“Social Engagement Embeddings of Parkinson’s Disease through Autoencoders”

Authors: T. Xiao, P. Guha, A. S. Tajuddin

IARIA HEALTHINFO 2024 CONFERENCE

Presenter- Poulomi Guha
PhD Scholar, University of North Texas

Nice, France, November 3rd - 7th , 2024

UNT

IARIA

Poulomi Guha

Education:

- University of North Texas, PhD, Computer Science and Engineering
- EURECOM, Masters1 in *Data Science and Engineering*
- Techno India University, Bachelors in *Data Science*

Experience:

- Graduate Teaching Assistant at the University of North Texas

“I often say now I don't have any choice whether or not I have Parkinson's, but surrounding that non-choice is a million other choices that I can make.”

—**Michael J. Fox**

Dataset Overview

Source: Public dataset from King’s College London (KCL).
Device: Audio recordings from a Motorola Moto G4 smartphone.

Assessments:

Hoehn & Yahr (H&Y) Scale:	**UPDRS III-18**:	**UPDRS II-5**
‘2’: Unimpaired bilateral movements	‘0’: Normal	‘0’: Normal
‘3’: Postural impairment	‘1’: Slight loss of dictation	higher scores indicate increasing severity.
‘4’: Needs assistance with regular activities	‘2’: Moderately impaired	

•**Participants**:15 PD subjects, 21 controls

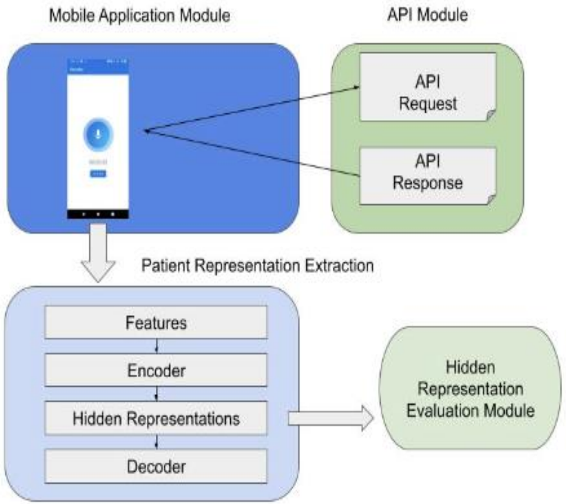
Material and Methods

API Module

- ✓ Built using Node.js & Google Speech-to-Text API.

Representation Learning

- ✓ Autoencoder used to derive hidden representations of subjects, capturing key patterns in the data.



Mobile Application

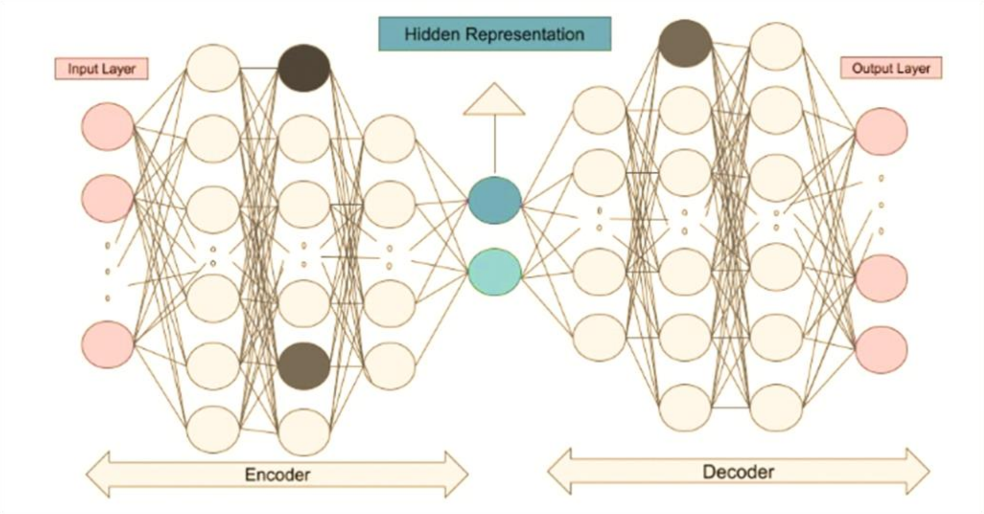
- ✓ Records user audio and extracts features via the API.

Classification & Evaluation

- ✓ Used KNN, SVM, XGBoost, Random Forest to classify PD vs. controls.
- ✓ Hidden representations correlate with UPDRS scores

Model Architecture

- We propose a 9-layer,a fully connected architecture stacked autoencoder.
- The autoencoder (AE) comprises three main parts: an encoder, a decoder, and a middle code representing the hidden layer
- The encoder consists of an input layer followed by a second layer with 100 neurons, a dropout layer set at 0.1, and a third layer with 30 neurons. The decoder mirrors the encoder in reversed order, aiming to reconstruct the original input at the output layer.
- The middle layer is fixed at 2 neurons, serving as the hidden representation.

