Detecting Parkinson's Disease Using Machine Learning from Movement and Memory Data

Emma Richard [a*], Mary Kelly [b]
[a] Department of Cognitive Science, Carleton University
0000-0003-1164-6862
[b] Department of Cognitive Science, Carleton University
0000-0002-6541-2992

ABSTRACT: This study explored the potential of machine learning techniques, specifically decision trees and artificial neural networks (ANNs), to detect Parkinson's Disease (PD) using data collected from the mPower mobile iOS app. Released in 2016 by Sage Bionetworks, mPower enables individuals, both with and without PD, to assess their cognitive and physical abilities through various tasks related to memory, tapping, voice, and movement. The main focus of this study is on the walking task within the app's version 1.0 build 7. Participants were required to walk unassisted for approximately 20 seconds in a straight line, followed by a 30-second period of standing still, and then returning with 20 seconds. The smartphone's accelerometer and gyroscope captured three-dimensional (3D) rotation data (x, y, z) during these movements, with the device placed in the participant's pocket or bag. A convolutional neural network was applied to the movement dataset to assess confirmed PD cases, utilizing accelerometer and gyroscope readings during outward walking, return walking, and rest periods.

KEYWORDS: Parkinson's, artificial intelligence, movement, memory

https://doi.org/10.25071/2817-5344/55
* Corresponding Author - Email Address: emmarichard@cmail.carleton.ca
Received 10 June 2023; Received in revised form 03 Aug 2023; Accepted 12 Aug 2023
© 2023 The Author(s). This is an open access article under the CC BY-NC-SA 4.0 license

PD occurs spontaneously, as it affects the nervous system based on motor and non-motor symptoms.[1] The exact cause in most cases is not known or may not be attributed to a specific external factor, but it does not negate the presence of prodromal symptoms or other factors that can contribute to its development. It mainly affects older adults and is less common for people under 40 years old, and it is the second most common neurodegenerative disorder.[2] If you are diagnosed with PD under 40 years old it is considered early onset and it tends to be because of genetics.[3] The main risk factors related to PD are aging, genetics, and exposure to chemicals in the environment.[4] A known environmental risk factor is exposure to the pesticide paraquat, which is used to control weeds.[5] There have been animal models that show similar pathology to PD when they are exposed to paraquat. These models would show animals having an increase in anxiety one-week post-treatment and impaired motor skills 3 weeks post-treatment.[6] There are other neurodegenerative diseases that are similar to PD, such as Dementia with Lewy Bodies or Multiple System Atrophy, but what separates PD is the tremor at rest, rigidity, slowed movement, and poor posture.[7] There are about 60,000 new cases of PD diagnosed each year and 7-10 million people are affected worldwide.[8] When diagnosing PD, it is mainly based on the symptoms because there are no definitive tests to classify it.[9] It can be expensive to diagnose PD because doctors have to eliminate other neurological diseases to correctly diagnose the patient with PD, this can take hours to days.[10] PD typically starts with non-motor symptoms such as depression or anxiety and progresses into the patient having motor symptoms such as slowed movement or akinesia.[11] The later stages of PD can result in the patient having dementia due to the degradation of the midbrain.[12]

Released in 2016, Sage Bionetworks created a mobile iOS app called mPower to allow people with or without PD to measure their cognitive and physical capabilities. They created various tasks that people could complete on the app regarding memory, tapping, voice, and movement. All of these tasks are related to deficits in people with PD. This study will focus mainly on the walking task, where in version 1.0 build 7 the participant was required to walk unassisted for about 20 seconds in a straight line then stand still for

30 seconds then walk 20 seconds back. Any updates to the mPower app after version 1.0 build 7 did not include the participant required to walk 20 seconds back. The participant's phone was asked to be in their pocket or a bag. While using smartphones in pockets or bags may not achieve the same precision as attaching sensors to specific body segments, it remains a valuable and practical approach for movement data collection in certain research settings. The accelerometer and gyroscope data from the phone can capture important features related to walking, gait patterns, and tremors related to Parkinson's disease classification. The motion documentation from the device recorded the phone's three-dimensional (3D) rotation in an (x,y,z) plane based on the movement detected. There is also a demographics survey that asks for the participant's personal and medical information.

