International Journal of
INTELLIGENT SYSTEMS AND APPLICATIONS IN ENGINEERING
IJISAE
ISSN:2147-6799 www.ijisae.org Original Research Paper

Automated Prognosis of Alzheimer's Disease using Machine Learning Classifiers on Spontaneous Speech features

Shridevi Karande*[1], Vrushali Kulkarni[2]

Submitted: 13/11/2022 **Accepted:** 15/02/2023

Abstract: Early prediction of Alzheimer's disease and related Dementia has been a great challenge. Recently, preliminary research has shown that neurological symptoms in Covid-19 patients may accelerate the onset of Alzheimer's disease. With such a further rise in Alzheimer's and related Dementia cases, having an early prediction system becomes vital. Speech can provide a non-invasive diagnostic marker for such neurodegenerative diseases. This work mainly focuses on studying significant temporal speech features extracted directly from the recordings of the Dementia bank dataset and applying Machine Learning algorithms to classify the Alzheimer's disease related Dementia Group and the healthy control group. The result shows that Support Vector Machine outperformed other machine learning algorithms with an accuracy of 87%. Compared to prior research, which used manual transcriptions provided with the dataset, this study used audio recordings from the Dementia bank dataset and an advanced Automatic Speech Recognizer to extract speech features from the audio recordings. Furthermore, this method can be applied to the spoken responses of subjects during a neuropsychological assessment.

***Keywords*:** *Alzheimer's disease, Covid-19, Dementia, Machine Learning Algorithm, Neuropsychological Assessment, Support Vector Machine*

1. Introduction

Alzheimer's disease and related Dementia are becoming more widespread in the world. As per World Health Organization Report [1], currently, dementia cases are found to be 47 million and may increase to 75 million by 2030. By 2050, dementia cases are expected almost triple. Alzheimer's Disease (AD) is the most common form of Dementia. Generally, families and friends tend to ignore the patient's first clinical symptoms and behavioural changes as it is often confused with normal ageing. Early diagnosis may help provide timely assistance and support to increase a patient's lifespan [2]. As per a recent study in AAIC 2021[3], it is seen that there is a strong correlation between AD and Dementia with the neurological symptoms found in Covid-19 patients. Therefore, it is imperative for patients to have access to a robust prediction system. Traditionally, Dementia has been diagnosed through neuropsychological assessment tests [4]. These tests help determine the extent to which the brain's cognitive functions are impaired. Mini-Mental State Examination (MMSE) is one of the most widely used screening tools consisting of a series of questions and images which take approximately 10-20 minutes to administer [5].

A person's cognitive level is said to be declining progressively through several stages at a preclinical stage, where symptoms are not noticed in a person visibly, but brain changes have begun.

Mild Cognitive Impairment (MCI) is an early stage where a person finds it difficult to remember and later, it advances into Dementia. Speech is one of the factors that is affected in these patients [6,9]. With the disease progression, it goes on deteriorating. In its early stage, a person finds it challenging to get the right word in spontaneous speech; this stage often goes unnoticed. In the middle stage, poor use of language and vocabulary is found in daily interactions [6] and at later stages, a person tends to use very few words. The progression of this deterioration can be reduced if these changes are caught at their mild stages.

The final aim of this study is to develop a prediction system which would analyse the speech parameters extracted from a spontaneous speech. This paper mainly explores the speech parameters which can help distinguish Alzheimer's Disease group from the healthy group. Section 2 covers a review of the relevant related work on speech features, Section 3 describes the dataset, method of feature extraction and various ML models used in the analysis, Section 4 discusses results and Section 5 is the conclusion.

