Social Engagement Embeddings of Parkinson's Disease through Autoencoders

Ting Xiao
Department of Information Science
University of North Texas
Denton, USA
email: ting.xiao@unt.edu

Poulomi Guha
Department of Computer Science
University of North Texas
Denton, USA
email: poulomiguha@my.unt.edu

Ali Shah Tajuddin
Department of Computer Science
University of North Texas
Denton, USA
email: alishahlnu@my.unt.edu

***Abstract*—Various survey tools are available to measure social engagement, but they often suffer from infrequent measurement and recall bias. To address this, we developed a mobile application that estimates turn-taking in conversations and generates engagement features. These features were used to create an autoencoder-based hidden representation of individuals, which distinguishes between Parkinson's Disease and control subjects. The study aims to create reduced representations to robustly compare speaker-test outcomes with limited samples. An autoencoder was employed to reduce the number of features related to social engagement. This tailored assessment tool was applied to extract 42 speaker assessment scores, which were distilled into two-dimensional embeddings using a 9-layer autoencoder. We compared the proposed hidden representation with Principal Component Analysis, assessing metrics such as conversation percentage, turn-taking, and total pauses. These embeddings enabled a cross-validated reconstruction of all 42 features, accounting for 58% of the variance and were validated using multiple classification methods, including K-Nearest Neighbors (KNN), Support Vector Machine, Random Forest, and XGBoost. The KNN model, using the embeddings features, achieved a 90% macro precision score. Our results suggest that autoencoder representations provide a concise and effective tool for the holistic assessment of speaker behavior in limited data scenarios.**

Index Terms*—*Autoencoder*; *Parkinson's Disease*; *Hidden Representation*; *Mobile Application.

I. Introduction

Communication disorders affect 5% to 10% of the US population [1][2]. Individuals affected by these disorders exhibit impaired abilities in listening, speaking, writing, reading, and social interaction. The benchmark for rehabilitation is that these individuals must achieve the ability to communicate effectively and independently in natural environments following hospital discharge [3]. However, the treatment of cognitive communication issues are challenging due to the high variability in therapeutic approaches required for different patients. Notably, the potential disability of everyday communication effectiveness and independence termed Participation Restriction cannot be predicted from the nature or severity of the underlying speech or language impairment as assessed in the hospital [4][5]. Furthermore, to reduce financial expenditures, medical practitioners should focus on most outcomes in the shortest period of time due to the short duration of inpatient and outpatient rehabilitation. Even after in-hospital rehabilitation, residual impairments and associated participation restrictions are common and may last a lifetime [6], which may result in negative social outcomes [7] and emotional problems [8] for the person with the communication disorder. Thus, a valid assessment for potential communicative participation restriction is essential to augment long-term health outcomes and patient quality of life while reducing societal costs.

Language impairment is difficult to assess in neurodegenerative conditions such as Parkinson's Disease (PD) and Alzheimer's Disease (AD). There have been few studies on how to measure individuals with communication difficulties in their natural setting, such as at home, and the evaluation methods used can be highly diverse, such as post-hoc and point-in-time self-report scales [9][10]. As a result, bridging the gap is a significant issue for researchers. Additionally, speech evaluations require detailed protocols and are influenced by factors like geographic location and physician availability. Fortunately, as technology has advanced in at-home devices, cell phone applications to monitor speech have been developed [11][12]. Mobile phones are affordable and accessible and have enormous potential to assess speech [13][14]. In this study, we investigated the use of speech and language features to distinguish between PD and control subjects and developed a mobile application designed to easily capture these features.

Communication impairment is a prevalent symptom of PD. Research demonstrates varied impacts: Lang et al. highlighted impaired pragmatic communication in PD [15], Robinsons et al. observed reduced spontaneity [16], and Dushanova et al. linked changes in verb and noun usage to the disease [17]. Despite these studies, there is no automated globally approved scale that considers speech and language features. This gap has spurred the use of data-driven deep learning techniques to refine speech assessment. For example, Yeung et al.'s analysis of speech characteristics confirmed clinician agreement on features such as word finding difficulties, which correlate with the number of pauses, word duration, and syntactic complexity—key indicators of language impairment [17][18]. Orozco-Arroyave evaluated nonlinear dynamics features and showed that up to 76.81% accuracy could be achieved using the utterance of vowels [19]. Berus et al. achieved an accuracy of 86.47% by applying multiple feature selection processes and proposed feedforward artificial neural networks to classify PD [20]. Additionally, Tsanas et al. developed a machine learning model integrated system that assesses both PD subjects and the quality of their speech [21]. These studies indicate the potential of using deep learning algorithms to assist clinicians in accurately diagnosing PD. In this essence, we adopted a data-driven approach by extracting speech and language features to accurately identify PD.

