APPROACH FOR SCREENING AND EARLY DIAGNOSIS OF ALZHEIMER'S DISEASE THROUGH DETECTION OF LINGUISTIC DEFICIENCIES AND OTHER BIOMARKERS

Maja Nikolova[1,2] , Zorica Nanovic[2], Sonja Gievska[1]

Faculty of Computer Science and Engineering[1] - MSc and Faculty of Medicine[2] - P.Grad.Dip. majanikolova89@gmail.com
Faculty of Medicine[2] - MD,PhD znanovic@hotmail.com
Faculty of Computer Science and Engineering[1] - PhD sonja.gievska@finki.ukim.mk
Ss. Cyril and Methodius University - Skopje, North Macedonia

Abstract. Introduction. Manually diagnosing neurodegenerative disorders like probable Alzheimer's disease and associated dementias has proven to be an arduous task. Their diagnosis is achieved (or performed) by using a combination of neuropsychological testing and particular clinical diagnostic criteria.The use of machine learning algorythms in developing an automated diagnostic model based on linguistic obtained from verbal interviews can become a crucial aid in the diagnostic process of these particular disorders. Material and methods.Based on a clinical dataset from the Dementia Bank, which includes personal and demographic information, the findings of physical and other medical examinations, and transcripts of audio recordings (I.E.interviews of each patient), we developed various machine learning models to meet that purpose..The collection of data records included 99 participant in each of the two groups, the group with probable AD and the control group. The models are based on distinct syntactic and lexical linguistic biomarkers to be able to discriminate the group of patients with probable Alzheimer's disease from a control group.. Results and Discussion.It was shown that patients with probable Alzheimer's disease have particulary increased their use of lexical components, while and dramatically decreasing their use of syntactic components in their speech when compared to the healthy control group. The use of machine learning algorithms to identify linguistic biomarkers in the verbal utterances of an older group of patients is an adequate and powerful tool, according to experimental and statistical evaluation, since they may help in the clinical diagnosis of probable Alzheimer's disease.

Keywords: Neurodegenerative disorders · Alzheimer's disease · Dementia Bank · Biomarkers

1 INTRODUCTION

The brain and spinal cord are components of the nervous system, which is assembled of neurons. Since neurons cannot be replaced or reproduced, they cannot

repair the brain damage. The term ”neurodegenerative disorder” is applied to a number of derangements that predominantly impact the neurons in the human brain. Examples of neurodegenerative diseases are Parkinson’s, Alzheimer’s and Huntington’s disease. Neurodegenerative disorders are incurable.They deplete the brain by causing nerve cells to gradually deteriorate and/or die, with subsequent movement difficulties (ataxias) or issues with mental function (dementias). Aproximately 70 percent of dementia cases are caused by Alzheimer’s disease. [1] [2].
Following are a few dementia subtypes: 1. Probable and possible Alzheimer’s disease (AD), 2. Vascular dementia, 3. Mixed dementia; 4. DLB; 5. Parkinson’s disease.

In most dementias, the neuropathology originates from the diffuse degeneration of cortical or subcortical structures and neural pathways, or chemical changes that affect neuronal function [1]: 1. Structural changes – include neurofibrillary tangles and neurotic plaques, often in correlation with AD, and the loss of neural pathways responsible for memory and new learning; 2. Chemical changes – representing cholinergic deficiency in subcortical structures (as in AD), and chemical imbalances associated with metabolic disorders.

Alzheimer’s disease is the most common neurodegenerative disorder, affecting millions of people worldwide[2].Clinical symptoms and signs include progressive dementia, irritability, confusion, and memory loss. Characteristic neuro pathology of AD, described for the first time by Alois Alzheimer in 1906, involves progressive neuronal degeneration and death due to aggregation and deposition of intra- and extracellular proteins, in the form of plaques and neurofibrillary tangles. Despite the fact that plaques and tangles consist of variety of well-characterized proteins (i.e. amyloids), the molecular mechanisms that lead to AD are still incompletely explored and understood. The time course and spatial distribution of amyloid deposition indicates thatAD progresses with certain regional specificity [2][3].

