

DiPark: A Novel Approach to Contactless & Multimodal Parkinson's Disease Diagnosis and Severity Analysis via Computer Vision and Ensemble Learning

Scienteer Project #175865

Introduction

Parkinson's Disease (PD) is a neurodegenerative disease with no cure. Additionally, existing methods of diagnosis and treatment are not the most accurate or cost-effective. The symptoms that entail this disease affect the movement of the patient's eyes, speech, and other visual attributes. Currently, there are no definitive ways to diagnose PD, as existing methods are often inaccurate, inaccessible, and expensive. Because over 10 million people worldwide have PD, it is imperative that an effective diagnosis technology be made available.

Background

- Parkinson's Disease has a high misdiagnosis rate of over 25%, meaning 1 in 4 patients are misdiagnosed.
- Existing technologies are not only costly, but are invasive and often inaccurate
- Existing machine learning models are unimodal and look for only one symptom, which may not be present in the patient.

Engineering Goal

- To provide a cost-effective and accurate diagnosis of Parkinson's Disease
- To help reduce the misdiagnosis rate of Parkinson's Disease and save lives
- To aid in the shift to contactless and virtual diagnosis technology

Materials & Methodology

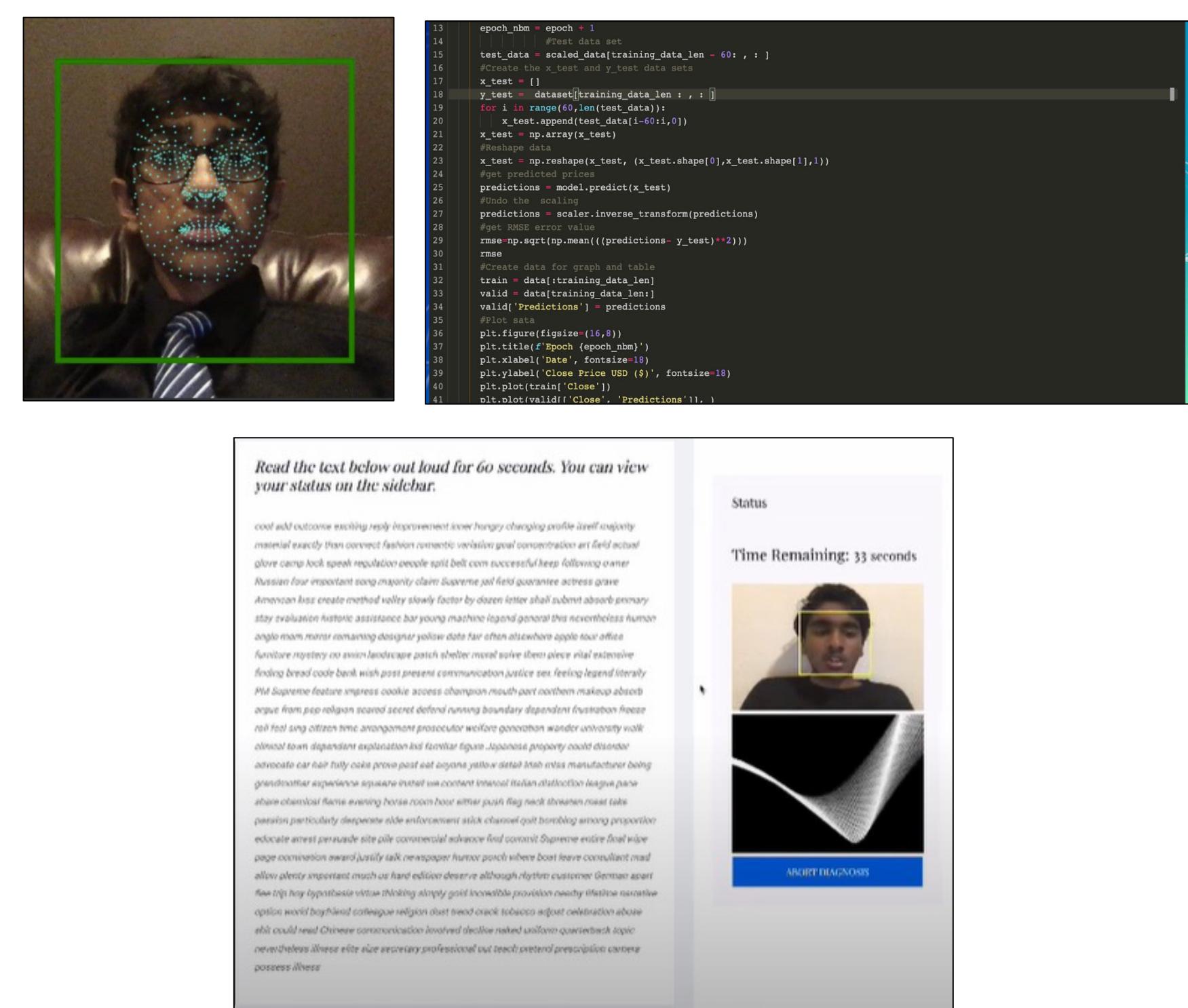
- Design the model and application using Canva.
- Create a Google Colab notebook to begin creating the model.
- Download University of California, Irvine Data Repository's Parkinson's Disease (PD) Data for audio.
- Download Data from the Parkinson's Progression Markers Initiative on eye movement, facial cues, and speech.
- Pair similar subjects and pad incomplete modalities with 0s and place data in a CSV file.
- Sort the data in combination groups, with combinations representing the state of missing modalities.
- Gather the feature sets via the Audio-Visual Emotion Challenge, and reduce over 1200 features using the Minimum Redundancy Maximum Relevance algorithm.
- Using Stratified K-Fold Cross Validation, split each combination dataset into training and testing data.
- Using Keras, develop a Convolutional Neural Network, with dropout layers set at 15%, 4 hidden layers with 128, 64, 32, and 16 neurons respectively, and activation function ReLU for each individual modality.
- Using Multi-Task Deep Learning, combine the results from each neural network to produce a single classification result.
- Using a Gradient Boosted Classifier for the audio data to produce a single modality classification result.
- Utilize Support Vector Machine-Recursive Feature Elimination for Eye Data, extracted via Google's MobileNetV2.
- Utilize a General Support Vector Machine for Visual Data, extracted via MIT's WebGazer.
- Utilize Ensemble Learning to use various classifiers for each modality.
- Fuse modality prediction results via Late Fusion.
- Evaluate the model using the testing data and calculate accuracy, precision, recall, and f-1 score metrics.
- Deploy the model in an S3 bucket connected to the SageMaker Endpoint and E2 Instance.
- Develop a landing page for the website.
- Develop a login page for the website, using cPanel's MySQL database.
- Connect the endpoint to the website using the lambda script and API.

