthesis system. This method creates its own set of limitations, given the capacity a sensor has in that position to gather relevant data. Such determination may only be made considering the patient's needs–to wit, if the type of active assistance in a prosthesis (and its respective model) does not require additional data to perform efficaciously, then the number and locations of sensors will suffice in this form.

Another parameter that may be implemented in developing ML models, as conveyed by Lecomte et al. (2020), is the comparison of a functional joint center (FJC) and a conventional ankle joint. The FJC can help characterize and differentiate between various prosthetic foot designs. This could allow for a better fitting for amputees, which has been a significant source of discomfort.

It is hypothesized the neural network model will yield the best overall performance metrics, leading with the highest accuracy and higher precision and recall for most activities. Moreover, its implementation in deployment would allow edge computing to individualize the classification model by adding an extra hidden layer of training data derived from the end user's specific movements. A secondary hypothesis is that the VW model will train the data fastest and achieve results comparable to or better than the random forest classifier. This study aims to add to the literature by:

1) Replicating the current benchmark model for this dataset in the literature (Badawi et al., 2018).
2) Exploring model metric scores of a neural network framework and VW model for individualization of functionality.
3) Assessing each model as a viable method to implement in active prostheses, considering their ability to adapt to each potential user.

3 Methods

The HuGaDB dataset was collected from 18 participants, four females and 14 males. Participants were identified as healthy adults with an average age of 23.67 (+/- 3.69σ) years and an average weight of 73.44 (+/-16.67σ) kg. Data was gathered using inertial sensors that acted as accelerometers and gyroscopes, tools used to collect information on linear acceleration and rate of rotation. Sensors were placed on each leg's thigh, shin, and foot for a total of six sensors. Each sensor collected acceleration and gyroscope data along the x, y, and z axes, yielding 36 features across all six sensors. Muscle activity was captured using two EMG sensors on the quadriceps.

After adding the participant ID, the complete data set comprised 38 explanatory features across 2,111,869 samples. The target variable was a classification of the activity being performed when the data was recorded. The activities ranged from basic actions, such as 'walking,' to more unique, activity-specific actions, like 'sitting in a car,' giving twelve unique responses. Across all participants, data were acquired over a collective total of 10 hours of activity. It is important to note that Chereshnev and colleagues state in their original paper for this data set that some corrupt data had to be discarded. They also state that the raw data were filtered with moving averages to "remove the bias drift of [their] inertial sensors." No further data cleaning or transformations were conducted beyond retrieving the data from the public GitHub repository.

Using the moment-to-moment sensor data, data were trained and tested using three different types of machine learning models to classify these actions: a random forest classifier, a VW algorithm, and a neural network model. Each model's performance was evaluated using their overall accuracy score first, as a general comparative metric and then for precision, for each class. Keeping in mind the end-user's functionality and safety, this model needs to maximize the degree to which it correctly identifies a given movement, as it is necessary to minimize the proportions of which a given class is incorrectly classified. Although our primary goal in assessing these models is to identify the one with the highest accuracy, given that these models are intended to optimize daily movement with prosthesis use, they are also crucial to yield low Type I and Type II errors. Neglecting these error rates in our model selection could bring unintended results upon implementation, such as improvement with specific activities at the cost of added discomfort in others. As such, accuracy scores were supplemented by recall and precision scores. These additional metrics could then help us identify if there are any preferences in actions between models; models may have different tendencies to identify one action over another.

Lastly, the performance speed was monitored when training for each model. In practice, users will need to train a classifier for each fitted with a prosthesis, possibly using more than 10 hours of data per person. With the added possibility of retraining the data over continued use, the practicality and financial viability of these models may rely heavily on the speed of performance. While the exact method of implementation and dissemination is beyond the scope of this study, assessing the performance speed serves as a gauge for what should and should not be a realistic expectation for each model.

3.1 EDA

The raw data comprised 2,111,869 observations across 40 features and 1 target variable. The target variable was the type of physical activity being performed, such as 'walking'. Explanatory features were accelerometer, gyroscope, and EMG data. The output values of the 18 accelerometer and 18 gyroscope features mostly ranged between –32,768 and 32,767. Six observations contained accelerometer or gyroscope values far outside this range and were excluded as outliers. The left and right EMG features had values ranging from 0 to 254. The id feature had values 1 through 18, each number representing one of the eighteen participants for the study. The target variable was a multiclass feature with 12 unique physical activities.