[1] School of Computer Engineering and Technology, Dr. Vishwanath Karad MIT World Peace University, Maharashtra, India.
ORCID ID: 0000-0001-8219-2189
[2]School of Computer Engineering and Technology, Dr. Vishwanath Karad MIT World Peace University, Maharashtra, India.
ORCID ID: 0000-0003-2982-8998
** Corresponding Author Email: shridevi.karande@mitwpu.edu.in*

2. Related Work

Earlier researchers have explored temporal, acoustic, syntactic and linguistic features extracted from speech to support the classification of disease groups. These features were extracted from various speech stimuli viz spontaneous speech, narrative tasks, picture description, event description, word list or paraphrasing. It was observed that the temporal parameters of speech can provide information about subtle impairments and can be a significant indicator in diagnosing AD. As the disease advances from mild to severe, the duration of hesitation (i.e., absence of speech for more than 30ms) was found to increase while the total number of phonemes decreased [7]. In another study [8], high pause rate and phonation time extracted from the spoken responses were found to be significant biomarkers for MCI diagnosis.

Researchers have compared the performance of the speech features extracted using manual and automatic transcription methods trained for a particular language. The features derived from Automatic Speech Recognizer (ASR) were observed to have accuracy similar to manually derived features [10]. The acoustic metrics like length of speech, number of silences, filled pauses and their duration were found to correlate with MCI diagnosis. In this study, SVM (Support Vector Machine) performed better on manually extracted features [14]. When the same study [19] was extended to automatic feature selection methods, it was noticed that the length of certain sounds within a content word indicated hesitation influencing the frequency of each phenome. This finding was missing in the manually extracted feature set.

Furthermore, the speech tempo and the number of pauses extracted from recordings of three different tasks viz. immediate Recall, spontaneous speech and delayed Recall were found to be significant. Here the SVM performance was better compared to other classifiers such as Naive Bayes and Random Forests [24]. Linguistic and Acoustic features [11-13] [15,17,18] can also help predict and identify AD. In another study [16], a speech analysis using discrimination of voiced and unvoiced segments over the groups (i.e., HC, MCI and AD) proved to be a better mechanism to evaluate the early stage of Dementia.

As most AD symptoms vary across individuals, a combination of various biomarkers and speech can help in sound examination at the earliest stage [20]. To support multimodal classification, along with the speech feature, language and image features can even help classify the groups [21]. Research showed that silence between the questions and the user's response (Gap) and coefficient of variation per utterance for the Japanese Language were significantly different in people with AD and healthy controls when trained with SVM [21-22]. For Detecting MCI from Speech Transcripts, a distinct approach [23] was used by creating a complex network (graph) of nodes (distinct words) and edges connecting similar words obtained using word embeddings.

3. Materials and Methods

The overall process consists of Dataset preparation, feature extraction and classification.

3.1. Dataset

The speech recordings from the Dementia bank project were used to create a dataset for this work. This data was created as a part of a clinical study on Alzheimer's and related Dementia. The information about the study cohort is available from Becker et al. [25]. The dataset was composed of English audio transcripts of verbal interviews conducted in English. The participants were categorized into Probable AD, Healthy Control, MCI. In this study, the participants were asked to describe a "Cookie Theft" picture which is a part of the Boston Diagnostic Aphasia Examination [26] and a narrative speech is recorded. This audio transcript starts with the examiner asking the patient to describe everything visible in the picture. In the case of less expressive patients, the examiner prompted them to speak more.

For this experiment, audio recordings comprising of 242 participants from the elderly Healthy CG (Control Group) and 309 elderly patients from the Alzheimer's Disease and related dementia Group (ADRG) were used. Within ADRG, 233 people were diagnosed with Probable AD, 21 Possible AD, 3 having Memory problems, 43 diagnosed with MCI, 5 with Vascular Dementia (VD) and 4 others (an unidentified form of Dementia).

3.2. Feature Extraction

The main objective of this study was to learn various speech parameters extracted from audio recordings of the Dementia bank project and use these features for classifying AD and related Dementia. Fig. 1 shows our system encompassing two modules, viz. Speech Features Extraction module and Model Training module. The input for the system was audio recordings from the Dementia bank and the Speech processing was done using IBM Watson's Speech-To-Text cloud service [29]. This service helped us extract audio transcripts and primary features like spoken words, start time and end time of all words and hesitations (such as 'ehm', 'umm'). As the recordings were of a conversation between the Examiner (Speaker1) and Patient (Speaker2), the primary features were extracted only for the Patient (Speaker 2).

