HS-MTCS-243-T2

Using AI to diagnose Parkinson's Disease from Voice Recordings

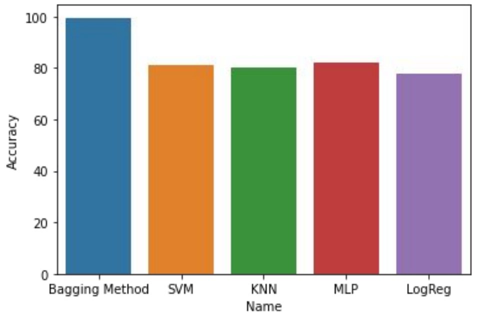
Maya Poghosyan & Diya Vij -- Dublin High School, Dublin, California, USA

Q1: Problem or Question

- Engineering Goal: A complex machine learning algorithm that can predict a Parkinson's Disease (PD) diagnosis based on recorded speech features to enable remote diagnosis that complements new telehealth systems

Q3: Findings

- Reached **perfect accuracy** ⅔ trials
- Outperformed any other method



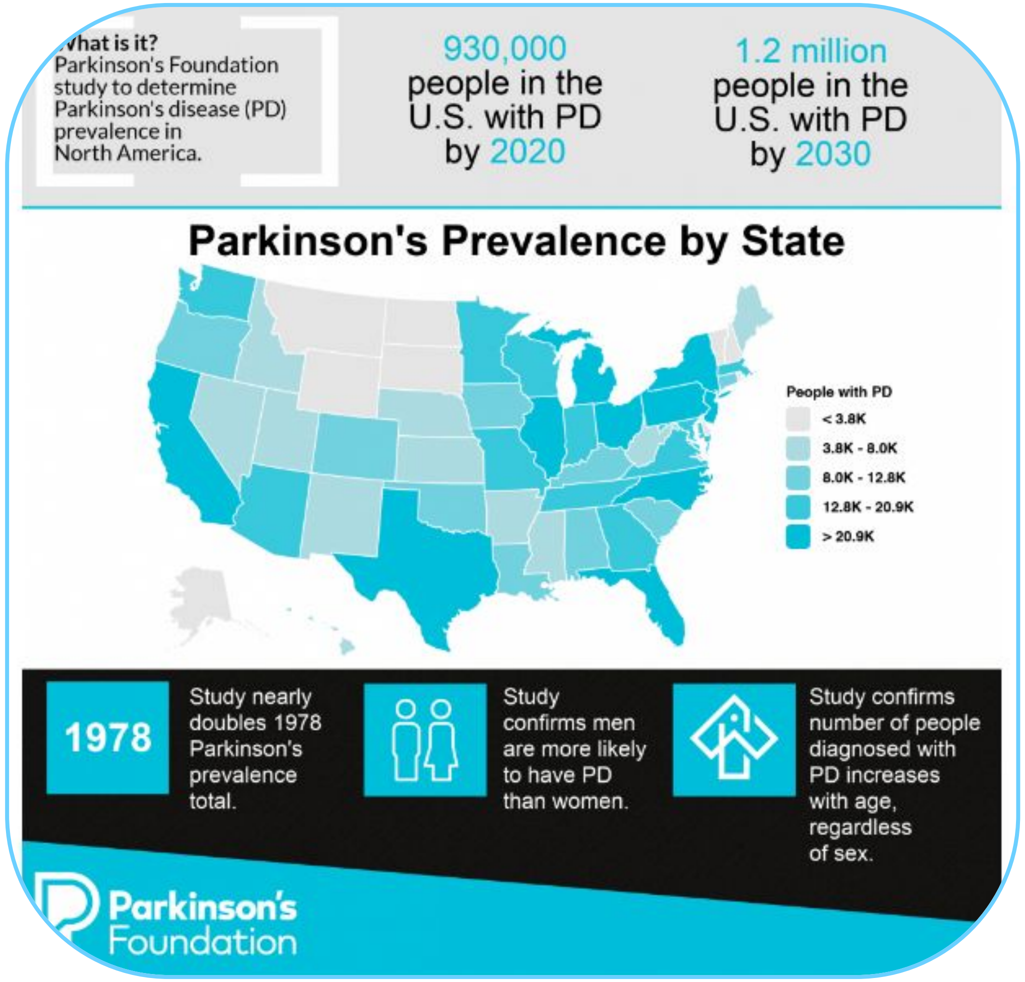
Q2: Framework

- Expe
- Principal component analysis to reduce dimensionality
- Bootstrap aggregating algorithm that classifies patients as sick or healthy
- KFolding to ensure result validity
-

Q4: Conclusions

- Great potential in telehealth -- can be used to aid in remote diagnosis in any kind of scenario