The modified data set used for analyses comprised 2,111,863 observations across 40 features and 1 target variable. Groups within the target variable were imbalanced; over 30% of observations were classified as 'walking', while other classifications like 'down by elevator' and 'sitting in car' only composed around 1% of the data. This imbalance was not addressed during data preprocessing but will be important to note when interpreting the results.

Table 1. Number of samples per activity in the complete dataset

Activity	Observations
Walking	679,073
Running	328,655
Going up	241,756
Going down	180,573
Sitting	156,560
Sitting down	131,604
Standing up	116,637
Standing	89,144
Bicycling	71,653
Up by elevator	69,729
Down by elevator	24,112
Sitting in car	22,373

3.2 Random Forest

The Random Forest algorithm was chosen due to its efficacy in classification modeling within the data science field, in addition to its lack of complexity in modeling and interpretability. It also serves to replicate the benchmark accuracy achieved by Badawi et al. (2018). Both machine learning models were prepared by and through Python's *scikit-learn* packages and trained with an 85-15 train-test split, with a stratified split to reflect the class distribution of the target variable. As outlined in Badawi (2018), the number of estimators was set to 256, and the split criterion was based on Gini impurity.

3.3 Neural Networks

Given its current widespread application in the data science field, a neural network model was also trained to identify its advantages, if any, in recognizing locomotive intentions. The neural network was developed using Python's *tensorflow* sequential model, creating a linear stack of five layers. The five layers had 300, 400, 500, 300, and 13 (output) neurons. The model utilized '*categorical_crossentropy'* as the loss function and '*adam'* as the optimizer with a *learning_rate* of 0.0001. A Dropout layer was added as the last hidden layer with a rate of 0.2 in order to prevent overfitting. Early stopping was implemented with patience set at 4 and mode set at '*max*' to monitor '*val_accuracy*.' A description of the model architecture is found in the table below:

Table 2. Architecture of Neural Network Model

Layer	Neurons	Activation
Input	39	N/A
Dense	400	Sigmoid
Dense	400	Sigmoid
Dense	400	Sigmoid
Dense (Output)	13	Softmax

3.4 Vowpal Wabbit

Finally, a VW model was trained, a system well-known for its advantage in speed by streaming data and using what is known as the hashing trick. The model was developed using the vowpal package in Python. The model was trained with 1000 '*passes'* using the 'hinge' loss function and *'oaa'* (One Against All) comparisons. The '*l1'* and *'l2'* regularizations were set to 0, and quadratic and cubic functions were created using *'-q::'* and '*--cubic:::'*, respectively. It is important to note that VW is likely faster to train and optimize using the command line. The model created for this study used the vowpal package in Python to keep conditions consistent across all models for ease of comparison.

4 Results

The performance of the three models was compared using three common evaluation metrics: accuracy, precision, and recall. The neural network model achieved an accuracy of 94%, a precision of 95%, and a recall of 93%. The random forest model, on the other hand, achieved an accuracy of 96%, a precision of 97%, and a recall of 94%. Lastly, the VW model had an accuracy of 56%, a precision of 80%, and a recall of 91.5%.

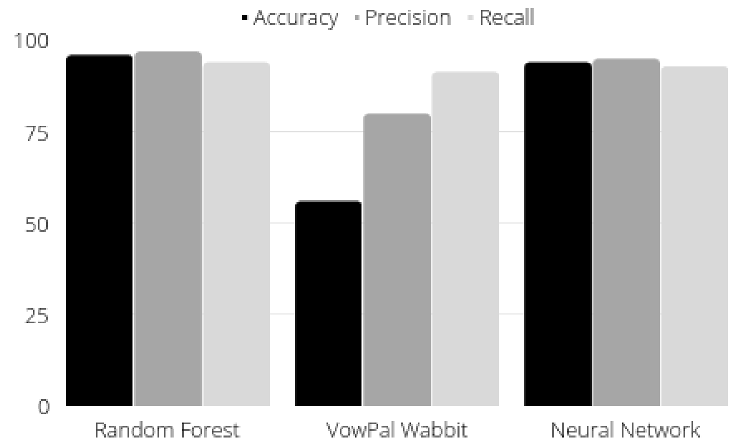

Fig. 1. A chart comparing the three models by Accuracy, Precision, and Recall

As shown in Figure 1, the results of this study favored the neural network and random forest models, with the random forest model having slightly better performance than the neural network model. The high accuracy and precision values suggest that the models were able to correctly classify almost all the instances in the test set. The recall values indicate that the models were able to identify a significant proportion of the positive instances in the test set.