The audio transcripts were produced for both the examiner and the patient. To choose highly correlating features, the p-Values were computed. The variables that correlated among themselves and those with a lower correlation coefficient value with the target variable were removed from the analysis. The metrics, as shown in Table1, were generated; it shows the various features and their p-Values. In order to identify significant features at each stage of the disease, a statistical analysis was conducted on various combinations of the Control group (CG) and all three groups MCI, ADRG, and Probable AD.

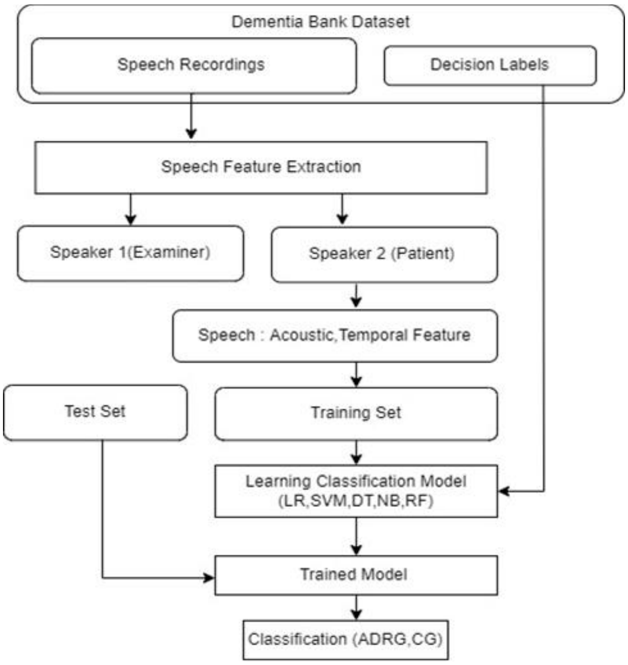

Fig. 1. Block Diagram

Furthermore, regression was also performed on the MCI group with Probable AD, to identify the influential parameters in disease progression from MCI to AD. The regression result showed that features such as Pause Count, Pause Rate, Mean Pause Duration, Hesitation Time and Hesitation Count significantly distinguish ADRG or Probable AD from CG. However, for distinguishing ADRG from CG, only Hesitation Time and Hesitation Count were significant. This indicates that people with AD and other related Dementia tend to produce more hesitation compared to the control group. Besides this finding, Phonation Time was also found to be shorter in probable AD than in CG. This finding correlates to the later stage of Dementia, where the total speech time excluding the pauses decreases, restricting the talk to very few words. However, the Hesitation Time and Hesitation Count were found to be significant in classifying MCI from CG.

Table 1. Logistic Regression Results

Features	p-Value			
	CG vs ADRG	CG vs MCI	CG vs Probable AD	MCI vs Probable AD
Pause Count	***0.044**	0.803	***0.027**	0.273
Pause Rate	0.428	0.563	0.469	0.661
Mean Pause Duration	***0.031**	0.272	***0.025**	0.556
Pause Time	0.728	0.109	0.916	0.333
Phonation Time	0.096	0.195	*0.012	0.069
Phonation Rate	0.837	0.065	0.978	0.241
Standardized Phonation Time	0.801	0.371	0.804	0.141
Verbal Rate	0.894	0.121	0.745	0.27
Hesitation Count	***0.029**	***0.006**	0.497	0.098
Hesitation Time	***0.006**	***0.003**	0.164	0.133
Word Count	0.349	0.154	0.069	0.183

*Indicate statistically vital features (p-value < 0.05)

ADRG-Alzheimer and related dementia Group(n=309); CG- Control Group(n=242); Probable AD(n=233), MCI (n=43)