Clinicians rely on a variety of clinical measures, many of which are collected by automated systems and used for judgment. However, the abundance of individual metrics can create challenges in synthesizing information and drawing clear conclusions. To overcome this issue, lowering the size of the data and developing summary metrics can help. Conventional strategies for reducing dimensions include Principal Component Analysis (PCA) [22] and Linear Discriminant Analysis (LDA). However, these linear approaches fail to capture the non-linear correlations between input data, resulting in a less efficient summary representation. Recently, neural network techniques such as autoencoders have shown promise in learning lower representations of high-dimensional data for natural language processing and computer vision tasks. AutoEncoders (AEs) are neural network algorithms that learn hidden representations of high-dimensional data [23][24].

Previous works have shown that AEs can learn meaningful hidden representations that lead to better downstream tasks, including Ng et al. using speech signals and demonstrated autoencoder-based

 ISBN: 978-1-68558-204-3

representation to distinguish disordered speech PD [25]. In order to correctly classify PD by accurate interpretation of the speech and vocal data, Caliskan et al. suggested stacked autoencoder embeddings [26]. Gunduz et al. used a convolution neural network with a vocal feature set to differentiate between PD and control subjects [27]. Hoq et al. compared several models to classify PD and control subjects. Based on a Support Vector Machine (SVM) [28] integrated with a Principal Component Analysis (PCA) and a Sparse AutoEncoder (SAE), the results demonstrated that the proposed SAE-SVM model outperformed not only the PCA-SVM and standard models such as Multilayer Perceptron (MLP), XGBoost [29], K-Nearest Neighbors (KNN), and Random Forest (RF), but also surpassed two recent studies on the same dataset. Additionally, performance was further enhanced by applying SMOTE for oversampling and dataset balancing [30].

Using a data-driven approach will enhance early diagnosis and reduce the amount of time for the diagnostic process. For this project, we developed a prototype mobile application that simplifies the extraction of necessary attributes. We utilized an AutoEncoder (AE) to learn the hidden representations of subjects, classifying between PD and control using embeddings in KNN, SVM, RF, and XGBoost. Most importantly, we demonstrated that hidden representations could effectively capture the full spectrum of an individual's social participation, proving that AE is a robust method for generating meaningful representations of individuals.

The rest of the paper is structrued as follows. In Section II, we present the materials and methods used. The results are shown and discussed in Section III. We conclude the work in Section IV.

II. MATERIALS AND METHODS

A. Study Design

We have four modules in this study, as shown in Figure 1. The first two modules are designed to extract features efficiently, while the last two demonstrate how these features can be utilized in downstream analysis.

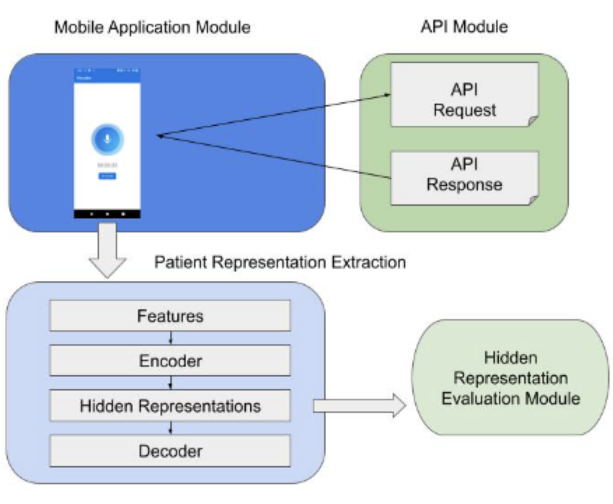

Figure 1. Study Design: Four modules: 1) API, 2) Mobile Application, 3) Patient Representation Extraction, and 4) Evaluation.

1) API: We have built a web API using node.js and Google Speech-to-Text API. This API facilitates direct feature extraction and can be integrated with various types of applications, including web, mobile, and desktop platforms. The API is publicly available on GitHub [31].