Data

- 106 Correctly Diagnosed Patients
 - All Positive patients correctly diagnosed
 - All but one Negative patient correctly diagnosed
 - 0 False Negatives
 - 1 False Positive
 - 1 Incorrectly Diagnosed Patients
- Accuracy of .9905
 - Precision of .9905
 - Recall of 1
 - F-1 Score of 0.9952
 - Average Confidence Score of 0.9681

Performance measures	Naïve Bayes	Logistic Regression	Boosted Trees	Random Forests	SVM	
Training	Testing	Training	Testing	Training	Testing	
Accuracy (%)	94.67 ± 0.59	93.12 ± 1.49	96.50 ± 0.60	95.63 ± 1.21	100 ± 0.02	99.99 ± 1.26
Sensitivity (%)	94.50 ± 0.68	92.67 ± 2.19	97.38 ± 0.54	96.78 ± 1.54	100 ± 0.04	96.18 ± 1.29
Specificity (%)	95.07 ± 0.97	93.52 ± 3.17	94.56 ± 1.00	93.26 ± 2.82	100 ± 0	93.15 ± 3.74
AUC (%)	98.66 ± 0.29	96.77 ± 0.21	99.20 ± 0.21	98.66 ± 0.77	100 ± 0.82	98.40 ± 0.91
	± 0.29 ± 1.33	± 0.21 ± 0.77	± 0.21 ± 0.77	± 0.21 ± 0.77	± 0.09 ± 0.91	± 0.16 ± 0.62

Photos



Data Analysis/Statistics

- Mean accuracy across individual classifiers was 94.87%
- Mean confidence across all test prediction was 0.9681
- Median Confidence across all test prediction was .9703
- Lower Confidence Bound on test prediction was .8005
- Highest Confidence Bound on test prediction was 0.9991
- Distribution of both accuracy & confidence was skewed left

Error Analysis

- Due to the scarcity of data, the datasets used were non-parallel, which could introduce confounding factors that are difficult to identify.
- The lack of true clinical studies prevents a robust method of verifying the accuracy of this model.

Results

- The three symptoms that DiPark analyzes are rarely all absent in a patient, thus allowing us to account for symptoms not present in a patient.
- The accuracy of DiPark is around the metric of state-of-the-art models for the diagnosis of other diseases, including neurodegenerative diseases, such as Alzheimer's. However, this is the first attempt at Parkinson's Disease (PD) diagnosis using multimodal learning that utilizes both speech and video data. The accuracy of DiPark is much greater than that of unimodal models that diagnose PD, which typically have accuracy rates that range from 60%-84%. Additionally, the confidence metrics tended to have a strong positive correlation to the number of years a patient has had PD

Conclusion

- DiPark proves that Multimodal diagnosis of Parkinson's Disease is more effective than conventional methods of diagnosis, and unimodal designs
- Contactless forms of Diagnosis are highly accurate, even more so than physical forms of diagnosis
- Analyzing multiple modalities is vital to ensuring an accurate Diagnosis of Parkinson's Disease
- Ensemble learning proves to be a capable method of increasing accuracy, decreasing variance, and decreasing bias
- Non-Parallel data, while not as potent as Parallel Data, can still be used to create a highly accurate model, especially when multiple modalities are vital
- Just a few dozen features are enough to accurately diagnose and analyze the severity of Parkinson's Disease