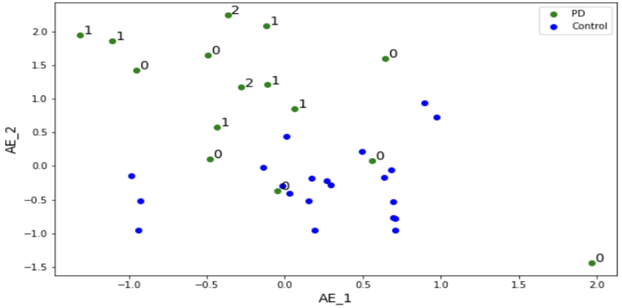


Hyperparameter Tuning

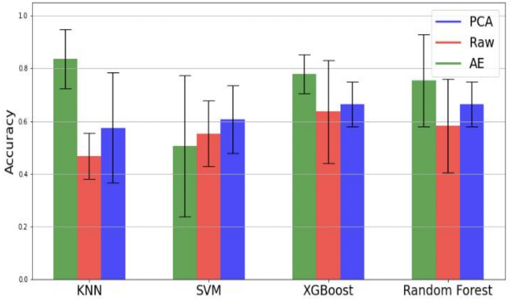
- **Neuron Counts**:
  - ✓ Encoder: First layer (64, 96, 100, 128), Third layer (8, 10, 16, 20, 24, 30, 32)
  - ✓ Decoder: Reverse of encoder structure
- **Activation Functions**: ReLU, Tanh
- **Learning Rates**: 0.1, 0.01, 0.001

Classification of Models

- **Classifiers**: KNN, SVM, Random Forest, XGBoost
- **Feature Sets Used**:
  - ✓ Autoencoder (AE) hidden representations
  - ✓ Raw features
  - ✓ 2D Principal Component Analysis (PCA) features
- **Autoencoder Hyperparameters**:
  - ✓ **Encoder**: Dense (100, 30 units, tanh), Dropout (0.1), Code (2 units)
  - ✓ **Decoder**: Reverse structure of encoder
- **Scaling**: Min-Max scaling applied to all data
- **Evaluation**: 5-fold cross-validation, average validation results reported



Autoencoder hidden representation visualization for control (blue) and PD (green) with UPDRS-II-5 rating



Mean accuracy comparisons using PCA (blue), AE (green), and raw (red) features.

Results

- The Autoencoder (AE) captures more variance than PCA, enabling better feature representation for downstream analysis.

- PD embeddings, normalized against control subject embeddings, were classified using KNN, XGBoost, Random Forest, and SVM.

Model	Features	Macro precision	Macro recall	Macro F1
KNN	AE	0.90	0.80	0.80
	RAW	0.26	0.40	0.32
	PCA	0.45	0.53	0.47
XGBoost	AE	0.81	0.76	0.75
	RAW	0.66	0.62	0.62
	PCA	0.67	0.61	0.57
Random Forest	AE	0.76	0.73	0.72
	RAW	0.62	0.55	0.54
	PCA	0.62	0.60	0.54
SVM	AE	0.48	0.50	0.48
	RAW	0.57	0.54	0.52
	PCA	0.56	0.55	0.50

Discussions

- This study aims to enhance insights into conversations for individuals with speech impairments through a mobile application that utilizes a feature extraction API powered by Google Speech Recognition to extract 42 features.

- An autoencoder (AE) was developed, capturing more variance than PCA and improving classification accuracy for Parkinson's Disease (PD) versus control subjects, with KNN achieving 90% macro precision. XGBoost and Random Forest also showed notable improvements using AE features.

- Although raw-SVM outperformed both AE and PCA.

➢ **Why?**

This is because KNN, XGBoost, and Random Forest benefit from the noise reduction and complex pattern representation in the transformed space. SVM, however, performed better with the original features, likely due to its effectiveness in using simpler, direct features for maximizing class separation.

Conclusions

- In summary, we have developed a user-friendly application that extracts features from interactive conversations.

- We then introduced an autoencoder-based model that generates reduced representations of individuals’ social engagement features.

- These hidden representations demonstrated the model’s ability to learn from diverse data and effectively distinguish between PD patients and control subjects

What’s Next?

Future Studies

- In the future, with greater data availability, we can develop more accurate models for predicting the severity of communication impairments.
- With more data availability, we can apply deep learning models to help the clinicians have better reports.

Acknowledgement

Dr. Ting Xiao

Ali Shah Solanki

Thasina Tabashum

Dr. Mark Albert

THANKS!

QUESTIONS?