Table 3. Accuracy, precision, recall, and training speed of each model

Model	Accuracy	Precision	Recall	Speed
Random Forest	96%	97%	94%	40 minutes
Neural Network	94%	95%	93%	8 minutes
Vowpal Wabbit	56%	80%	91.5%	16 minutes

Models were also evaluated on speed of training. This was accomplished by running all models in a standard GPU runtime Google Colab environment. The neural network model took 8 minutes, the random forest model 40 minutes, and the VW model 16 minutes. The reason for the random forest's significant training time can be attributed to the sheer number of estimators that was included in the model. Using GridSearch, the random forest model was optimized for performance, which resulted in a total of 256 estimators. This could be a significant factor to consider when choosing a model for practical applications, particularly when working with large datasets. In the predicted use case for our study, training time and resources are of high importance, since the model would be fine-tuned to each user's specific gait in order to improve performance.

Table 4. Random Forest Precision and Recall Scores by Class_id

Class_id	Precision	Recall
1 (Walking)	0.984	0.978
2 (Bicycling)	0.976	0.980
3 (Sitting Down)	0.947	0.963
4 (Standing Up)	0.934	0.958
5 (Going Up)	0.989	0.996
6 (Down by Elevator)	0.967	0.910
7 (Sitting in Car)	0.954	0.887
8 (Running)	0.897	0.947
9 (Sitting)	1.000	0.999
10 (Standing)	0.796	0.763
11 (Up by Elevator)	0.850	0.673
12 (Going Down)	1.000	1.000

Table 5. Neural Network Precision and Recall Scores by Class_id

Class_id	Precision	Recall
1 (Walking)	0.981	0.976
2 (Bicycling)	0.957	0.976
3 (Sitting Down)	0.946	0.956
4 (Standing Up)	0.931	0.941
5 (Going Up)	0.986	0.994
6 (Down by Elevator)	0.942	0.885
7 (Sitting in Car)	0.921	0.899
8 (Running)	0.897	0.895
9 (Sitting)	0.999	0.998
10 (Standing)	0.772	0.565
11 (Up by Elevator)	0.712	0.638
12 (Going Down)	0.999	1.000

Additionally, all models were most accurate when identifying 'walking', their precision and recall scores ranging between 97% and 100%. The neural network model achieved a precision of 97.3% and a recall of 98.2% for ‘walking”. These results were not surprising, given that 'walking' comprised nearly one-third of the true responses in the entire data set. Precision and recall scores for the random forest and neural network model are presented in Table 4 and Table 5, respectively. Overall precision and recall score comparisons are shown in Figure 2.

Fig. 2. Neural Network Precision and Recall Scores by Class_id

In conclusion, both the neural network and random forest models proved to be effective in this classification task, substantively above the VW model. Between the top performing models, the random forest showed slightly better performance and the neural network was faster to train. Further analysis and experimentation may be necessary to determine the best model for the specific data and problem at hand.

5 Discussion

The main objective of this study was to investigate the most effective way for a user of a prosthesis, specifically a lower-leg amputee, to benefit from a machine learning application through their prosthesis. Specifically, we aimed to assess and compare three machine learning classification methods: a random forest, a neural network, and a VW model across multiple criteria. To compare the three models, we assessed their effectiveness based on the evaluative metrics, accuracy, precision, and recall. We also compared the training time to determine ease of model customization and feasibility of using these models to fit prostheses for individuals.

The evaluative metrics used in this study—accuracy, precision, and recall—provided confidence in how well the models would perform when applied to prostheses in the real world. The results suggest that the neural network and random forest models are capable of accurately classifying human gait based on the inputs provided, yielding scores in the mid- to upper- 90% range. While these models fell short compared to the benchmark model in the original paper, both models performed much better than the VW model, which yielded an accuracy of 56%. These results indicate that VW may not be appropriate for the sensory data used in this study or that further fine-tuning is necessary. It is also possible that this model may see improvements after applying other methods, such as feature engineering.

We also examined the training time, taking into consideration the feasibility of training these models as part of the process of fitting a prosthesis. The neural network model had the fastest training speed at 8 minutes, VW trained in 16 minutes, and the random forest yielded the longest training time at 40 minutes. The considerable difference in training time is important to note, given that the additional time it takes to train a model on each user becomes a gauge of how expensive this process could become. At five times the speed of the random forest, the neural network model appears to be most feasible in implementation. This also implies that the random forest model, while competitive in its evaluative performance metrics, may not be a viable option in practice.