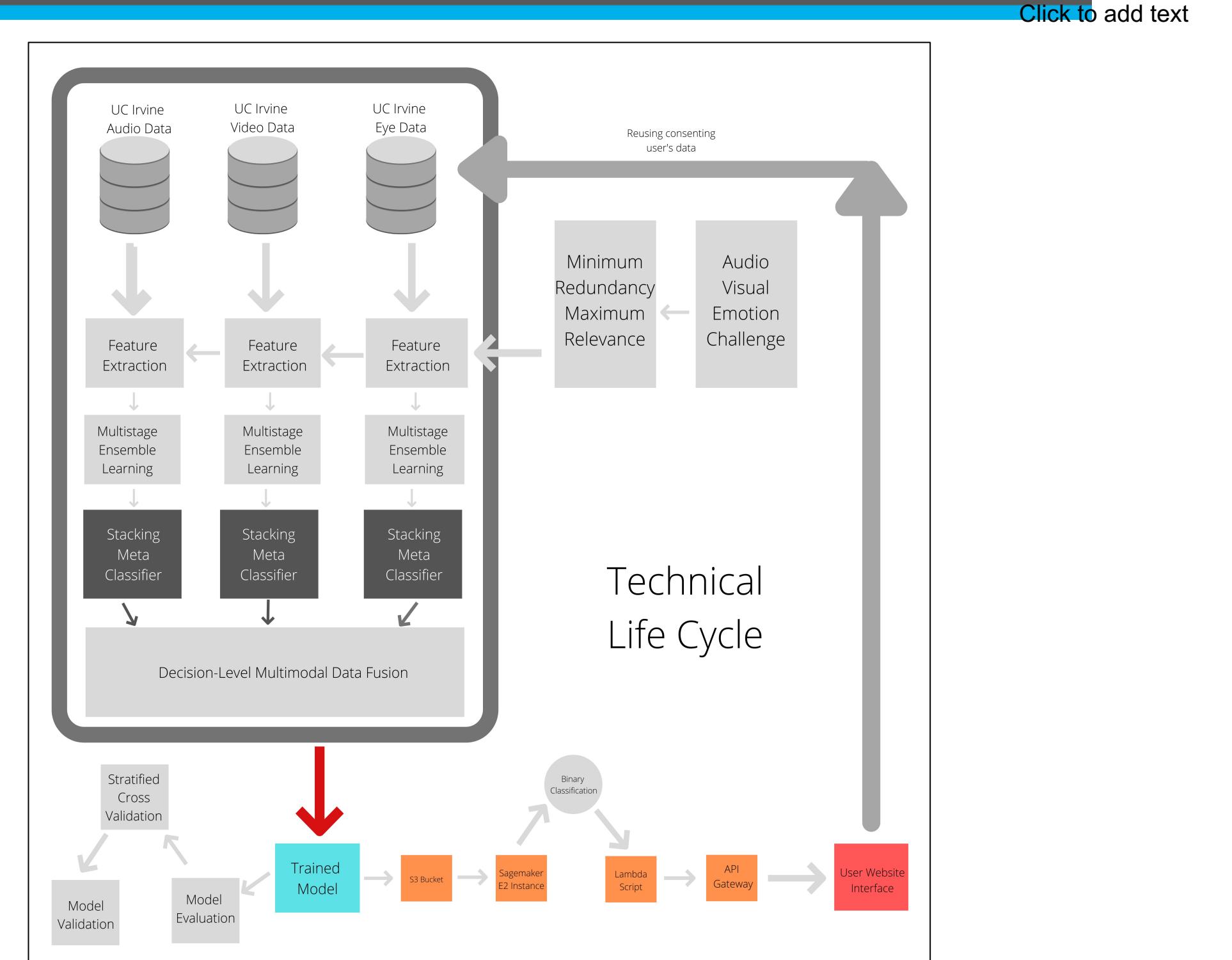
Applications

- Online diagnosis technology
- Replacement to questionaries
- Retirement Homes
- Licensed Government Research Technology
- Hospitals (as opposed to DaTScans)

Photo Credit

All photos, images, and graphics were done by the researcher or parent

Designs



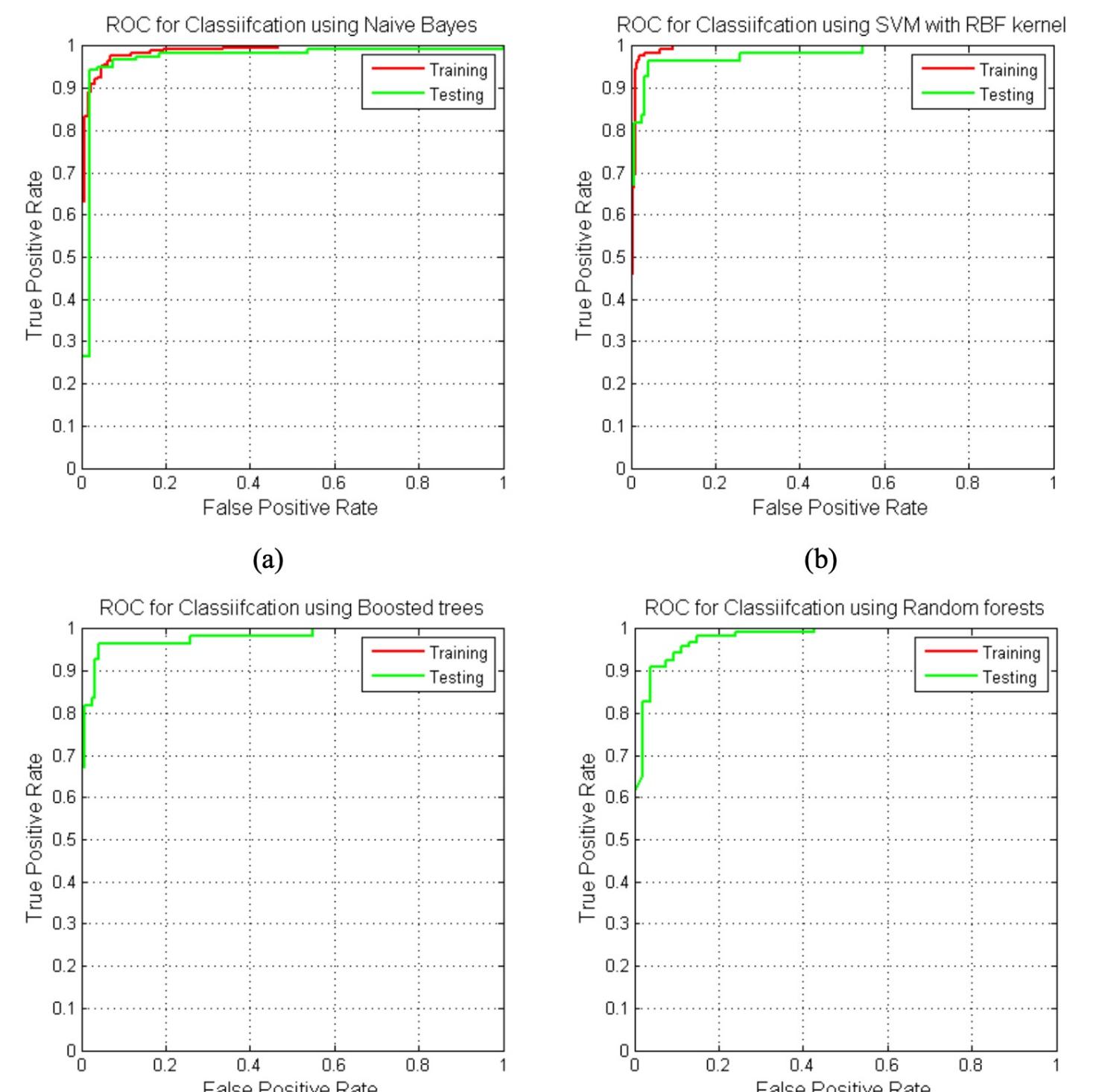
Observations

- Greater movement-based measures are often correlated with higher severity
- Greater Confidence Scores typically correlated to greater severity in Parkinson's-Positive subjects
- Accuracy amortized after roughly 26 epochs, despite running nearly double that
- Vocal Analysis was the most consistent with the final result after model fusion
- Facial analysis was the least consistent with the final result after model fusion
- Ensemble learning increased the accuracy of the model by nearly 8.3% after added to all modalities
- Despite over 1200 features, only 62 were deemed relevant through MRMV
- The Fusion prediction was more accurate than each individual prediction