- Improves PD testing accessibility for seniors of all kinds and is a financially sound alternative to current procedure

Introduction



PD is currently the NUMBER TWO neurological disease that affects the most amount of people in the United States.

📈 Prevalence rate is on the rise

A progressive disease - detection is critical

Long-distance testing is a great alternative in the current system, COVID-19 and other infectious diseases disproportionately impact seniors who want to minimize travel and contact. Our project tested whether or not long-distance testing could be accurate, and therefore a viable alternative to current methods.

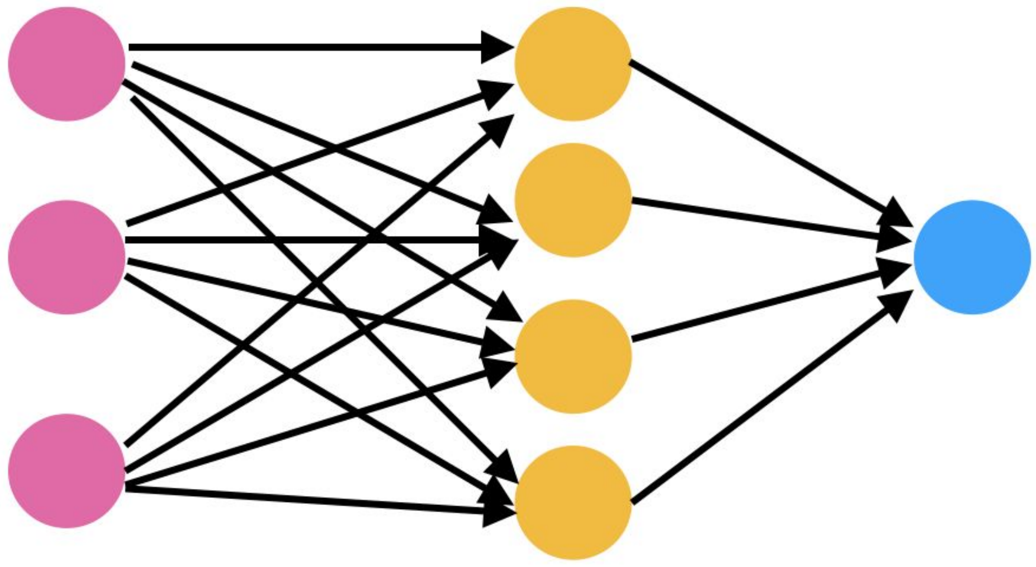
Algorithm goal → AID healthcare systems in setting up early action diagnosis and treatment plans