Taking all these results into consideration, the tuned neural-network model proved to be most useful for this problem. In making this determination, we considered and prioritized autonomy of the user by upholding precision with a higher weight in our evaluation and used training time as a measure of feasibility in practice. We further clarify that this selection is specific to the application outlined in this study and not a blanket assessment of the efficacy of the algorithms used for this dataset and/or datasets of this nature. Further experiments and analysis may be necessary to determine the best algorithm for a specific dataset and problem.

The results of this study indicate that the models are effective in the classification and prediction of movements but have yet to be applied to the movements of a lower-leg amputee using a prosthesis. We expect there to be a small difference in gait and thus in performance of the models; with this in mind, the use of layer architecture inherent in the neural network algorithm allows for an extra layer with data specific to the user to be integrated into the model, thereby fine-tuning it to its specific application. Furthermore, the findings of this study have important implications for the development of technology for people with physical disabilities, such as lower-leg amputees, who may benefit from the improved mobility and quality of life provided by these models.

The limitations of this study include the inability to conduct proof-of-concept due to time and funding constraints. Although the models achieved a high accuracy in detecting gait activities, these models were not tested on actual prostheses. The available data was also gathered from non-disabled subjects; thus, these results being generalized to individuals with lower-limb amputations should be interpreted with caution. It is also important to note that it may be impractical to require the end-user to integrate 6 different sensors in 6 different areas of the body, as was done to create the dataset used in this study. The use of external sensors (with respect to the prostheses) may, generally, not be recommended, as this would require additional elements to be created, integrated, and maintained. For lower-leg amputees, we surmise that two internal sensors may be plausibly used as an input mechanism for the model with satisfactory results. Further investigation would be required to fine-tune the sensors and model to this end. Given that these models were trained on publicly available data, there is also a restriction in terms of which variables can be included in building the model. While the results of this study favor the neural network model, future research should aim to overcome these limitations before fully integrating the models into a device for practical use.

The ethical considerations of this study are important to ensure that the technology discussed in this study is used in a responsible and appropriate manner. Although the machine learning models are intended to improve the mobility and quality of life for individuals with lower-extremity prostheses, there is a need to address multiple factors that have historically affected those who are disabled--specifically, in matters of autonomy, privacy, and fairness.

In recent decades, there have been significant advances in the rights of people with disabilities, such as with the passing of the Americans with Disabilities Act (ADA)(1990) and the Convention on the Rights of Persons with Disabilities (CRPD)(2006). Such developments have helped promote greater accessibility and inclusion for people with disabilities and have given them a stronger voice in shaping the policies and technologies that affect their lives.

At the same time, advances in assistive technology, such as active prostheses, have improved the mobility and independence of people with physical disabilities. These technologies have also helped to break down barriers to full participation in society, allowing people with disabilities a greater opportunity to live, work, and participate in their communities. However, there have been numerous cases where technological advances and implementations have discriminated against people with disabilities.

It is, thus, essential that novel technological advances and implementations respect the autonomy and dignity of the intended users. Said principle holds an even higher burden of consideration when the use-case is intended for a marginalized group, such as people with disabilities. For the case highlighted in this study, the end goal of developing these machine learning models is to adjust contact points between the prosthesis and the individual's leg. It is important that the collected sensory inputs and the adjustments made to those contact points do not infringe on the privacy of the user. It is also crucial that the model is trained in a way that respects the autonomy of the user and allows them to control their own movement and actions.

The models' functions are also centered around collecting sensory input data, measures of individual gaits. Training these models requires a lot of data directly related to the user's gait. The model's usage is completely dependent on continuously obtaining gait information, namely sensory input, to classify their activities at every moment of their daily lives. Given that gait is unique to each individual and noting its close connection to each user's identification, this information should be handled with caution. Any attempts to use this data as personal identifying information should be accounted for and prevented.

Additionally, it is important that the implementation and application of these models and prostheses are not biased or unfair with respect to the end users. There is a need to recognize that tailoring each model to each user can lead to discrimination, either with respect to that individual or groups of people. Each model should be trained and optimized for each user without bias. Use of these models must ensure that there is no discrimination against individuals or people of different backgrounds and that this technology is not used to promote discrimination of any kind. The collected data, trained model, and tailored prosthesis should only be used to improve the mobility and well-being of the end-user.

This study provides important insights into the development of ML models for the control of active prostheses for lower-leg amputees. The results suggest that both the neural network and random forest models are effective in the classification of human gaits and have important implications for the development of technology for people with physical disabilities. Further research is necessary to fully realize the potential of these models and make them available and accessible to those who can benefit from them.