The use of neuropsychological tests to diagnose AD and associated dementia may be confined. The diagnostic process is based on the following: amnestic or non-amnestic presentation for detecting cognitive deficit through recounting a nearly observed event, a linguistic (lexical, syntactic, etc.) presentation, visuospatial description of an object and its semantic interpretation (based on cognitive processes necessary to identify, integrate, and analyze spatial and visual forms, details, structures, and spatial relationships in more than one dimension), and executive functions.[4]. In contrast to neuropsychological examinations, verbal utterances can provide an accurate indication of AD and associated dementia. [5]. According to the premise, neurodegenerative disease affects the nerve cells that regulate speech, cognition, and language functions, which is how we really derive the words and the sentence structure from a patient’s verbal output. [6][7]. A basis for ongoing research and the discovery of efficient syntactic strategies, according to certain writers [8],is provided by the syntactic procedures of syntactic processing in particular language disorders, such as adult aphasia. Similarly, it is important to emphasize the lexical-semantic components of the language that can be observed in the acquisition of verbal expressions in the younger population[9]. parallel to the growth of lexical capacity, syntactic processing becomes automated, changing the language as a result. Thus, it was determined that lexical and syntactic processes that control language and verbal expressions may change as a result of the consequences of a specific language disorder.[7][9].

When language and thinking are viewed as one system, it is obvious that both are functions of the central nervous system, and simultaneously involved in the brain working. Exchange of information takes place between language and perception, and memory and consciousness, in both directions. Language is involved in the reciprocal and recursive exchange of information in every element of thought. Language is closely related to thought, and it is normal to assume that language is part of the thinking process. The study of language is basically necessary for understanding how humans think. The more we study the language used by people, the more we will understand the structure of thought.
Comprehensive neuropsychological exams utilizing a battery of cognitive tests are applied in AD diagnostics (i.e. a set of questions and images). [10][11][12] [13].Decades of extensive research in the field of dementia have led to the identification of Mild Cognitive Impairment as "an intermediate condition of cognitive function, that lies between the alterations noticed in aging and those fitting the criteria for dementia and frequently AD" (MCI) [14][15]. A crucial QUESTION in the field is the identification of early, non-invasive biomarkers for the detection of pre-clinical or pre-symptomatic AD and other dementias, particularly in terms of research objectives, the design of preventive clinical trials, and the creation of population-based health care policies. Recent developments in computational linguistics have built upa potent tools which demonstrate that fully automated analyses of speech and language can accurately identify dementia patients and distinguish between different types of dementia, such as the early stage of Alzheimer's disease (AD) VS advanced stage of the disease, and non-fluent progressive aphasia from semantic dementia.[16][17].

The interest in automated spoken language analysis (NLP approaches) and the availability of several algorithms for speech analysis and classification have opened up new perspectives.[15]. Research advances in the field of discourse analysis, language modeling and text classification may be applicable to this area and may lead to progress. By using interpretable features, the statistical analysisand machine learning (ML) could be performed between groups, features, and dementia stage predictions.

In the context of previous one, the goals in conducting our research were: : 1. To investigate the performance of natural language processing (NLP) combinedwith reliable machine learning algorithms,in the analysis of spontaneous speech and changes in language performanceindicative of early cognitive decline; 2. To examine whether theseresultscan be used as linguistic biomarkersor early indices (i.e. ML prediction model) to identify preclinical or asymptomatic stages of AD and related dementia.

2 RELATED WORK

In 2011, Roark et al. carried out a study employing 21 linguistic elements of speech and syntactic measurements, including pauses, to demonstrate the effectiveness of using sophisticated syntactic features to categorize MCI, which is a precursor of AD. It was determined that seven language characteristics were statistically important for quick logical memory [18]. Our study used low-level syntactic and lexical indicators that are more representative than the linguistic pause in both participant groups to identify patients with probable AD. Additionally, our study makes extensive use of word n-grams to identify patients with probable AD by using more than 1000 practical and distinguishing characteristics. The significance of the lexical and syntactic characteristics of verbal narratives in individuals with probable AD is also examined.They exhibit phonemic paraphasia, word revision, semantic substitution, and obstacles in finding words as lexical characteristics. Coordinated sentences, subordinate clauses, and sentences scanty in words are among the syntactic aspects that are examined. For example, we distinguish semantic dementia where the use of nouns is increased, and progressive non-fluent aphasia where the use of verbs is increased. Bucks et al [19] also introduced a novel CNG approach, inspired from the authorship attribution which uses character N-grams to model consistencies in author style and used it in the Control vs. Dementia classification task with accuracy ranging from 80 percent to 94 percent. Wankerlet al [20] also used an N-gram based approach, but unlike the previous paper, used a word-N-gram model instead of a byte-level-N-gram model and used the perplexity of these models to analyze the data. The perplexity is used to evaluate how well an N-gram model fits the test data, with lower perplexity resulting in better test predictions. Using a binary classifier the study shows the following best results: 63.5percent, 59.1percent, 77.1percent for three different perplexity values as models. The problem with these kinds of models is the interpretability of the results. Although higher accuracy may be gained, the black box nature of these techniques where the reasoning behind the decision cannot be explained is what makes them mostly optimization problems and not methods of disease knowledge discovery.