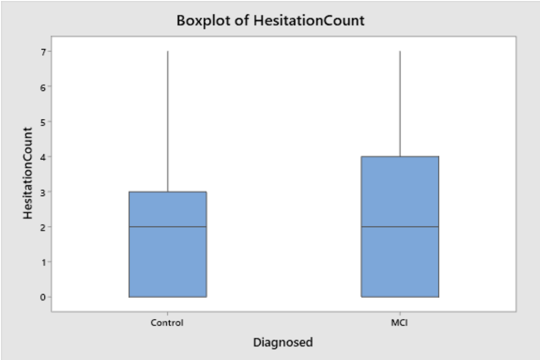

Fig. 2. Box Plot of Control Vs MCI on the basis of Hesitation Count

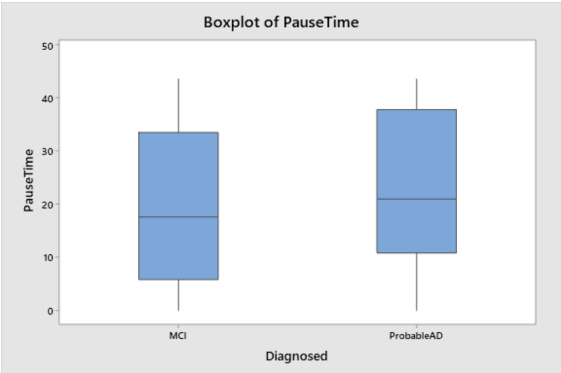

Fig. 3. Box Plot of MCI Vs Probable AD on the basis of Pause Time

As seen in Fig. 2, the hesitation count is found to be taking large range of values in case of MCI when compared with Control Group. However, when MCI and Probable AD classes are compared, it is observed that Probable AD participants having less hesitation count compared to MCI participants which could be possibly because in later stages instead of hesitation, they take longer pauses while describing the picture and can even be seen in Fig. 3.

3.3. Machine Learning Models

Here, AD detection is formulated as a binary classification problem. Five classifiers were executed on all features (shown in Table 1) and the significant features such as Age, Pause Count, Mean Pause Duration, Hesitation Count and Hesitation Time were identified in the feature selection stage.

A review of data and its distribution among both groups revealed that the SVM classifier performed well. Hence, a primary focus was provided to SVM in the model training. The Objective of SVM classifier is to find the hyperplane with maximum margin. The hyperplane equation dividing the two different groups of points can be represented as:

$$H: w^T(x) + b = 0$$

From a given point vector $\Phi(x0)$, distance d can be written as:

$$d_H(\phi(x_0)) = \frac{|w^T(\phi(x_0)) + b|}{||w||_2}$$

$$w^* = arg_w max\, [min_n\, d_H(\phi(x_n))]$$

On the training set, if the prediction is from the AD group in the hyperplane equation, $w^T(\Phi(x)) + b > 0$ will have a value greater than 0 (zero) and if the point is from non – AD group, then it would have negative values. As a result, SVM minimizes the aggregate distance between the maximum-margin hyperplane and the support vectors of AD and non-AD classes and given as:

SVM outperformed by giving an accuracy of 87% on the selected feature set.

$$w^* = arg_w max \left[min_n \frac{|w^T(\phi(x_n)) + b|}{||w||_2} \right] = arg_w max \left[min_n \frac{y_n|w^T(\phi(x_n)) + b|}{||w||_2} \right]$$

Table 2. Performance of ML Models considering All features Vs Selective Feature Set for classifying ADRG AND CG

Model	Accuracy (Overall Features)	Accuracy (Significant Feature)
SVM	82	**87**
LR	79	79
NB	74	79
DT	79	77
RF	74	74

Table 3. Evaluation Metrics of ML Models considering Selective Feature Set for classifying ADRG AND CG

ML Algo	Precision (ADRG/CG)	Recall (ADRG/CG)	F-Measure (ADRG/CG)	Accuracy
SVM	0.91/0.86	0.71/0.96	0.73/0.83	**87**
LR	0.71/0.84	0.71/0.84	0.71/0.84	79
NB	0.80/0.79	0.57/0.92	0.67/0.85	79
DT	0.62/0.83	0.71/0.76	0.67/0.79	74
RF	0.69/0.87	0.79/0.80	0.73/0.83	79

A comparison of the performance of SVM with other classifiers, viz Logistic Regression (LR), Random Forest (RF), Naïve Bayes (NB) and Decision Tree (DT) was done. Tables 2 and 3 exhibit the performance of all four classifiers on all feature set and particular feature set, respectively. Metrics such as Precision, Recall, F-measure and Accuracy were used to evaluate these models as this is a classification problem. All the experiments were performed using Python scikit-learn [27]. The results of all classifiers can be seen in Tables 2 and 3.