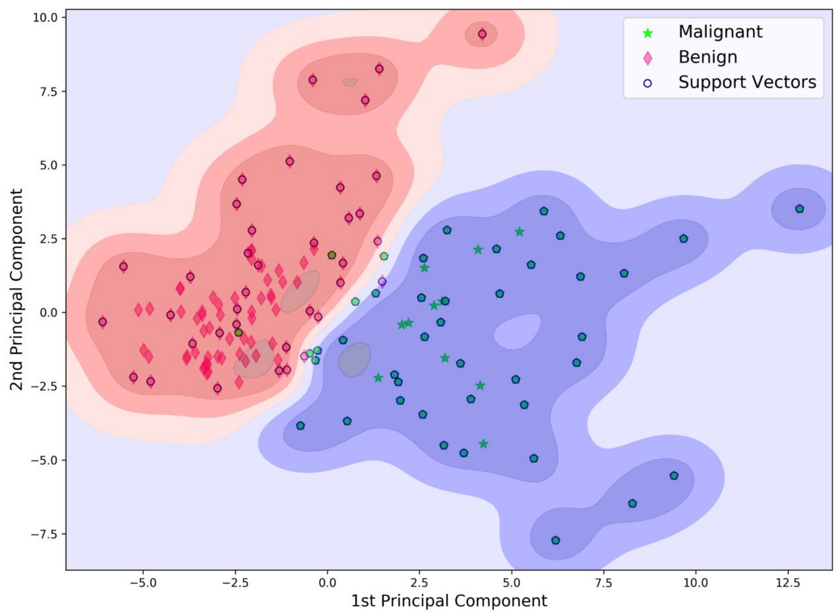
Framework (I) - Exploration

Attempted to replicate previous results @ University of Istanbul study

We used simpler models including:

- Logistic Regression - predicting a 0 or a 1
- Support Vector Machine (SVM)
- K - Nearest Neighbors (KNN) - a grouping classifier used in the past study
- Multilayer Perceptron Network (MLP) - a pretrained neural net architecture





Framework (II) - Our Solution

Principal Component Analysis - Reduced dimensionality, or number of features, of our data --“boiled down” to 4 principal components
→ allows for quicker & more robust data processing

	principal component 1	principal component 2	principal component 3	principal component 4	label
0	-1.345768	-0.561464	-1.012609	-0.137229	1
1	-1.457105	-0.497098	-1.117513	0.188325	1
2	-1.929426	-0.180870	-0.951920	0.269965	1
3	-0.386014	-1.766024	1.813981	-0.398953	1
4	-0.161909	-1.189702	1.912959	-0.745900	1

Training - used a bootstrap aggregating or “bagging” model that is an ensemble of decision tree models.

- Model makes a “decision” based on decision rules it has learned w/ training data

Cross-Validation - used KFold to shuffle + split our data
→ test validity of model on novel data

Findings

<u>Basic Models</u> - Accuracy

- Logistic Regression - 78%
- Support Vector Machine (SVM) - 81%
- K - Nearest Neighbors (KNN) - 79.6%
- Multilayer Perceptron Network (MLP) - 82%

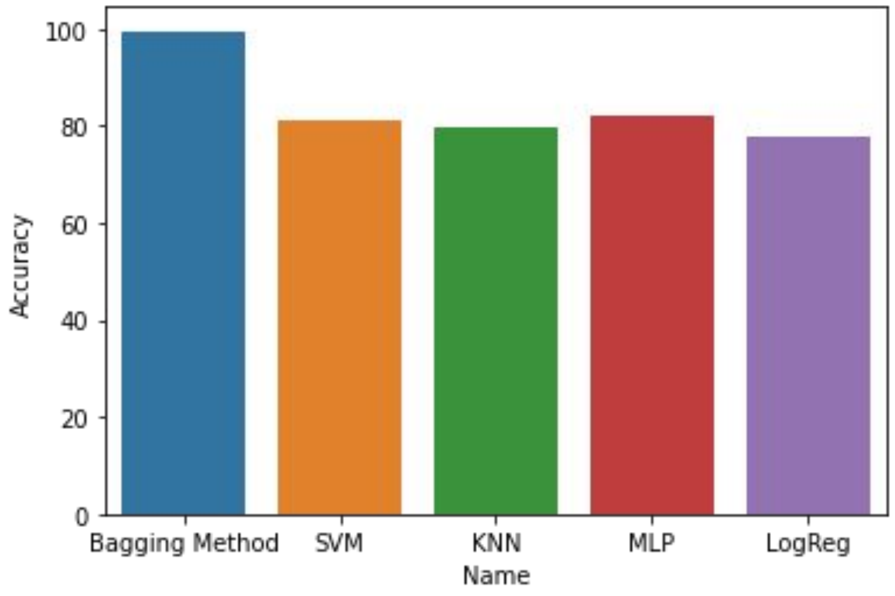
```
Our accuracy scores:
 [1.0, 1.0, 0.9880952380952381]
```