2) Mobile Application: To utilize the API, we have developed a mobile application that records the audio, allowing users to extract features directly in a structured format. The prototype of the mobile application is illustrated in Figure 2.

TABLE I
NUMBER OF SUBJECTS AND H&Y, UPDRS II-5, UPDRS III-18 SCORES

	Severity	Number of Subjects
H&Y	2	8
	3	6
	4	1
UPDRS II-5	0	7
	1	6
	2	2
UPDRS III-18	0	6
	1	6
	2	3

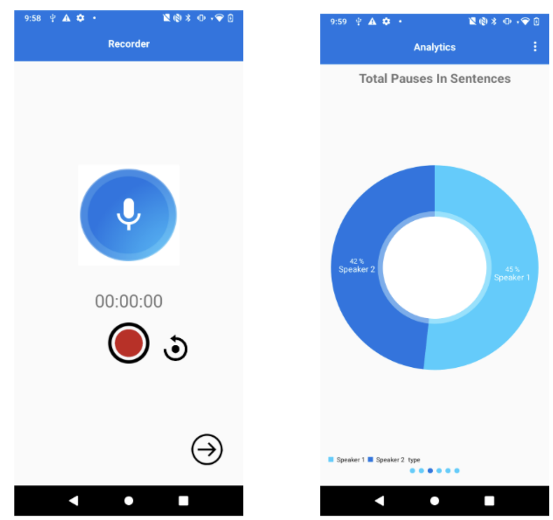

Figure 2. User Interface of Mobile Application: a) (left) recorder, b) (right) demonstration of total speaking time between two speakers.

3) Patient Representation Extraction: The extracted features were then utilized to derive the subjects' hidden representations. We employed an AE approach to learn these representations.

4) Evaluation: We used the hidden representations to distinguish between PD and controls by applying classification algorithms such as KNN, SVM, XGBoost and RF. Moreover, we demonstrated that these hidden representations of the subjects are distinguishable based on their Unified Parkinson's Disease Rating Scale (UPDRS) scores.

B. Dataset

We used a publicly available dataset from King's College London (KCL) [32], with audio recordings made using a Motorola Moto G4 smartphone. This dataset includes assessments such as the Hoehn and Yahr (H&Y) scale, UPDRS III-18, and UPDRS-II-5 scores. H&Y indicates PD progression stages: '2' for unimpaired bilateral movements, '3' for postural impairment, and '4' for needing assistance with regular activity [33]. The UPDRS III-18 score, assessing the motor examination of speech, ranges from '0' (normal), '1' (slight loss of dictation), to '2' (moderately impaired) [34][35]. The UPDRS II-5 score evaluates daily speech, where '0' is normal and higher scores indicate increasing severity. The study comprised 15 PD subjects and 21 control subjects. The distribution of patients across H&Y and UPDRS is detailed in Table I.

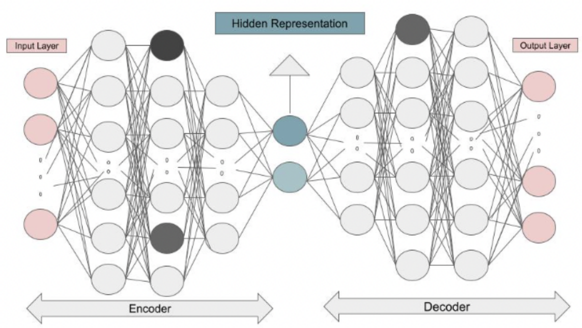

Figure 3. 9-layered stacked autoencoder architecture. Encoder and decoder both consist of 3 fully connected layers and one drop out layer.

C. Feature Extraction

Using our API [31], we extracted a total of 42 features as listed below, focusing on speaker segmentation and conversation dynamics.

- Total time of each speaker with all the pauses
- Total pause time during a conversation
- Gap between the turns
- Continuous repeating word
- Percentage of speaking time in the conversation
- Total turns
- Total unique words
- Average word length
- Percentage of total first half speaking time
- Percentage of total last half speaking time
- Total conversation duration between 2 people

Additionally, we quantified each Part Of Speech (POS) in the conversations [36] utilizing the Python Natural Language ToolKit (NLTK) [37].

D. Autoencoder Architecture

We proposed a 9-layer stacked AE, depicted in Figure 3, featuring a fully connected architecture. The 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 reverse order, aiming to reconstruct the original input at the output layer. The middle layer is fixed at 2 neurons, serving as the hidden representation.