3.4. Evaluation Metrics

The performance of the classifiers was evaluated using Precision, Recall, F-measure and Accuracy. Here Recall is the proportion of the count of correct ADRG predictions to the total actual ADRG occurrences and Precision is the proportion of the count of correct ADRG predictions to the total returned ADRG occurrences. F-measure is the harmonic mean of Precision and Recall. The performance of these classifiers is also studied with the selective feature set. Table 2 shows the performance of ML models by considering 5-features having pause and hesitation metrics. It was found that SVM provided a precision of 0.91 and recall of 0.86 respectively on disease group whereas LR had a good recall after SVM.

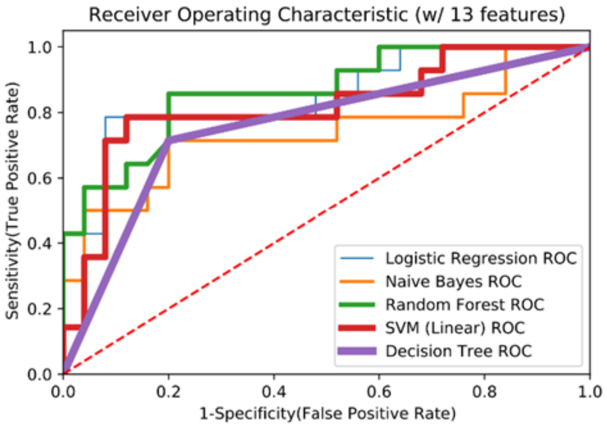

Fig 4. ROC Curve for all Features

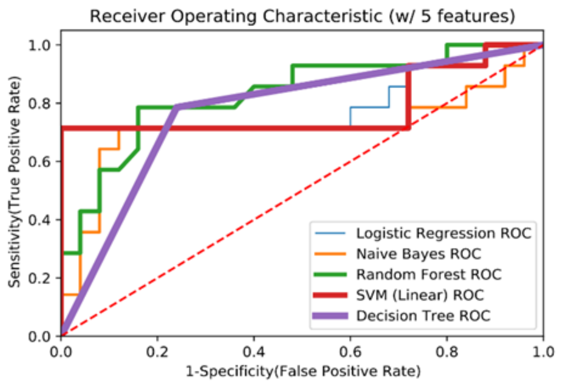

Fig 5. ROC Curve for Significant 5-Features Set

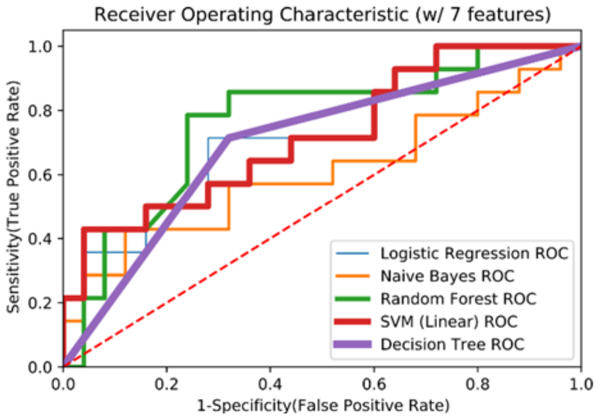

Fig 6. ROC Curve for Significant 7-Features Set

4. Results and Discussion

After comparing all the models, it is evident that SVM outperformed over all the classifiers for both the feature set. In this experiment, a comparison of the performance of 5 classifier algorithms was done using all features and significant features set obtained from feature selection. Our method of extracting the speech metrics directly from audio recordings could classify the groups with an accuracy of 0.87. The highest Accuracy of 0.87 was attained using the SVM algorithm with 0.91 precision in the disease group. The next highest Accuracy of 0.79 was attained by LR. Though the performance of RF was average on all feature sets, they gave a good recall of 0.79 for the disease ADRG group, which is better than the Recall from the other two classifiers. Figures 4, 5 and 6 show the ROC (Receiver Operating Characteristic) curve plotting for all classifiers on the 13, 5 and 7 features set.