Charts



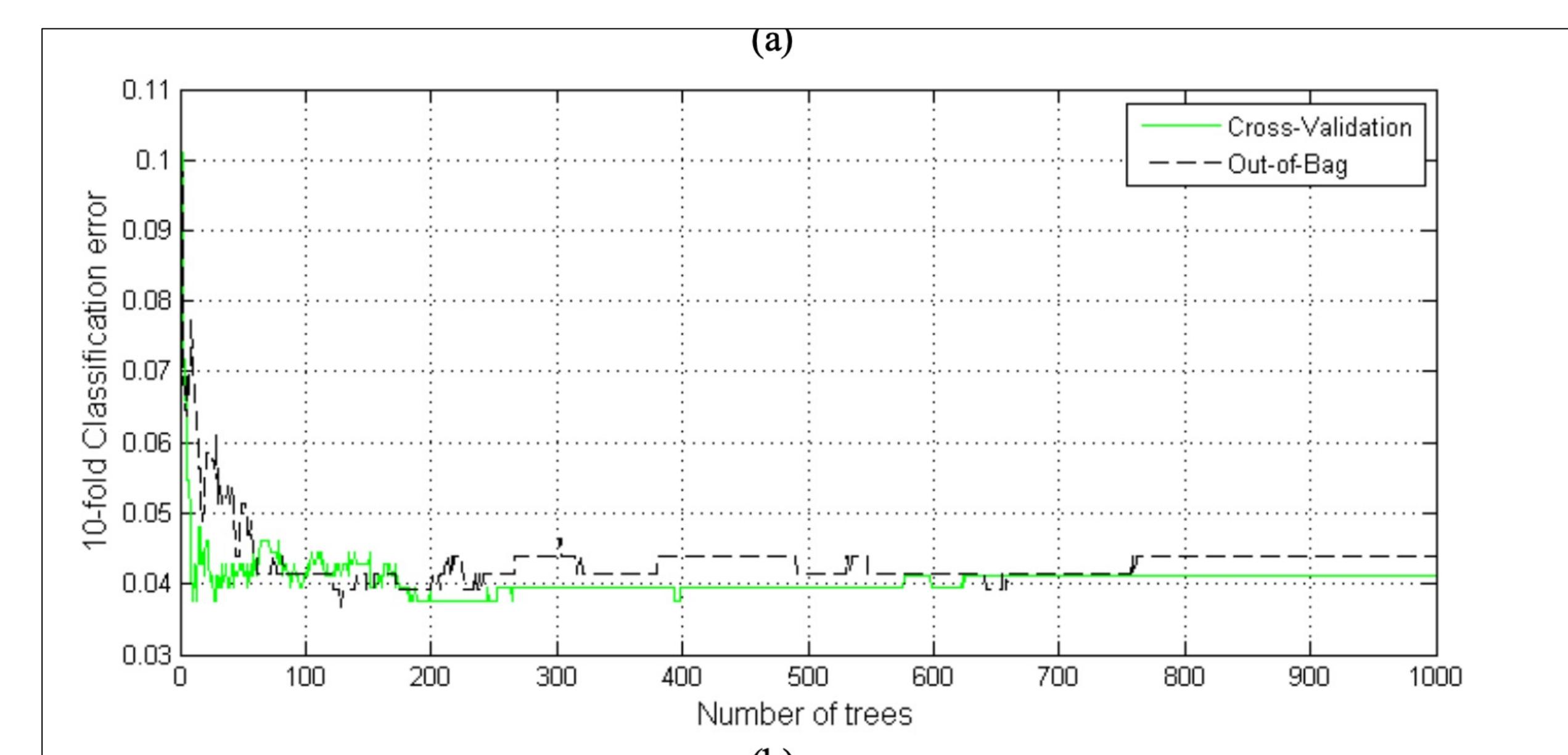
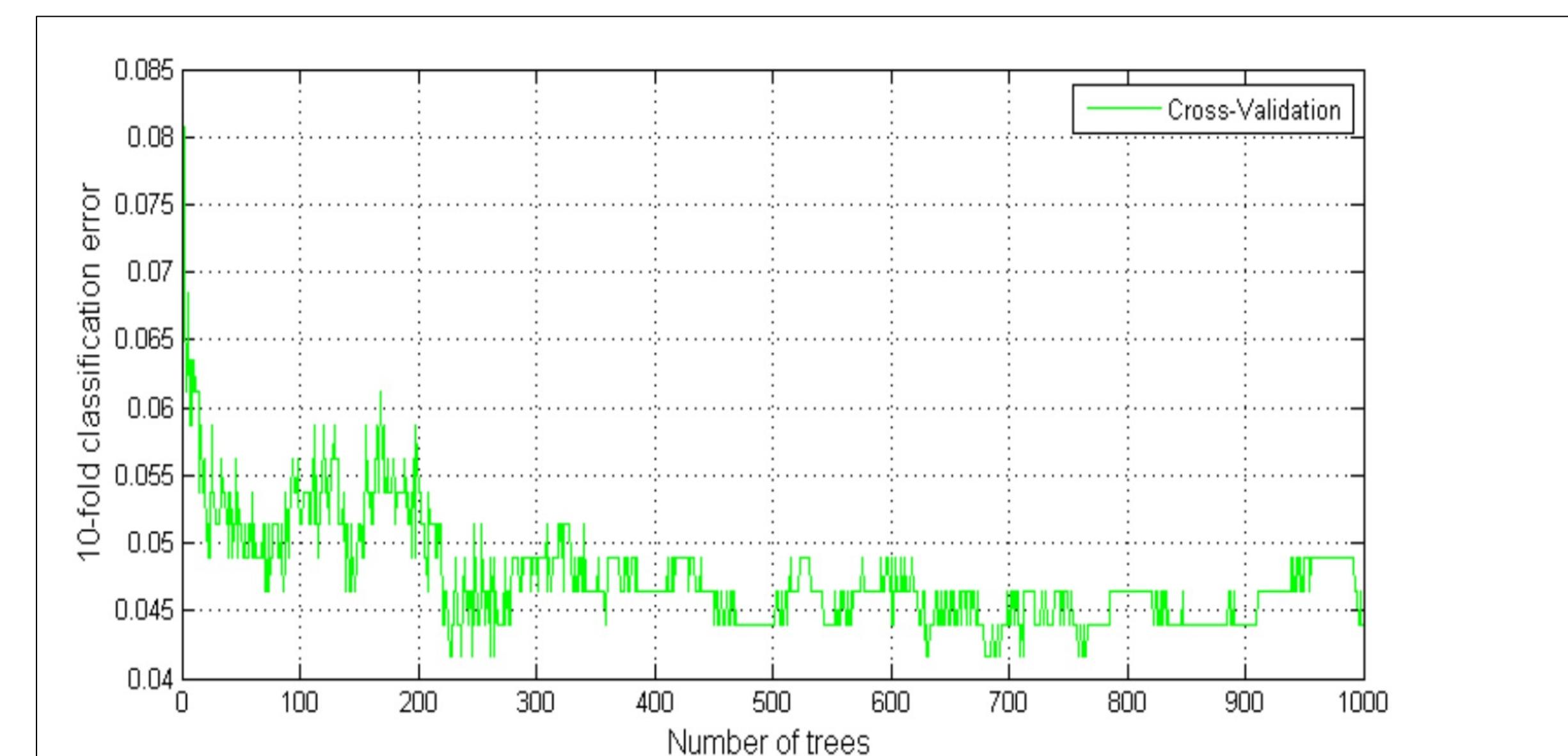