For each subject, we extracted two-dimensional hidden representations. To standardize these embeddings, we normalized the PD subjects' data relative to the control group by calculating the mean and standard deviation for each dimension among the controls and adjusting the PD values using the following equation:

$$AE\_scaled_{PD} = (Mean(AE_{control})) - (AE_{PD}/std(AE_{control}))$$

E. Hyperparameter Tuning

To optimize the hyperparameters of the AE, we employed a grid search strategy. We tested various neuron counts for the first dense layer (64, 96, 100, 128) and the third dense layer (8, 10, 16, 20, 24, 30, 32) in the encoder. In the decoder, the layers were structured with the same number of neurons as the encoder but in reverse order. Rectified Linear Unit (ReLU) and tanh were applied in hidden layers as activation functions and different learning rates (0.1, 0.01 and 0.001) were explored. After the grid search, we chose the hyperparameters that captured the most variance, as listed in Table II.

TABLE II
HYPERPARAMETERS OF AE

	Layers	Units	Activation Functions
Encoder	Input Layer	42	
	Dense Layer	100	tanh
	Dropout	0.1	
	Dense Layer	30	tanh
Code	Dense	2	tanh
Decoder	Dense Layer	30	tanh
	Dropout	0.1	
	Dense Layer	100	tanh
	Output Layer	42	linear

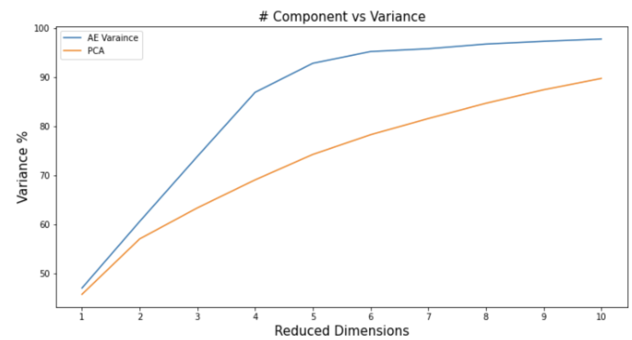

Figure 4. Reduced dimension vs variance. AE captures more variance than PCA.

F. Classification Models

For comparisons, we employed traditional classifiers including KNN, SVM, RF and XGBoost algorithms. These algorithms were fed with two sets of AE hidden representations, raw features, and two-dimensional PCA features to classify PD versus control subjects. All data were scaled using min-max scaling. Hyperparameters were selected via a grid search conducted on the raw 42 features, and these parameters were then applied consistently across all models. Model performance was assessed using 5-fold cross-validation, and we reported the average validation results from the 5 folds.

III. RESULTS AND DISCUSSION

Figure 4 illustrates the variance, explained by the trained AE compared to PCA, clearly showing that AE captures more variance. For downstream analysis, we took our 2 hidden representations generated by the encoder and applied normalization to the PD embeddings based on the control subject embeddings. Then, we classified these embeddings using KNN, XGBoost, RF, and SVM algorithms. Figure 5 compares the accuracy of three data modalities across four classifiers and Figure 6 shows the 2D embeddings with UPDRS-III-18 scores, while the mean macro precision, recall, and F1-scores from a 5-fold validation are reported in Table III.

The aim of this study is to enhance insights into conversations via our mobile application, particularly for individuals with speech impairments who need monitoring of their social engagement. The feature extraction API leverages the Google speech recognition system to extract 42 features [38], which are readily accessible through our mobile application designed for two-speaker settings, given the dataset constraints.

Utilizing these features, we developed an AE that captures more variance than PCA. The hidden representations extracted from the AE were used in downstream analysis, showing a significant improvement in classification accuracy for PD versus control subjects, with KNN achieving a 90% macro precision. XGBoost and RF also

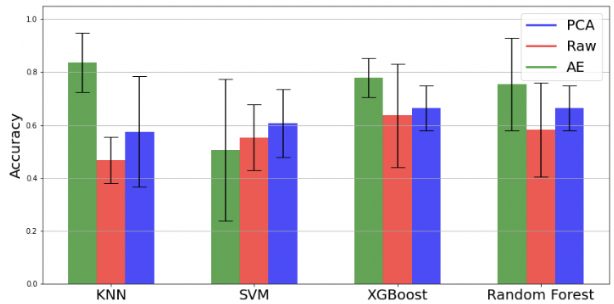

Figure 5. Mean accuracy comparisons using PCA (blue), AE (green), and raw (red) features.