Because of the transcript's availability with the Dementia Bank dataset, existing work done [12,17,23,30] considers only linguistic features. Our experiment mainly focuses on studying temporal features extracted directly from recordings of the Dementia Bank Dataset. The result showed an improved Precision of 16% compared to the earlier work [12] and improved Accuracy of 7% & 28% compared to works [17] and [30], respectively, for the same dataset.

As in the case of Probable AD and CG, the class distribution was well balanced and found that the pause can be one of the indicators to distinguish the probable AD group from the control group, which correlates with the study [7,8,24]. In our study, pause metrics were found to be significant indicators in the control Versus Probable AD group correlating with advanced stage AD symptoms. Some studies have confirmed that speech analysis can also aid in Mild AD and MCI diagnosis [30-32].

5. Conclusion & Future direction

The results of Machine Learning experiments and statistical evaluations suggest that learning temporal speech features from the narratives can aid the determination of Alzheimer's disease and related Dementia. This study was mainly based on speech temporal and acoustic features, which are different from the earlier work of syntactic and lexical features using the same dataset. This work does not claim to be a robust system for classification but a potential experimental work in temporal features. This dataset and experiments can also be extended to observe specific linguistic and temporal features contributing to disease progression. The Accuracy of the existing model (0.87) may improve by increasing the dataset with a balanced number of participants in each class. Also, using current technology for high-quality recording may provide better values for metrics such as pause and hesitation.

5.1. Acknowledgment

Special thanks to University of Pittsburgh for providing unhindered access to the Dementia Databank.

References

[1] 2020 Alzheimer's disease facts and figures. Alzheimers Dement. 2020 Mar 10. doi: 10.1002/alz.12068. Epub ahead of print. PMID: 32157811.

[2] K. S. Santacruz and D. Swagerty, "Early Diagnosis of Dementia," *Amer. Family Phys.*, vol. 63, no. 4, pp.703–713, Feb.15, 2001.

[3] *Alzheimer's Association International Conference (AAIC) 2021* – "Covid-19 Associated with Long-term Cognitive Dysfunction, Acceleration of Alzheimer's Symptoms," [Online]. Available: https://www.alz.org/

[4] James E. Galvin, "Using Informant and Performance Screening Methods to Detect Mild Cognitive Impairment and Dementia," *Current Geriatrics Reports*, Springer Science and Business Media, LLC, part of Springer Nature, vol.7, pp. 19–25, Jan.26, 2018.

[5] M. F. Folstein, S. E. Folstein, and P. R. McHugh, "Mini mental state: A practical method for grading the Cognitive State of patients for the clinician.," *J. Psychiatric Res.*, vol. 12, no. 3, pp.189-198, 1975.

[6] K. L. de Ipina, J.-B. Alonso, C. M. Travieso, J. Sol-Casals, H. Egi- ˜ raun, M. Faundez-Zanuy, A. Ezeiza, N. Barroso, M. Ecay-Torres, P. Martinez-Lage, and U. M. de Lizardui, "On the selection of non-invasive methods based on Speech Analysis oriented to Automatic Alzheimer disease diagnosis," *Sensors*., vol.13, no.5, pp.6730–6745, May 23, 2013.

[7] Ildiko Hoffmann, Dezso Nemeth, Cristina D. Dye, Magdolna Pakaski, Tamas Irinyi, Janos Kalman, "Temporal parameters of Spontaneous Speech in Alzheimer's Disease," *International Journal of Speech-Language Pathology*., vol.12, no.1, pp.29–34, Nov. 4, 2010.