Related Work

Previous research has been conducted on the classification of PD using movement data, particularly through the utilization of mPower. In a study conducted by Zhang et al., it was discovered that individuals with PD exhibited shorter steps accompanied by longer strides, along with the presence of a resting tremor. Additionally, Zhang et al. identified certain correlations, such as education level, marital status, and retirement, which may serve as proxy measures for age. However, no correlation was found between the severity of symptoms and the duration of the disease since diagnosis. To achieve optimal model performance, it was recommended analyzing data from participants who engaged with the iOS app between 3 to 5 times.

In a similar vein, Pittman et al. conducted a study focusing on the classification of PD using movement data collected via mPower. To evaluate the data, they employed a 10-fold cross-validation technique. The results indicated that Pittman et al.'s logistic regression model achieved an accuracy of 79%, k-nearest neighbors exhibited 75% accuracy, grid search for a decision tree yielded 86% accuracy, and support vector classification demonstrated 82% accuracy. It is worth noting that a bias was identified in models

predicting a positive PD diagnosis, as there were no features employed to indicate a negative PD diagnosis.

As well, Mehrang et al. leveraged the mPower dataset to investigate the potential of machine learning in distinguishing participants with PD from those without PD based on movement data. Notably, when analyzing the data, they only considered a single measurement from participants with multiple entries. To mitigate confounding variables such as medication time and time of day, Mehrang et al. repeated a random selection of movement data 100 times. The study findings revealed that PD could be accurately recognized, and the utilization of smartphones for remote movement monitoring showcased promise in refining PD diagnosis.

In a similar investigation, Giuliano et al. focused on classifying PD based on voice recordings obtained from mPower. Participants recorded themselves saying "/a/," and their voice data was combined with demographic information, including age, sex, years since diagnosis, and medication start year. The study included 2,253 participants with unique health codes, all of whom were aged 35 or older. After removing inconsistent data based on the time of diagnosis and medication, the final dataset consisted of 933 participants with PD and 1,289 participants without PD.[13] The models were divided into 70% training data and 30% testing data. To simplify the analysis, the researchers reduced the number of vocal parameters from 62 to 5. Utilizing a two-layer multilayer perceptron network with the 5 vocal parameters, along with age and sex as additional attributes, Giuliano et al. achieved an accuracy of 76%. Furthermore, using the same attributes with a logistic regression model, they achieved an accuracy of 73%. These findings highlight the potential of using voice recordings and demographic data to classify PD accurately.

In the context of the current study, it is important to consider the significance and relevance of these related studies in shaping and advancing the research landscape in PD classification using movement data. By acknowledging the contributions of these previous works, the current study can build upon existing knowledge and contribute to the continued progress in this important area of research.

Methods

Movement Dataset

The mPower walking dataset from Sage Bionetworks consists of 3,101 unique participants: 658 participants have a confirmed diagnosis of PD, 2165 participants have not been diagnosed with PD, and 278 participants did not fill out if they have been diagnosed with PD. There are 35,410 unique records of the walking task, in 24,001 of these tasks the participant has a diagnosis of PD, which is 67.7% of unique tasks. We reduced the number of unique records to 21,046 rows to make sure we had a 50/50 split of participants with a PD diagnosis and participants without one. We removed every other datapoint that was recorded by the accelerometer and gyroscope in order to use our data with the convolutional neural network.

Participants who had a professional diagnosis were asked to complete the walking task three times a day, right before taking their medication, after taking their medication, and at another time of their choice. The participants who did not have a professional diagnosis of PD were asked to complete the walking task 3 times a day, but could do it at any time. The participants could also complete other tasks that involve memory, tapping, and vocals. Each participant was asked to fill out a demographics survey indicating their medical history, if they have been professionally diagnosed with PD, the year of diagnosis, and socio-demographic factors such as age, sex, employment, etc. There were 6,805 people who filled out this survey and each participant had a unique healthCode identifier that was matched to their results on the walking activity. The use of an iPhone during the walking task allowed there to be an accelerometer and gyroscope to detect the rotation of the device during movement. This can help recognize slowed movement, tremors, stiffness, and poor coordination. When trying to classify PD we used a 10-fold cross-validation and splitting the data uniquely by healthCode IDs into 70% training and 30% testing we analyzed the data using a convolutional neural network.