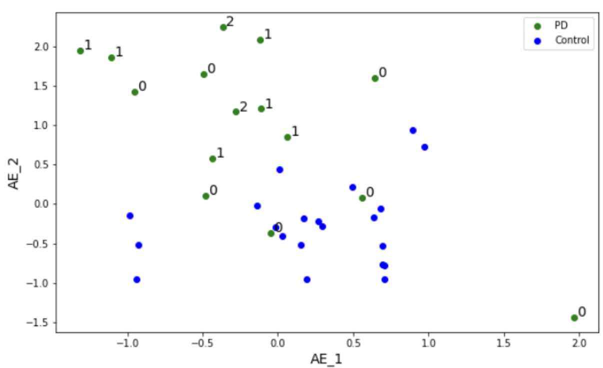

Figure 6. Autoencoder hidden representation visualization for control (blue) and PD (green) with UPDRS-II-5 rating.

showed notable improvements using AE features, although raw-SVM outperformed both AE and PCA. This is because KNN, XGBoost, and RF 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.

Furthermore, in Figure 6, the hidden representations of three patients, who are close to controls and in the early stages according to the H&Y, UPDRS II-5, and UPDRS III-18 scales, were misclassified as controls by the classification algorithms. This misclassification suggests that their communication skills may not be significantly affected. Additionally, the existing rating systems vary from one to another. For instance, there are seven subjects at stage 0 according to UPDR II-5, but only six according to UPDRS III-18. To address these inconsistencies, we propose using an AE to develop a unified scaling system for assessing subjects. A limitation of this study is the small sample size and the use of control subjects' hidden representations as a baseline for scaling. With a larger dataset, we could directly establish a scale that more accurately evaluates communication skills. This pilot study demonstrates the utility of embeddings in distinguishing PD from control subjects, with less severe PD cases tending to cluster closer to controls, as depicted in Figure 6.

Although several previous studies have tried to distinguish between PD and control subjects using deep learning algorithms to assist clinicians in accurately diagnosing PD, many of them are not designed for easy integration with new environments or for extracting features from different datasets. Our proposal includes an application and a versatile API that can be integrated into any platform, offering clinicians deeper insights into patient conversations and social interactions. In the future, with greater data availability, we can develop more accurate models for predicting the severity of communication impairments.

TABLE III
MEAN MACRO PRECISION, RECALL AND F1 SCORE COMPARISONS

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

IV. CONCLUSION

In this paper, we have developed a user-friendly application that extracts features from interactive conversations. We then introduced an AE-based model that generates reduced representations of individuals' social engagement features. Through these hidden representations, the model enabled classification between PD and control groups using embeddings employed by KNN, SVM, RF, and XGBoost classifiers. Our findings showed that these embeddings capture diverse data patterns and effectively distinguish PD patients from control subjects, demonstrating that these representations can encapsulate the full range of an individual's social participation. This highlights the AE model's value in creating meaningful representations for assessing social engagement.

Moving forward, enhancing this model with additional conversational and multimodal features, such as gesture or facial expression data, could further improve its accuracy and adaptability to real-world scenarios. Such expansions would support clinicians in monitoring PD symptoms more effectively and could potentially enable early interventions based on real-time social interaction insights. This study highlights the potential of embedding-based methods in healthcare, offering a pathway toward practical, non-invasive tools that can aid in diagnosing and managing communication disorders through assessment of social engagement patterns.

ACKNOWLEDGEMENT

The authors would like to thank Thasina Tabashum and Dr. Mark Albert for their help and support in making this work possible. Their expertise and encouragement have been instrumental in enabling us to achieve our research goals.

REFERENCES

[1] R. J. Ruben, "Redefining the survival of the fittest: Communication disorders in the 21st century," *The Laryngoscope*, vol. 110, no. 2, pp. 241–241, 2000.

[2] R. Body, M. Perkins, and S. McDonald, "Pragmatics, cognition, and communication in traumatic brain injury," in *Communication Disorders Following Traumatic Brain Injury*. Psychology Press, 2016, pp. 81–112.