[8] B. Roark, M. Mitchell, J.-P. Hosom, K. Hollingshead, and J. Kaye, "Spoken language derived measures for detecting Mild Cognitive Impairment," *IEEE Transactions on Audio, Speech, and Language Processing*, vol.19, no.7, pp.2081–2090, Sep. 1, 2011.

[9] C. Laske, H. R. Sohrabi, S. M. Frost, K. L. de-Ipina, P. Garrard, M. Buscema, J. Dauwels, S. R. Soekadar, S. Mueller, C. Linnemann, S. A. Bridenbaugh, Y. Kanagasingam, R. N. Martins, S. E. O'Bryant, "Innovative Diagnostic tools for Early Detection of Alzheimer's Disease," *Alzheimer's & Dementia*, pp.1-18,2014.

[10] M. Lehr, E.T. Prudhommeaux, I. Shafran, and B. Roark, "Fully Automated Neuropsychological Assessment for detecting Mild Cognitive Impairment," *Proceedings of Interspeech*, pp.1039-1042,2012.

[11] W. Jarrold, B. Peintner, D. Wilkins, D. Vergryi, C. Richey, M. L. Gorno-Tempini, and J. Ogar, "Aided Diagnosis of Dementia type through Computer-based Analysis of Spontaneous Speech," *Proceedings of ACL Workshop Computational Linguistics and Clinical Psychology*, Baltimore, Maryland, USA, pp. 27– 37, Jun.27,2014.

[12] S. O. Orimaye, J. S.M. Wong, and K. J. Golden, "Learning predictive linguistic features for Alzheimer's Disease and related Dementias using Verbal Utterances," *Workshop on Computational Linguistics and Clinical Psychology: From Linguistic Signal to Clinical Reality*, Baltimore, Maryland USA, pp. 78–87, Jun.27,2014.

[13] Ali Khodabakhsh, Serhan Kuscuoglu, Cenk Demiroglu, "Natural Language Features for Detection of Alzheimer's Disease in Conversational Speech," *IEEE-EMBS International Conference on Biomedical and Health Informatics (BHI)*, pp. 581-584, 2014.

[14] L. Toth, G. Gosztolya, V. Vincze, I. Hoffmann, G. Szatloczki, E. Biro, F. Zsura, M. Pakaski, J. Kalman, "Automatic Detection of Mild Cognitive Impairment from Spontaneous Speech using ASR," *Proceedings of Interspeech*, Dresden Germany, pp. 2694–2698, 2015.

[15] Nikhil Yadav, Christian Poellabauer, Louis Daudet, "Portable Neurological Disease Assessment Using Temporal Analysis of Speech," *ACM BCB'15*, Atlanta GA USA, pp. 77-85, Sep. 9-12, 2015.

[16] Alexandra Konig, Aharon Satt, Alexander Sorin, Ron

Hoory, Orith Toledo-Ronen, Alexandre Derreumaux, Valeria Manera, Frans Verhey, Pauline Aalten, Phillipe H. Robert, Renaud David, "Automatic Speech Analysis for the Assessment of Patients with predementia and Alzheimer's Disease, *Alzheimer's Dementia, Diagnosis, Assessment and Disease Monitoring*, vol. 1, no.1, pp.112-124, 2015.

[17] K. C. Fraser, J. A. Meltzer, F. Rudzicz, "Linguistic features identify Alzheimer's disease in narrative speech," *Journal of Alzheimer's Disease*. Vol. 49, no.2, pp.407-422, Aug.20,2015.

[18] E. Aramaki, S. Shikata, M. Miyabe, and A. Kinoshita, "Vocabulary size in speech may be an early indicator of cognitive impairment," *PloS ONE*, vol.11, no.5, pp.155-195, 2016.

[19] Gabor Gosztolya, Laszlo Toth, Tamas Grosz, Veronika Vincze, Ildiko Hoffmann, Greta Szatloczki, Magdolna Pakaski, Janos Kalman, "Detecting Mild Cognitive Impairment from Spontaneous Speech by Correlation-Based Phonetic Feature Selection," *INTERSPEECH*, San Francisco, USA, pp. 107-111, 2016.