Convolutional Neural Network for Walking Dataset

To test if movement data is able to detect Parkinson's we used a pre-existing convolutional neural network from Brownlee. The creators used the convolutional neural network for human activity recognition based on if a person is walking, walking upstairs, walking downstairs, sitting, standing, or laying. Brownlee used the human activity recognition dataset from Davide Anguita, et at. at the University of Genova. From the mPower dataset we used the recordings of the accelerometer and gyroscope of outbound, rest, and return from the participants on the convolutional neural network. The mPower movement data was manipulated through a Python script that we created to split the data into fixed windows of 2.56 seconds, which is 128 data points, with a 50% overlap. We did this because this is how the convolutional neural network was programmed to receive data. When training and testing the convolutional neural network, we had to match when the timesteps were being recorded to when the timesteps were being recorded in the example dataset for the convolutional neural network. In the example dataset their timesteps were being recorded every 0.02 seconds and in the mPower movement dataset they were being recorded around every 0.04 seconds. The number of times that timesteps the were being recorded for the mPower movement dataset were around double the number of times they were being recorded in the example dataset used in the pre-existing code. We were able to use every other timestep in the mPower movement dataset to reduce the number of times the timesteps were being recorded to around 0.02 seconds, which is a close match to the example dataset. We split the data up into 70% training and 30% testing and so there is a 50/50 split between participants having a Parkinson's diagnosis and participants not having a Parkinson's diagnosis. In order to prevent bias each healthCode ID only appeared in either training or testing. There are a total of 9 features that we used in the model. First, is the total acceleration from the phone's accelerometer in gravity units for the x, y, and z coordinates. Next, we use the body acceleration from the phone's x, y, and z coordinates, we calculated this by subtracting the gravity from the total acceleration. Last, we use the angular velocity from the gyroscope for the x, y, z coordinates, this is considered the body gyroscope reading in radians per second.

Outcome Matrix

The outcome matrix used for each test using demographics and movement data was accuracy. Accuracy is the sum of correct predictions (true positives and true negatives) divided by the total number of predictions. A high accuracy indicates that the model is making accurate predictions, while a lower accuracy suggests that the model is not performing well on the given dataset, with 100 being the highest accuracy. Therefore, we had an accuracy for each individual test ran.

Results

Classification of Parkinson's Diagnosis using Movement Data

Classifying PD based on outbound movement data resulted in a 71% accuracy and 63% accuracy for when the participant was resting when using a convolutional neural network. Using the convolutional neural network to detect PD based on return movement data resulted in 68% accuracy based on 16,482 rows. The data manipulations for return were the same that was used for outbound and rest. For outbound, rest, and return the convolutional neural network ran 10 times and the final accuracy was the mean of the 10 trials (see figure 1). When combining the data for outbound and return it resulted in 71% accuracy based on 37,528 rows using the convolutional neural network.

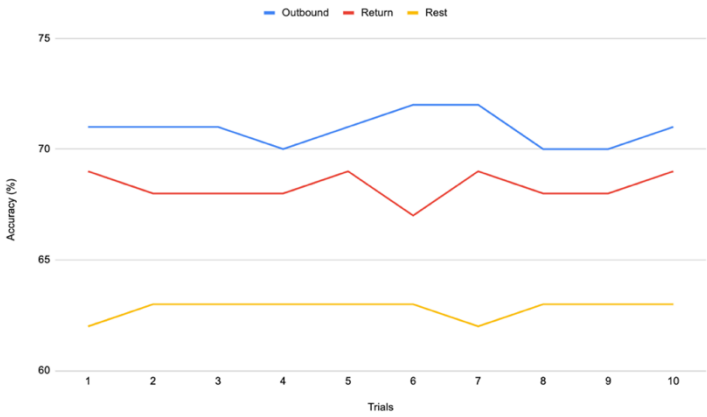

Figure 1. The comparison between the accuracy of each trial tested for outbound, rest, and return of the convolutional neural network.

Type of Model	Test Option	Attributes Used	Accuracy(%)
Decision Tree	10 Fold Cross-Validation	Demographics	97
Decision Tree	80/20 healthCode ID Split	Demographics	99
ANN	10 Fold Cross-Validation	Demographics	91
ANN	80/20 healthCode ID Split	Demographics	99

Table 1. Summary of the results when classifying PD with decision trees and ANNs using a 10-fold cross-validation test and splitting the data uniquely by

Classification of Parkinson's Diagnosis using Demographics

When using the 10-fold cross-validation for detecting a confirmed PD diagnosis (see Table 1) for the decision tree we got 97% accuracy when only demographic data was used to train the model. When using the 10-fold cross-validation for detecting a confirmed PD diagnosis for the ANN we got 91% accurate when only demographic data was used to train the model.