[3] C. K. Thompson, "The American Speech-Language-Hearing Association Functional Assessment of Communication Skills ASHA FACS: Test description and pilot data," in *A Paper Presented at the Clinical Aphasiology Conference, Sunriver, OR, 1995*, 1995, pp. 75–78.

[4] World Health Organization, *International Classification of Functioning, Disability, and Health: Children & Youth Version: ICF-CY.* World Health Organization, 2007.

[5] R. O'Halloran, M. Carragher, and A. Foster, "The consequences of the consequences: The impact of the environment on people with aphasia over time," *Topics in language disorders*, vol. 37, no. 1, pp. 85–100, 2017.

[6] H. L. Flowers, S. A. Skoretz, F. L. Silver, E. Rochon, J. Fang, C. Flamand-Roze, and R. Martino, "Poststroke aphasia frequency, recovery, and outcomes: A systematic review and meta-analysis," *Archives of Physical Medicine and Rehabilitation*, vol. 97, no. 12, pp. 2188–2201, 2016.

[7] K. Brown, B. Davidson, L. E. Worrall, and T. Howe, ""Making a Good Time": The role of Friendship in Living Successfully with Aphasia," *International Journal of Speech-Language Pathology*, vol. 15, no. 2, pp. 165–175, 2013.

[8] K. Hilari, S. Northcott, P. Roy, J. Marshall, R. D. Wiggins, J. Chataway, and D. Ames, "Psychological distress after stroke and aphasia: The first six months," *Clinical Rehabilitation*, vol. 24, no. 2, pp. 181–190, 2010.

[9] D. R. Paul, C. Frattali, A. L. Holland, C. K. Thompson, C. J. Caperton, and S. C. Slater, *Quality of Communication Life Scale: Manual*. American Speech-Language Hearing Association, 2004.

[10] P. Sukanya and B. S. Rao, "Parkinson's disease classification based on stacked denoising autoencoder." *Computer Science*, vol. 24, no. 4, 2023.

[11] T. Xiao, T. Tabashum, B. Metwally, M. Albert, A. Du, R. Jebamalaidass, M. Leal, and E. Oliviera, "Conversation moderator: A mobile app for tracking individual speaking in group conversations," in *2020 IEEE 14th International Conference on Semantic Computing (ICSC)*. IEEE, 2020, pp. 430–433.

[12] C. G. Goetz *et al.*, "Testing objective measures of motor impairment in early parkinson's disease: Feasibility study of an at-home testing device," *Movement Disorders*, vol. 24, no. 4, pp. 551–556, 2009.

[13] J. Rusz, J. Hlavnička, T. Tykalová, M. Novotnỳ, P. Dušek, K. Šonka, and E. Ržička, "Smartphone Allows Capture of Speech Abnormalities Associated with High Risk of Developing Parkinson's Disease," *IEEE Transactions on Neural Systems and Rehabilitation Engineering*, vol. 26, no. 8, pp. 1495–1507, 2018.

[14] A. Tsanas, M. A. Little, and L. O. Ramig, "Remote assessment of parkinson's disease symptom severity using the simulated cellular mobile telephone network," *IEEE Access*, vol. 9, pp. 11 024–11 036, 2021.

[15] P. McNamara and R. Durso, "Pragmatic Communication Skills in Patients with Parkinson's Disease," *Brain and Language*, vol. 84, no. 3, pp. 414–423, 2003.

[16] G. A. Robinson, "Primary Progressive Dynamic Aphasia and Parkinsonism: Generation, Selection and Sequencing Deficits," *Neuropsychologia*, vol. 51, no. 13, pp. 2534–2547, 2013.

[17] J. Dushanova, *Diagnostics and Rehabilitation of Parkinson's Disease*. BoD–Books on Demand, 2011.

[18] A. Yeung *et al.*, "Correlating natural language processing and automated speech analysis with clinician assessment to quantify speech-language changes in mild cognitive impairment and alzheimer's dementia," *Alzheimer's Research & Therapy*, vol. 13, no. 1, p. 109, 2021.

[19] J. R. Orozco-Arroyave, *Analysis of Speech of People with Parkinson's Disease*. Logos Verlag Berlin GmbH, 2016, vol. 41.

[20] L. Berus, S. Klancnik, M. Brezocnik, and M. Ficko, "Classifying parkinson's disease based on acoustic measures using artificial neural networks," *Sensors*, vol. 19, no. 1, p. 16, 2018.