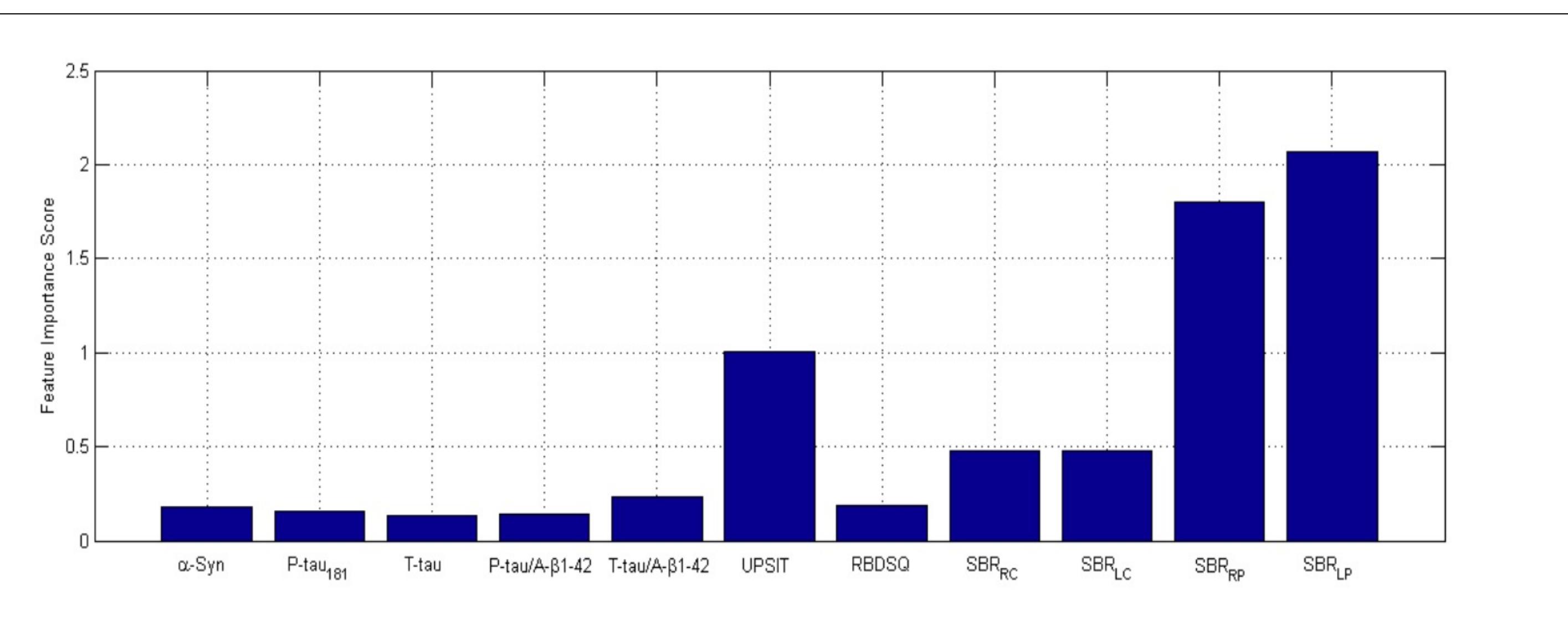
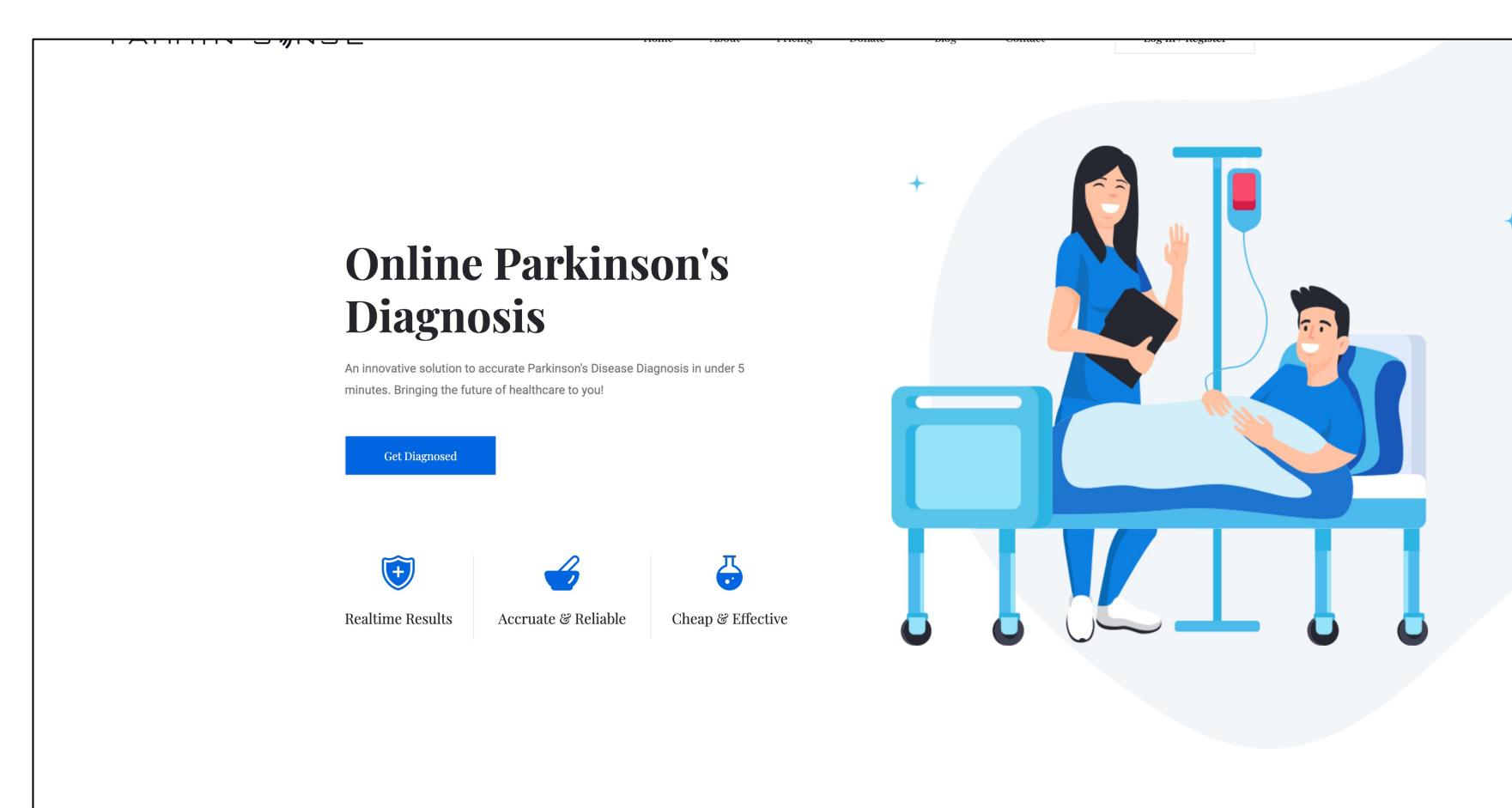
Graphs