When testing the models with demographic data we used their age, education, employment, gender, marital status, and if they smoke. The results of splitting the dataset into 80% training and 20% testing uniquely by healthCode IDs for the decision tree were 99% accurate for the demographic data. Splitting the dataset into 80% training and 20% testing uniquely by healthCode IDs for the ANN were 99% accurate for the demographic data.

Discussion

For both the decision tree and ANN, the accuracies are highest (99%) when trained on demographics data when splitting the data uniquely by healthCode IDs. Since the results of splitting the data uniquely by healthCode IDs were just as high as the 10-fold cross-validation, this shows that the 10-fold cross-validation test was not just using overfitting to specific participants and using training data from a given participant to recognize the participant again at test.

Instead, the models are able to predict PD from participant demographics even given new participants not seen before. The attribute that is most informative when looking at detecting PD using demographics is age, as the decision tree splits the data by participants younger than or equal to 49 or if they are older than 49. Age is an important attribute when classifying PD because PD is much more common in older adults than in younger people. Another attribute that is important when classifying PD in decision trees is employment. I think employment is an important factor in classifying PD because a retired participant is more likely to be >60", resulting in higher chances of having PD. Additionally, education is another attribute that is important when classifying PD, this might be because people with higher education may have access to more resources to get diagnosed with PD than someone with little or no education. When looking at outbound, rest, and return for the movement data, the convolutional neural network was the most accurate at 71% for outbound movement. The amount of data used for outbound and rest was the same; however, the amount used for return was less since versions created after 1.0 build 7 of the mPower app did not include return. In the future, when trying to classify PD using movement data, I would recommend synthetically generating more movement data so the convolutional neural network would have more data to analyze, which could increase the accuracy.

Conclusion

Through leveraging a convolutional neural network, we surpassed chance-level classification by incorporating outbound, resting, and return movements. The ability to employ machine learning for PD classification holds promising implications, such as enhancing testing efficiency. Advancements in this area have the potential to help classify PD accurately, with discussions focusing on distinguishing early, mid, and advanced stages of the disease. Examining the robustness of the algorithm to classify the same person into the same category would be a valuable avenue to explore and discuss. The potential link between younger age and lower levels of education raises important considerations. Describing the population

under study, ensuring the balance of two groups (gender, age-matched), and verifying potential correlations between factors are essential aspects that require further exploration and clarification. Addressing these issues would enhance the reliability and applicability of the findings in real-world scenarios and contribute to the understanding of PD classification using movement data.

Notes

[1] Beitz, Janice M. "Parkinson's Disease a Review." *Frontiers in Bioscience* 6 (2014): 65–74. https://doi.org/10.2741/s415.
[2] Beitz, Parkinson's Disease Review, 65-74
[3] Klein, C., and A. Westenberger. "Genetics of Parkinson's Disease." *Cold Spring Harbor Perspectives in Medicine* 2 (2012). https://doi.org/10.1101/cshperspect.a008888.
[4] Beitz, Parkinson's Disease Review, 65-74
[5] Litteljohn, D., E. Mangano, N. Shukla, and S. Hayley. "Interferon-γ Deficiency Modifies the Motor and Co-Morbid Behavioral Pathology and Neurochemical Changes Provoked by the Pesticide Paraquat." *Neuroscience* 164, no. 4 (2009): 1894–1906. https://doi.org/10.1016/j.neuroscience.2009.09.025.
[6] Litteljohn, Interferon- γ and Parkinson's, 1894-1906.
[7] Beitz, Parkinson's Disease Review, 65-74
[8] Beitz, Parkinson's Disease Review, 65-74
[9] Beitz, Parkinson's Disease Review, 65-74
[10] Singh, Sanjana, and Wenyao Xu. "Robust Detection of Parkinson's Disease Using Harvested Smartphone Voice Data: A Telemedicine Approach." *Telemedicine and e-Health* 26, no. 3 (2020): 327–34. https://doi.org/10.1089/tmj.2018.0271.
[11] Schapira, Anthony H.V., K. Ray Chaudhuri, and Peter Jenner. "Non-Motor Features of Parkinson Disease." *Nature Reviews Neuroscience* 18, no. 7 (2017): 435–50. https://doi.org/10.1038/nrn.2017.62.
[12] Taguchi, Katsutoshi, Yoshihisa Watanabe, Atsushi Tsujimura, and Masaki Tanaka. "Expression of α-Synuclein Is Regulated in a Neuronal Cell Type-Dependent Manner." *Anatomical Science International* 94, no. 1 (2018): 11–22. https://doi.org/10.1007/s12565-018-0464-8.
[13] Giuliano, Monica, Alfonsa Garcia-Lopez, Silvia Perez, Fransisco Diaz Perez, Osvaldo Spositto, and Julio Bossero. "Selection of Voice Parameters for Parkinson´s Disease Prediction from Collected Mobile Data." *2019 XXII Symposium on Image, Signal Processing and Artificial Vision (STSIVA)*, 2019. https://doi.org/10.1109/stsiva.2019.8730219.