6 Conclusion

Despite lower-limb amputations being one of the most common forms of amputations today, they remain one of the most negatively influential on quality of life for amputees. Multiple factors contribute to this effect, including the fitting and quality of the prosthesis, as well as its functionality. With regards to functionality, prostheses today are found in either passive or active models, with active models being powered in some form via a microprocessor. While the use of interfaces and sensors to operate active prostheses has been largely explored in the literature, machine learning methods and the advancement of microprocessors offers an opportunity to integrate adaptive/learning functions into these units.

Using the HuGaDB dataset, which provides comprehensive data about human gait for activity recognition, we explored several models well known for yielding high accuracy, including the random forest, neural network, and VW model with the ability to be adaptive to an individual amputee's specific movements. The objective of training these models is to aid mobility and individualized functionality. Findings from this study provide support for the practical applicability of machine learning in facilitating movement using prostheses.

References

1. Afzal, T., Iqbal, K., White, G., & Wright, A. B. (2017). A Method for Locomotion Mode Identification Using Muscle Synergies. IEEE Transactions on Neural Systems and Rehabilitation Engineering, 25(6), 608–617. https://doi.org/10.1109/TNSRE.2016.2585962

2. Anderson, C. B., Wurdeman, S. R., Miller, M. J., Christiansen, C. L., & Kittelson, A. J. (2021). Development of a physical mobility prediction model to guide prosthetic rehabilitation. Prosthetics & Orthotics International, 45(3), 268–275. https://doi.org/10.1097/PXR.0000000000000001

3. Badawi, A. A., Al-Kabbany, A., & Shaban, H. (2018). Multimodal Human Activity Recognition From Wearable Inertial Sensors Using Machine Learning. IEEE-EMBS Conference on Biomedical Engineering and Sciences (IECBES), 2018, pp. 402-407, doi: 10.1109/IECBES.2018.8626737.

4. Bolourani, S., Thompson, D., Siskind, S., Kalyon, B. D., Patel, V. M., & Mussa, F. F. (2021). Cleaning Up the MESS: Can Machine Learning Be Used to Predict Lower Extremity Amputation after Trauma-Associated Arterial Injury? Journal of the American College of Surgeons, 232(1), 102-113e4. https://doi.org/10.1016/j.jamcollsurg.2020.09.014

5. Bryant, M. (2019, March 29). How AI and Machine Learning Are Changing prosthetics. MedTech Dive. Retrieved October 2, 2022, from

https://www.medtechdive.com/news/how-ai-and-machine-learning-are-changing-prosthetics/550788/

6. Burçak, B., Kesikburun, B., Köseoğlu, B. F., Öken, Ö., & Doğan, A. (2021). Quality of life, body image, and mobility in lower-limb amputees using high-tech prostheses: A pragmatic trial. Annals of Physical and Rehabilitation Medicine, 64(1), 101405. https://doi.org/10.1016/j.rehab.2020.03.016

7. Chereshnev, R., Kertész-Farkas, A. (2018). HuGaDB: Human Gait Database for Activity Recognition from Wearable Inertial Sensor Networks. In: , et al. Analysis of Images, Social Networks and Texts. AIST 2017. Lecture Notes in Computer Science(), vol 10716. Springer, Cham. https://doi.org/10.1007/978-3-319-73013-4_12

8. Dillen, A., Lathouwers, E., Miladinović, A., Marusic, U., Ghaffari, F., Romain, O., Meeusen, R., & De Pauw, K. (2022). A data-driven machine learning approach for brain-computer interfaces targeting lower limb neuroprosthetics. Frontiers in Human Neuroscience, 16, 949224. https://doi.org/10.3389/fnhum.2022.949224

9. Dyson, M., Dupan, S., Jones, H., & Nazarpour, K. (2020). Learning, Generalization, and Scalability of Abstract Myoelectric Control. IEEE Transactions on Neural Systems and Rehabilitation Engineering, 28(7), 1539–1547. https://doi.org/10.1109/TNSRE.2020.3000310

10. Eshraghi, A., Osman, N. A., Gholizadeh, H., Ali, S., & Shadgan, B. (2013). 100 top-cited scientific papers in limb prosthetics. BioMedical Engineering OnLine, 12(1), 119. https://doi.org/10.1186/1475-925X-12-119