[20] Ane Alberdi Aramendi, Asier Aztiria, Adrian Basarab, "On the early diagnosis of Alzheimer's Disease from multimodal signals: A survey," *Artificial Intelligence in Medicine*, vol. 71, pp.1-29, 2016.

[21] Hiroki Tanaka, Hiroyoshi Adachi, Norimichi Ukita, Takashi Kudo, Satoshi Nakamura, "Automatic Detection of Very Early Stage of Dementia through Multimodal Interaction with Computer Avatars," *ICMI'16*, Tokyo, Japan, pp.12–16,2016.

[22] H Tanaka, H Adachi, N Ukita, M Ikeda, H Kazui, T kudo, and S Nakamura, "Detecting Dementia Through Interactive Computer Avatars," *IEEE journal of Translational Engineering in Health and Medicine*, vol. 5,2017.

[23] Leandro B. dos Santos, Edilson A. Correa Jr, Osvaldo N. Oliveira Jr, Diego R. Amancio, Letıcia L. Mansur, Sandra M. Aluısio, "Enriching Complex Networks with Word Embeddings for Detecting Mild Cognitive Impairment from Speech Transcripts," arXiv:1704.08088v1 [cs.CL],2017.

[24] Laszlo Toth, Ildiko Hoffmann, Gabor Gosztolya, Veronika Vincze, Greta Szatlbczki, Zoltan Banreti, Magdolna Pakaski, Janos Kalman, "A Speech recognition-based Solution for the Automatic Detection of Mild Cognitive Impairment from Spontaneous Speech," *Current Alzheimer's Research*, Benthem Science Publisher, pp.130-138,2018.

[25] Becker JT, Boiler F, Lopez OL, Saxton J, McGonigle KL, "The Natural History of Alzheimer's Disease: Description of study cohort and accuracy of diagnosis," *Arch Neurol.* Vol.51, pp.585-594,1994.

[26] Goodglass H, Kaplan E, "The Boston Diagnostic Aphasia Examination," *Lea & Febinger*, Philadelphia,1983.

[27] Fabian Pedregosa, Gael Varoquaux, Alexandre Gramfort, Vincent Michel, Bertrand Thirion, Olivier Grisel, Mathieu Blondel, Peter Prettenhofer, Ron Weiss, Vincent Dubourg, Jake Vanderplas, Alexandre Passos, David Cournapeau, Matthieu Brucher, Matthieu Perrot, and Edouard Duchesnay, "Scikit-learn: Machine Learning in Python," *Journal of Machine Learning Research*, vol.12, pp.2825–2830,2011.

[28] Gerard Biau, "Analysis of a Random Forests Model," *Journal of Machine Learning Research*, vol. 13, pp. 1063-1095, 2012.

[29] J. F. Pitrelli, R. Bakis, E. M. Eide, R. Fernandez, W. Hamza and M. A. Picheny, "The IBM expressive text-to-speech synthesis system for American English," *IEEE Transactions on Audio, Speech, and Language Processing*, vol. 14, no. 4, pp. 1099-1108, July 2006.

[30] Saturnino Luz, "Longitudinal Monitoring and Detection of Alzheimer's Type Dementia from Spontaneous Speech Data,", *IEEE 30th International Symposium on Computer-Based Medical System*, 2017.

[31] Roark B, Hosom J-P, Mitchell M, Kaye J, "Automatically derived Spoken Language markers for detecting Mild Cognitive Impairment," *proceedings of the 2nd international conference on technology and aging (ICTA)*, pp.1–4,2007.

[32] Coulston R, Klabbers E, Villiers J, Hosom J-P, "Application of speech technology in a home-based assessment kiosk for early detection of Alzheimer's disease," *INTERSPEECH, 8th annual conference of the international speech communication association*, Antwerp, Belgium: Aug.27-31,2007.