[21] A. Tsanas, M. A. Little, C. Fox, and L. O. Ramig, "Objective automatic assessment of rehabilitative speech treatment in parkinson's disease," *IEEE Transactions on Neural Systems and Rehabilitation Engineering*, vol. 22, no. 1, pp. 181–190, 2013.

[22] K. Pearson, "LIII. On Lines and Planes of Closest Fit to Systems of Points in Space," *The London, Edinburgh, and Dublin philosophical magazine and journal of science*, vol. 2, no. 11, pp. 559–572, 1901.

[23] D. E. Rumelhart, G. E. Hinton, and R. J. Williams, "Learning representations by back-propagating errors," *Nature*, vol. 323, no. 6088, pp. 533–536, 1986.

[24] S. Nandigam, T. Tabashum, and T. Xiao, "Autoencoders and embeddings: How unsupervised structural learning enables fast and efficient goal-directed learning," in *Bridging Human Intelligence and Artificial Intelligence*. Springer, 2022, pp. 47–63.

[25] S.-I. Ng and T. Lee, "Automatic Detection of Phonological Errors in child Speech Using Siamese Recurrent Autoencoder," *arXiv preprint arXiv:2008.03193*, 2020.

[26] A. Caliskan, H. Badem, A. Basturk, and M. Yuksel, "Diagnosis of the parkinson disease by using deep neural network classifier," *IU-Journal of Electrical & Electronics Engineering*, vol. 17, no. 2, pp. 3311–3318, 2017.

[27] H. Gunduz, "Deep learning-based parkinson's disease classification using vocal feature sets," *IEEE access*, vol. 7, pp. 115 540–115 551, 2019.

[28] N. Cristianini and J. Shawe-Taylor, *An Introduction to Support Vector Machines and Other Kernel-Based Learning Methods*. Cambridge University Press, 2000.

[29] T. Chen and C. Guestrin, "XGBoost: A Scalable Tree Boosting System," in *Proceedings of the 22nd ACM SIGKDD International Conference on Knowledge Discovery and Data Mining*, ser. KDD '16, vol. 11. ACM, Aug. 2016, p. 785–794. [Online]. Available: http://dx.doi.org/10.1145/2939672.2939785

[30] M. Hoq, M. N. Uddin, and S.-B. Park, "Vocal feature extraction-based artificial intelligent model for parkinson's disease detection," *Diagnostics*, vol. 11, no. 6, p. 1076, 2021.

[31] T. Xiao, P. Guha, T. Thasina, and A. Tajuddin (2024). Conversation Feature Extraction Tool. Deep 6 Lab. https://github.com/Deep6Lab/Conversation-Feature-Extraction-Tool. Accessed: 2024-10-31.

[32] H. Jaeger, D. Trivedi, and M. Stadtschnitzer, "Mobile device voice recordings at king's college london (mdvr-kcl) from both early and advanced parkinson's disease patients and healthy controls," *Zenodo*, 2019.

[33] M. M. Hoehn and M. D. Yahr, "Parkinsonism: Onset, progression, and mortality," *Neurology*, vol. 17, no. 5, pp. 427–427, 1967.

[34] Parkinson's Disease Research, Education and Clinical Centers, U.S. Department of Veteran Affairs. https://www.parkinsons.va.gov/resources/UPDRS.asp. Accessed: 2024-10-31.

[35] N. Diaz and J. M. Bronstein, "Parkinson's disease research education and clinical centers (PADRECC): background and overview," *NeuroRehabilitation*, vol. 20, no. 3, pp. 153–160, 2005.

[36] S. Bird, E. Klein, and E. Loper, *Natural Language Processing with Python*, 01 2009.

[37] B. Steven, E. Loper, and E. Klein (2009). Natural Language Processing with Python. O'Reilly Media Inc. https://www.nltk.org/. Accessed: 2024-10-31.

[38] F. Filippidou and L. Moussiades, "*A* Benchmarking of IBM, Google and Wit Automatic Speech Recognition Systems," in *16th IFIP International Conference on Artificial Intelligence Applications and Innovations (AIAI)*, ser. IFIP Advances in Information and Communication Technology, I. Maglogiannis, L. Iliadis, and E. Pimenidis, Eds., vol. AICT-583, no. Part I. Neos Marmaras, Greece: Springer International Publishing, Jun. 2020, pp. 73–82, part 1: Classification. [Online]. Available: https://inria.hal.science/hal-04050587