Additional Information and Citations

Further Data Analysis

We split our initial data using stratified K-fold cross-validation to obtain both training datasets and test datasets. The training dataset will be used to train the model, while the test dataset will be used to validate the model. Once the testing dataset is evaluated, our model will return four different metrics: accuracy, precision, recall, and f-1 score. Accuracy is a measure of the number of correct diagnoses over the total number of diagnoses. Precision is calculated as the number of true positives divided by the sum of true positives, and false positives. The recall is calculated by dividing the number of true positives by the sum of the number of true positives and false negatives. Finally, the F-1 scores two times the product of precision and recall divided by the sum of precision and recall. All of these will be values between 0 and 1, with values closer to one implying a more accurate model. These four metrics were evaluated and analyzed. Each metric will be analyzed against relevant literature that also attempts to address the issue of Parkinson's Diagnosis using machine learning. This will provide us with the conclusion to the question at hand of the effectiveness of DiPark at Parkinson's Disease diagnosis.



Features	Normal (n ₁ = 183)	Early PD (n ₂ = 401)	z-statistic	p-value*
RBDSQ score	2.60 ± 2.09	3.26 ± 2.67	-2.42	0.016
UPSiT score	34.08 ± 4.74	22.22 ± 8.23	15.21	≈ 0
A β 1-42	376.17 ± 109.65	372.73 ± 99.35	0.66	0.501
α-Syn	2197.9 ± 1085.1	1856.6 ± 795.1	3.45	≈ 0
P-tau ₁₈₁	17.98 ± 11.23	15.82 ± 10.20	3.04	0.002
T-tau	52.47 ± 27.10	45.08 ± 18.53	3.05	0.002
T-tau / A β 1-42	0.16 ± 0.19	0.13 ± 0.07	2.28	0.023
P-tau ₁₈₁ / A β 1-42	0.05 ± 0.06	0.04 ± 0.03	2.33	0.020
P-tau ₁₈₁ / T-tau	0.37 ± 0.19	0.37 ± 0.22	0.34	0.737
SBR _{RC}	2.95 ± 0.62	1.99 ± 0.60	14.52	≈ 0
SBR _{LC}	3.00 ± 0.65	2.00 ± 0.60	14.65	≈ 0
SBR _{RP}	2.15 ± 0.57	0.84 ± 0.36	18.31	≈ 0
SBR _{LP}	2.13 ± 0.57	0.81 ± 0.36	18.45	≈ 0