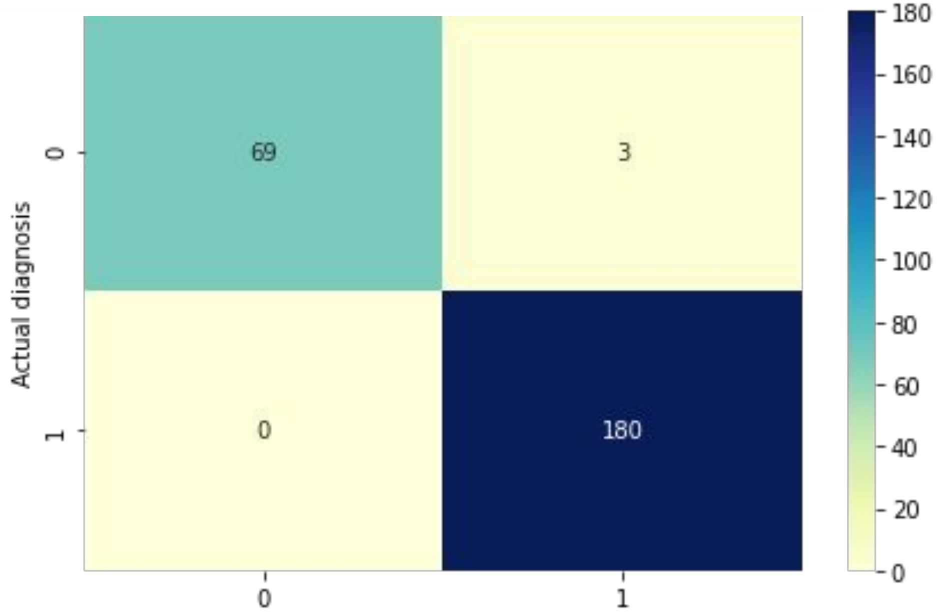
<u>Bagging Model</u>
Obtained 100% ACCURACY for 2 out of 3 validation sets, 98.81% accuracy for last one → averages out to 99.60% accuracy

For our last set, we had 3 errors:

- False Positives

Statistically more ideal - false positives are less damaging than false negatives





Conclusions (I)

Our Model

Significantly outperformed the last models trained on this dataset in the University of Istanbul study
Accuracy is high enough.

Our findings definitely addresses the question we proposed, and taught us that the results we managed to get were better than we could have thought. Our models were trained uniquely in a way never seen before

Room for Improvement

Dataset was small -- has two key impacts

a) Came from one demographic -- does not truly reflect diversity in medical cases
b) Runs the risk of overfitting

Both of these can be a setback in real-world application

Conclusions (II)

Applications

Packaged into software -- has great application in **telehealth** (distribution of health-related services and information via electronic information and telecommunication technologies)

The number of telehealth primary care visits increased 350-fold from pre-pandemic levels. (HHS)

Telemedicine consults resulted in **cost savings** by diverting patients from more expensive care settings. (The American Journal of Emergency Medicine)

→ Remote solutions mean more accessible, COVID-safe diagnoses. The accuracy of our model is unheard of

References

Bosworth A, Ruhter J, Samson LW, Sheingold S, Taplin C, Tarazi W, and Zuckerman R, Medicare Beneficiary Use of Telehealth Visits: Early Data from the Start of COVID-19 Pandemic. Washington, DC: Office of the Assistant Secretary for Planning and Evaluation, U.S. Department of Health and Human Services. July 28, 2020.