3 MATERIAL AND METHODS

3.1 Dataset

This study uses a data from clinical trial conducted as a part of DementiaBank, precisely the English Pitt Corpus of the DementiaBanksegment of the TalkBank database[21]. The DementiaBank[22] dataset consists of examinations of adults aged 44 and older, with mandatory prerequisites as follows: an initial MMSE score of over 10, at least 7 years of education and no history of disorders of the nervous system. Participants are assigned to either Dementia or healthy Control group, based on a battery of neuropsychological tests and on their medical histories.

English Pitt Corpus data were gathered during a longitudinal study conducted by the University of Pittsburgh, School of Medicine, on patients suspected for Alzheimer's disease and related dementia. The database consists of transcripts of verbal interviews in English language, collected at different time intervals (ranging from one to two years between the visits). Participants were asked to verbally describe the contents a given image, i.e. the "Cookie Theft" picture description task from the Boston Diagnostic Aphasia Examinations(BDAE)[23], [24].Of note, the "Cookie Theft" picture description is a clinically relevant information for diagnosing and claiming linguistic deficiency in patients with AD and Aphasia [25]. Further, the participants in Dementia group completed four categories: the fluency (i.e., nominating words of a given category), the letter fluency (i.e., nominating words with a particular first letter), sentence construction, and story recall tasks. The picture description and fluency tasks were professionally transcribed and annotated with instances of filled pauses [21].

The Dementia Bank English Pitt Corpus dataset defines three groups of participants: Dementia, Control, and the group of Unknown diagnosis [Fig. 1]. At the moment of writing this paper, the Control group consisted of 99 participants, and 169 participants with probable or possible AD were included in the Dementia group. Since the core of our work is binary designation between the two groups, the group of the first 99 participants with probable AD has been selected for the study, and compared with equal number of healthy controls from the database.It can be assumed that the inclusion of symptomatic patients may increase the sensitivity of the model for correctly predicting the level of linguistic deficit, with subsequent follow up for confirming (or not) its connection with fully expressed AD or dementia [26].

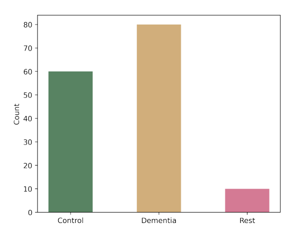

Fig. 1. The Dementia Bank English Pitt Corpus participants

3.2 Text Preprocessing

To get optimal results, one must perform various types of preprocessing on the initial transcripts. Depending on the feature in question, we performed different combinations of the following: special character removal, stop words removal, lemmatization and stemming. To accomplish this, we used the Python Natural Language Toolkit (NLTK). The resulting verbal statements of the patients were recorded as transcripts in CHAT (Codes for the Human Analysis of Transcripts) format [27].The CHAT transcripts are the result of computational techniques, designed in particular to expedite the automatic transcription of audio data for academic and research use. In our work, we took OUT sentences from the transcripts in CHAT files belonging to the patients, processed and analyzed them using Python's Natural Language Toolkit (NLTK), which revealed several characteristics that may be explored further. From the available demographic data, we only used the age in combination with extracted characteristics, in order to examine its significance for the disease [28]. Three fundamental linguistic elements are used by computational models: 1. n-gram model that identifies a distinctive string of words in the language of patients with probable AD and healthy control group. 2. Examination of the lexical data present in the vocabulary of patents with probable AD. 3. Using syntactic representation to understand variations in grammatical complexity in patients with probable AD and healthy control group.

The transcripts were used to extract a number of attributes. We started by removing every CHAT sign from the transcripts and cataloging them based on how frequently and where they appeared in each sentence. We stress that some CHAT symbols represent both explicit and implicit characteristics that characterize the patient's lexical competence. For instance, the CHAT symbol [//] in a particular position within a sentence suggests that the patient was going back and trying to repair a RHETORICAL FAILURE that happened before that position. The CHAT character [/] denotes instantaneous word repetition in a similar manner.[27], [29].