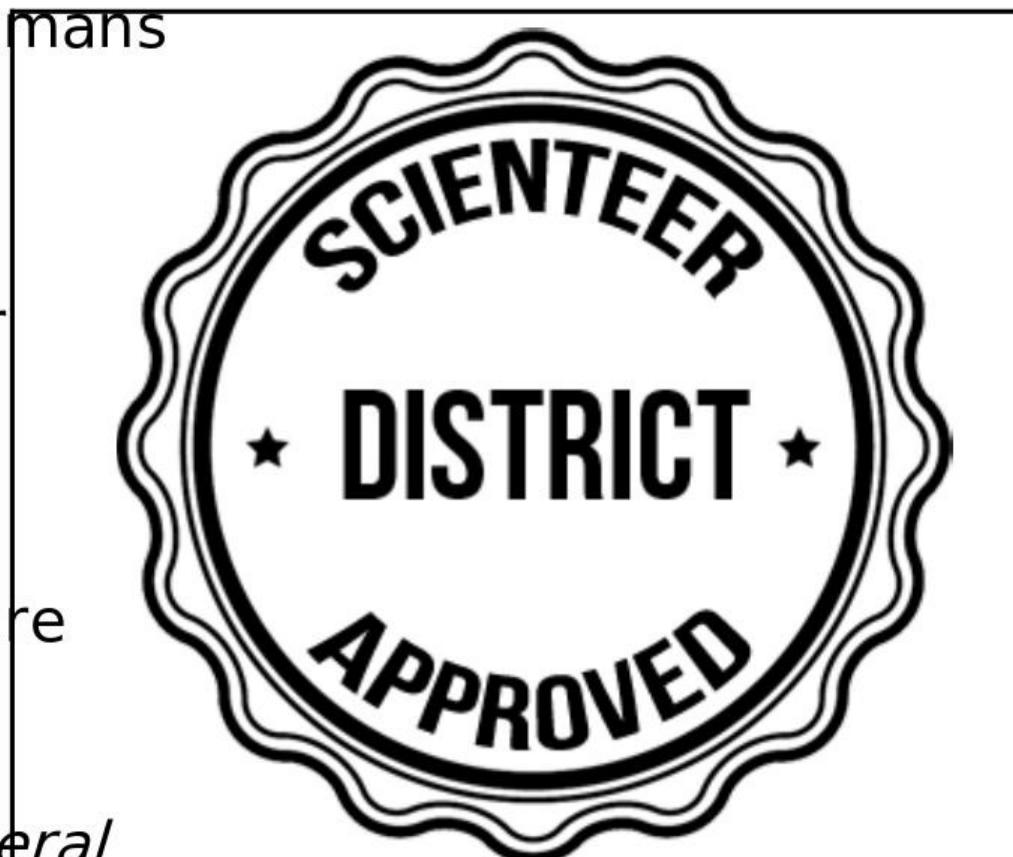
Citations

- Baltrušaitis, Tadas, Chaitanya Ahuja, and Louis-Philippe Morency. n.d. "Multimodal Machine Learning: A Survey and Taxonomy." <http://arxiv.org/abs/1705.09406v2>.
- 'Exploiting Nonlinear Recurrence and Fractal Scaling Properties for Voice Disorder Detection', Little MA, McSharry PE, Roberts SJ, Costello DAE, Moroz IM. BioMedical Engineering OnLine 2007, 6:23 (26 June 2007)
- Huang, Mei-Ling et al. "SVM-RFE based feature selection and Taguchi parameters optimization for multiclass SVM classifier." TheScientificWorldJournal vol. 2014 (2014): 795624. doi:10.1155/2014/795624
- Korosec, Marko et al. "Eyelid movements during blinking in patients with Parkinson's disease." Movement disorders: official journal of the Movement Disorder Society vol. 21,8 (2006): 1248-51. doi:10.1002/mds.20930
- Liu, Siqi et al. "Multimodal neuroimaging feature learning for multiclass diagnosis of Alzheimer's disease." IEEE transactions on bio-medical engineering vol. 62,4 (2015): 1132-40. doi:10.1109/TBME.2014.2372011
- Ma, Mengmeng, et al. "Smil: Multimodal learning with severely missing modality." Proceedings of the AAAI Conference on Artificial Intelligence. Vol. 35. No. 3. 2021.
- "Parkinson's Disease." Mayo Clinic, Mayo Foundation for Medical Education and Research, 14 Jan. 2022, <https://www.mayoclinic.org/diseases-conditions/parkinsons-disease/diagnosis-treatment/drc-20376062#:~:text=No%20specific%20test%20exists%20to,a%20neurological%20and%20physical%20examination>.
- Qi Wang, Liang Zhan, Paul Thompson, and Jiayu Zhou. 2020. Multimodal Learning with Incomplete Modalities by Knowledge Distillation. Proceedings of the 26th ACM SIGKDD International Conference on Knowledge Discovery & Data Mining. Association for Computing Machinery, New York, NY, USA, 1828–1838. DOI:<https://doi.org/10.1145/3394486.3403234>
- Rahate, Anil, et al. "Multimodal Co-learning: Challenges, applications with datasets, recent advances and future directions." Information Fusion 81 (2022): 203-239.
- Shahbakh, M. , Far, D. and Tahami, E. (2014) Speech Analysis for Diagnosis of Parkinson's Disease Using Genetic Algorithm and Support Vector Machine. Journal of Biomedical Science and Engineering, 7, 147-156. doi: 10.4236/jbise.2014.74019.
- Sibley, Krista G. et al. 'Video-Based Analyses of Parkinson's Disease Severity: A Brief Review'. 1 Jan. 2021 : S83 – S93.
- Thung, Kim-Han et al. "Multi-stage Diagnosis of Alzheimer's Disease with Incomplete Multimodal Data via Multi-task Deep Learning." Deep learning in medical image analysis and multimodal learning for clinical decision support: Third International Workshop, DLMIA 2017, and 7th International Workshop, ML-CDS 2017, held in conjunction with MICCAI 2017 Quebec City, QC,... vol. 10553 (2017): 160-168. doi:10.1007/978-3-319-67558-9_19
- Vasquez-Correa, Juan Camilo et al. "Multimodal Assessment of Parkinson's Disease: A Deep Learning Approach." IEEE journal of biomedical and health informatics vol. 23,4 (2019): 1618-1630. doi:10.1109/JBHI.2018.2866873
- Xie, X. Sampling Active Learning Based on Non-parallel Support Vector Machines. Neural Process Lett 53, 2081–2094 (2021). <https://doi.org/10.1007/s11063-021-10494-x>