Castillo-Barnes, Diego, et al. “Robust Ensemble Classification Methodology for I123-Ioflupane SPECT Images and Multiple Heterogeneous Biomarkers in the Diagnosis of Parkinson's Disease.” Frontiers in Neuroinformatics, vol. 12, 2018, doi:10.3389/fninf.2018.00053.

Sakar, Betul Erdogdu, et al. “Collection and Analysis of a Parkinson Speech Dataset With Multiple Types of Sound Recordings.” IEEE Journal of Biomedical and Health Informatics, vol. 17, no. 4, July 2013, pp. 828–834., doi:10.1109/jbhi.2013.2245674.

Shahbakhti, Mohammad, et al. “Linear and Non-Linear Speech Features for Detection of Parkinson's Disease.” The 6th 2013 Biomedical Engineering International Conference, 2013, doi:10.1109/bmeicon.2013.6687667

Skodda, Sabine, and Uwe Schlegel. “Speech Rate and Rhythm in Parkinson's Disease.” Movement Disorders, vol. 23, no. 7, 2008, pp. 985–992., doi:10.1002/mds.21996.

Farrús, Hernando, and Ejarque. “Jitter and Shimmer Measurements for Speaker Recognition”, TALP Research Center, Department of Signal Theory and Communications Universitat Politècnica de Catalunya, Barcelona, Spain

Image Sourcing:

- https://www.parkinson.org/understanding-parkinsons/what-is-parkinsons
-

Virtual Lab Notebook Excerpt 1

02/01
Went to the paper developed in Istanbul and did comprehensive research on the relevance of Parkinson’s it brought up.

02/02
Emailed professors for a copy of their paper to cite.

02/04
Received paper.

Decided to briefly change project focus— took it from a general project study to a more in-depth analysis of how voice analysis could directly impact the detection of Parkinson’s. We observed and read many studies which had attempted to look for similar results, but our experiment remained unique.

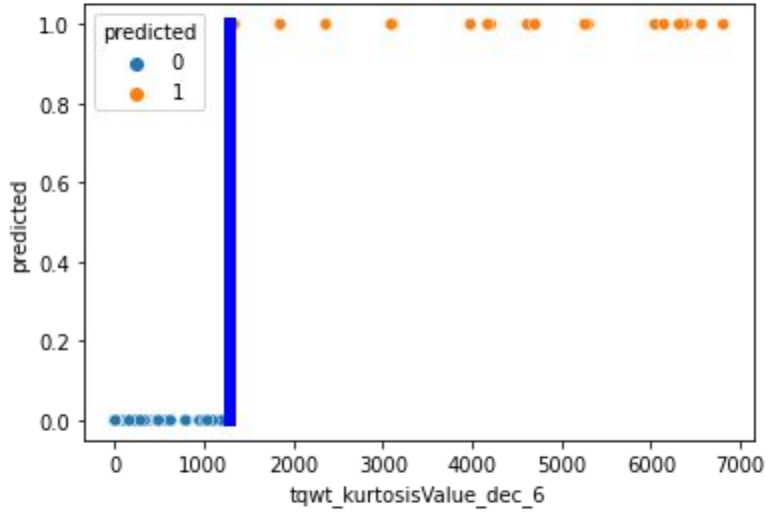
Saw most studies used KNN (K-Nearest Neighbors) algorithms. Proposed finding a way around that and potentially using an ANN (Artificial Neural Network) instead to maintain accuracy and fulfil project uniqueness.

Virtual Lab Notebook Excerpt 2

02/07
Performed a preliminary guess-and-check model using a boundary classifier. The accuracy was terrible, around 27%, but a good starting point.

Then implemented logistic regression using two values: the kurtosis values provided in the initial data set and the number of pulses provided. Results as seen:

- Accuracy - 0.78
- Precision - 0.83
- Recall - 0.95



02/10
Finished Principal Component Analysis. Decided upon which factors to keep in our final analysis and ended up with only four. Allows us to start data processing with robust numbers.