Bibliography

Beitz, Janice M. "Parkinson's Disease a Review." *Frontiers in Bioscience* 6 (2014): 65–74. https://doi.org/10.2741/s415.

Brownlee, Jason. "1D Convolutional Neural Network Models for Human Activity Recognition." MachineLearningMastery.com, August 27, 2020. https://machinelearningmastery.com/cnn-models-for-human-activity-recognition-time-series-classification/.

Giuliano, Monica, Alfonsa Garcia-Lopez, Silvia Perez, Fransisco Diaz Perez, Osvaldo Spositto, and Julio Bossero. "Selection of Voice Parameters for Parkinson´s Disease Prediction from Collected Mobile Data." *2019 XXII Symposium on Image, Signal Processing and Artificial Vision (STSIVA)*, 2019. https://doi.org/10.1109/stsiva.2019.8730219.

Klein, C., and A. Westenberger. "Genetics of Parkinson's Disease." *Cold Spring Harbor Perspectives in Medicine* 2 (2012). https://doi.org/10.1101/cshperspect.a008888.

Litteljohn, D., E. Mangano, N. Shukla, and S. Hayley. "Interferon-γ Deficiency Modifies the Motor and Co-Morbid Behavioral Pathology and Neurochemical Changes Provoked by the Pesticide Paraquat." *Neuroscience* 164, no. 4 (2009): 1894–1906. https://doi.org/10.1016/j.neuroscience.2009.09.025.

Mehrang, Saeed, Milla Jauhiainen, Julia Pietila, Juha Puustinen, Jari Ruokolainen, and Hannu Nieminen. "Identification of Parkinson's Disease Utilizing a Single Self-Recorded 20-Step Walking Test Acquired by Smartphone's Inertial Measurement Unit." *2018 40th Annual International Conference of the IEEE Engineering in Medicine and Biology Society (EMBC)*, 2018. https://doi.org/10.1109/embc.2018.8512921.

Pittman, Benjamin, Reza Hosseini Ghomi, and Dong Si. "Parkinson's Disease Classification of Mpower Walking Activity Participants." *2018 40th Annual International*

Conference of the IEEE Engineering in Medicine and Biology Society (EMBC), 2018. https://doi.org/10.1109/embc.2018.8513409.

Schapira, Anthony H.V., K. Ray Chaudhuri, and Peter Jenner. "Non-Motor Features of Parkinson Disease." *Nature Reviews Neuroscience* 18, no. 7 (2017): 435–50. https://doi.org/10.1038/nrn.2017.62.

Singh, Sanjana, and Wenyao Xu. "Robust Detection of Parkinson's Disease Using Harvested Smartphone Voice Data: A Telemedicine Approach." *Telemedicine and e-Health* 26, no. 3 (2020): 327–34. https://doi.org/10.1089/tmj.2018.0271.

Taguchi, Katsutoshi, Yoshihisa Watanabe, Atsushi Tsujimura, and Masaki Tanaka. "Expression of α-Synuclein Is Regulated in a Neuronal Cell Type-Dependent Manner." *Anatomical Science International* 94, no. 1 (2018): 11–22. https://doi.org/10.1007/s12565-018-0464-8.

Zhang, Hanrui, Kaiwen Deng, Hongyang Li, Roger L. Albin, and Yuanfang Guan. "Deep Learning Identifies Digital Biomarkers for Self-Reported Parkinson's Disease." *Patterns* 1, no. 3 (2020): 100042. https://doi.org/10.1016/j.patter.2020.100042.