Project Forms

Abstract

OFFICIAL ABSTRACT and CERTIFICATION		<u>Category</u>
<p>DiPark: A Novel Approach to Contactless Early Parkinson's Disease Diagnosis and Severity Analysis via Multimodal Learning and Multi-Task Deep Learning</p> <p>Rithik Duvva, Sharv Save, and Pulkith Paruchuri Heritage High School, Frisco, Texas, US</p> <p>Parkinson's Disease (PD) is one of the most prevalent neurodegenerative diseases, affecting over ten million people worldwide. Since certain medications for PD are only effective in the preliminary stages, it is imperative that patients are diagnosed early on. However, there is no definitive way to diagnose PD, and current methods of diagnosis are highly inaccurate, inaccessible, and expensive. Furthermore, nearly 1 in 3 patients are misdiagnosed for PD. Therefore, it is paramount a contactless, cost-effective, and accurate method of diagnosis is developed. To combat these issues, DiPark, an online method of PD diagnosis that uses a webcam and a microphone was created. Using a multimodal approach, DiPark analyzes speech irregularities, facial cues, and eye movement patterns to diagnose a user for PD, while also providing confidence scores for each modality and the overall classification that acts as PD severity metrics. Data collected from the UC Irvine Data Repository (64 Healthy Controls, 124 PD), and the Parkinson's Progression Markers Initiative (106 Health Controls, 51 PD), were used by Convolutional Neural Networks and a few classifiers including Support Vector Machines and Gradient Boosted Classifiers. The results from these classifiers and deep neural networks were fused via a Multi-task Learning algorithm on incomplete multimodal data, that returned a PD diagnosis binary classification and a confidence metric to the user. All neural networks and classifiers were evaluated on their accuracy, precision, recall, and F-1 score. Through this, DiPark was able to achieve over 97% accuracy on all four metrics for each modality network.</p>		Translation Medical Science
<p>1. As a part of this research project, the student directly handled, manipulated, or interacted with (check ALL that apply):</p> <p><input type="checkbox"/> human Participants <input type="checkbox"/> potentially hazardous biological agents <input type="checkbox"/> vertebrate animals <input type="checkbox"/> microorganisms <input type="checkbox"/> rDNA <input type="checkbox"/> tissue</p> <p>2. I/we worked or used equipment in a regulated research institution or industrial setting:</p> <p><input type="checkbox"/> Yes <input checked="" type="checkbox"/> No</p> <p>3. This project is a continuation of previous research:</p> <p><input type="checkbox"/> Yes <input checked="" type="checkbox"/> No</p> <p>4. My display board includes non-published photographs/visual depictions of humans (other than myself):</p> <p><input type="checkbox"/> Yes <input checked="" type="checkbox"/> No</p> <p>5. This abstract describes only procedures performed by me/us, reflects my/our own independent research, and represents one year's work only:</p> <p><input checked="" type="checkbox"/> Yes <input type="checkbox"/> No</p> <p>6. I/we hereby certify that the abstract and responses to the above statements are correct and properly reflect my/our own work:</p> <p><i>This stamp or embossed seal attests that his project is in compliance with all federal and state laws and regulations and that all appropriate reviews and approvals have been obtained including the final clearance by the Scientific Review Committee.</i></p>		



Project Board Rules

- A virtual project board must be uploaded prior to the deadline set for the fair.
- All [ISEF Display and Safety Rules](#) must be followed as well as any additional rules below.
- Students are allowed **three** 48" x 36" slides.
- Font on the boards must be 16pt or larger.
- If slides contain photographs with people other than the researcher, the researcher must have photo releases available upon request.
- The first two slides may NOT include an abstract. Only your official abstract from Scienteer should be included on slide 3.
- The slides may NOT include any animations, videos, links, or demonstrations.
- The slides may NOT include any offensive or inappropriate images or photographs.
- The slides may NOT include any reference to an institution, mentor that supported the work, or any patents pending other than what is displayed on Form 1C on slide 3.
- The slides may NOT include any personal information such as email address, home address, phone number, or social media contacts.

Slide 1

- The first slide should be modeled after a engineering fair board (see slide 1 example) and include all key information such as Problem/Question, Hypothesis or Engineering Goal, Background Research, Materials, Procedure, Data, Results, and Conclusion.
- The Scienteer Project # must be displayed under the Project Title on slide 1.
- **The example project board on slide 1 may be used as a guide by students. Formatting to change colors, text box sizes, move panels, etc. is acceptable.** We want this to be your board. Just make sure to follow the rules as to what needs to be included.
- The first slide must include a Photo Credit that gives credit for all photos, images, and graphics on the slides.

Slide 2

- The second slide may include any supporting information for the project such as additional photos, charts, and graphs. Make sure to use image citations as needed.

Slide 3

- The third slide must include the Official SRC Approved Abstract from Scienteer as well as Form 1C (Regulated Research Institutional/ Industrial Setting Form) and Form 7 (Continuation/ Research Progression Projects Form) if required for your project.