11. Griffiths, B., Diment, L., & Granat, M. H. (2021). A Machine Learning Classification Model for Monitoring the Daily Physical Behaviour of Lower-Limb Amputees. Sensors, 21(22), 7458. https://doi.org/10.3390/s21227458

12. Jamaludin, M. S., Hanafusa, A., Shinichirou, Y., Agarie, Y., Otsuka, H., & Ohnishi, K. (2019). Analysis of Pressure Distribution in Transfemoral Prosthetic Socket for Prefabrication Evaluation via the Finite Element Method. Bioengineering, 6(4), 98. https://doi.org/10.3390/bioengineering6040098

13. Karlik, B., Osman Tokhi, M., & Alci, M. (2003). A fuzzy clustering neural network architecture for multifunction upper-limb prosthesis, in IEEE Transactions on Biomedical Engineering, 50(11), 1255-1261, doi: 10.1109/TBME.2003.818469.

14. Keles, A. D., & Yücesoy, C. (2017). Development of Artificial Neural Network Based Active Ankle Prosthesis Algorithm Using Gait Analysis Data. 21st National Biomedical Engineering Meeting (BIYOMUT), 2017, pp. i-iv, doi: 10.1109/BIYOMUT.2017.8478881.

15. Kozak LJ, Owings MF. Ambulatory and inpatient procedures in the United States, 1995. National Center for Health Statistics. Vital Health Stat 13(135). 1998.

16. Lecomte, C., Starker, F., Guðnadóttir, E. Þ., Rafnsdóttir, S., Guðmundsson, K., Briem, K., & Brynjolfsson, S. (2020). Functional joint center of prosthetic feet during level ground and incline walking. Medical Engineering & Physics, 81, 13–21. https://doi.org/10.1016/j.medengphy.2020.04.011

17. Lee, J., Hong, W., & Hur, P. (2021). Continuous Gait Phase Estimation Using LSTM for Robotic Transfemoral Prosthesis Across Walking Speeds. IEEE Transactions on Neural Systems and Rehabilitation Engineering, 29, 1470–1477. https://doi.org/10.1109/TNSRE.2021.3098689

18. Nayak, S., & Das, R. K. (2020). Application of Artificial Intelligence (AI) in Prosthetic and Orthotic Rehabilitation. In V. Sezer, S. Öncü, & P. B. Baykas (Eds.), Service Robotics, 17. IntechOpen. https://doi.org/10.5772/intechopen.93903

19. Petron, A. J., Prosthetic Socket Design: From a Multi-Indenter Device for in vivo Biomechanical Tissue Measurement to a Quasi-passive Transtibial Socket Interface. Thesis: Ph.D, Massachusetts Institute of Technology, School of Architecture and Planning, Program in Media Arts and Sciences.

20. Ranger BJ, Feigin M, Pestrov N, Zhang X, Lempitsky V, Herr HM, Anthony BW. Motion compensation in a tomographic ultrasound imaging system: Toward volumetric scans of a limb for prosthetic socket design. Annu Int Conf IEEE Eng Med Biol Soc. 2015 Aug; 2015:7204-7. doi: 10.1109/EMBC.2015.7320054. PMID: 26737954.

21. Rathinam, Dr. S., Mp, Dr. K., Imram, Dr. M., & Lakshmi, Dr. S. (2021). Study on prevalence and the contributing factors of lower limb amputation in a tertiary health care centre. Journal of Case Reports and Scientific Images, 3(1), 08–11. https://doi.org/10.22271/27080056.2021.v3.i1a.24

22. Robbins, J. M., Strauss, G., Aron, D., Long, J., Kuba, J., & Kaplan, Y. (2008). Mortality Rates and Diabetic Foot Ulcers. Journal of the American Podiatric Medical Association, 98(6), 489–493. https://doi.org/10.7547/0980489

23. Schulte, R. V., Prinsen, E. C., Buurke, J. H., & Poel, M. (2022). Adaptive Lower Limb Pattern Recognition for Multi-Day Control. Sensors, 22(17), 6351. https://doi.org/10.3390/s22176351

24. Stokosa, J., (2022). Limb Prosthesis Preparation. Merck Manual Professional Version. Accessed from: https://www.merckmanuals.com/professional/special-subjects/limb-prosthetics/options-for-limb-prostheses.

25. Wang, P. T., King, C. E., Chui, L. A., Do, A. H., & Nenadic, Z. (2012). Self-paced brain-computer interface control of ambulation in a virtual reality environment (arXiv:1208.6057). arXiv. http://arxiv.org/abs/1208.6057