3.3 Feature Analysis

In the exploratory study we address the relevance and the predictive power of the following features: Clinical Dementia Rating (CDR)[30], Education[31], Mini-mental state examination (MMSE)[32], Age[33], Mattis Dementia Rating Scale (DRS) and MDRS-Second Edition (MDRS-2)[34], Education/Age[35], Blessed Dementia Scale (DS)[36], Hamilton Rating Scale for Depression (HRSD) and the so-called Modified Hamilton Rating Scale for Depression (MHRSD)[37], [38].

The complex syntactic processing is necessary for syntactic features. Three syntactic features were processed in this manner: 1. Coordinated sentences are those in which coordinating conjunctions are used to join clauses together. 2. Subordinate sentences, which are related to the major independent sentence and are tagged as part of the Part-Of-Speech (POS) process (CC); 3. Word-reduced sentences, which are presented as subordinate sentences without a conjunction but with nominal verb forms, are marked with the Part-Of-Speech (POS) tag (S). Assigned a tag as part of the Part-Of-Speech (POS) process (VBG and VBN).

The model's progress could be assessed by two lexical features: 1. Reactions (calculated as total number per patient). From the start of the verbal exchange to the subsequent verbal halt, each expression is calculated as a point or CHAT symbol that denotes a particular communication engagement. The utterances in a sentence can be one or more and can take the form of a word, phrase, or clause. Our theory is that the potential patient's language proficiency of a possible patient can be gauged by the total number of utterances in a single interaction; 2. Mean length of utterance (MLU) is used to measure the MLU is a measurement of how structural organisation of a given sentence. The MLU express the ratio between total count of words and the number of expressions. In our study, we investigate the value of MLU in identifying linguistic impairment in AD patients.

3.4 Prediction Models

In this paper we use a feature vector approach using metrics from different categories compiled from various studies. Different models have been developed from the features which confirm the hypothesis that the automated model can predict AD. A model with SVM (Support Vector Machines) has been developed, a kernel has been optimized and hyper-parameters were set with separate datasets. SVC (Support Value Classifier) was applied in cases with linearly separate data in two dimensions. A typical machine learning algorithm is one that tries to find a boundary dividing the data in a way that minimizes the error from the wrong classification. It differs from other classification algorithms in that it selects a decision boundary by maximizing the distance from the nearest data of each class.Dementia Bank comprises of multiple visitsof each patient. The transcripts from the second to the last visit, were used to generate a development set from which we may extract the hyper-parameters. This set consists of 40 random transcripts from each of the two included groups - the patients with probable AD and the control group. The datasets from the groups with probable AD and the control group that were used for the training and testing of our model are not included in the development set. Only the most recent visit of our two group's visits allowed us to obtain the training and testing sets. Our prediction models were evaluated based on the last visit. Gaussian NB- This algorithm has a probabilistic approach. It includes a priori and a posteriori calculation of the probability of the classes in the dataset, and the test data of the examined class.Feature analysis - A SHAP (SHapley Additive exPlanations) library was implemented. This feature extraction shows the structural organization of the sentence and the linguistic disorder, and builds the model.

4 RESULTS

4.1 Feature analysis

With the exception of the age feature, some of the features are not distributed equally, hence the presumption that the Dementia Bank is made to account for all participants' ages. Other characteristics of each patient are specific and may indicate other diseases over time. The research shows that the group with probable AD has less statistically significant syntactic traits than the control group. The AD group has certain difficulties in constructing a complex sentence, in contrast to the control group.It has been suggested that the use of reduced structures may be vital to the adequate measurement of linguistic abilities in patients with probable AD.

4.2 Correlations

Inter-feature correlation was calculated and is shown in Figure 2. We use this heat map to estimate the additional information that each of these features gives when added. A strong positive correlation is observed between Mattis and mms (0,96), htotal and hmtotal (0,86), Blessed and cdrfs (0,82) and cdrfs and mms-grp (0,83), due to the nature of these features and the way they are evaluated. Nevertheless,for most of the entities the heat map reveals low inter-feature correlation, meaning they all-together provide unique information with regards to predicting the diagnose and detecting the stage of the disease.

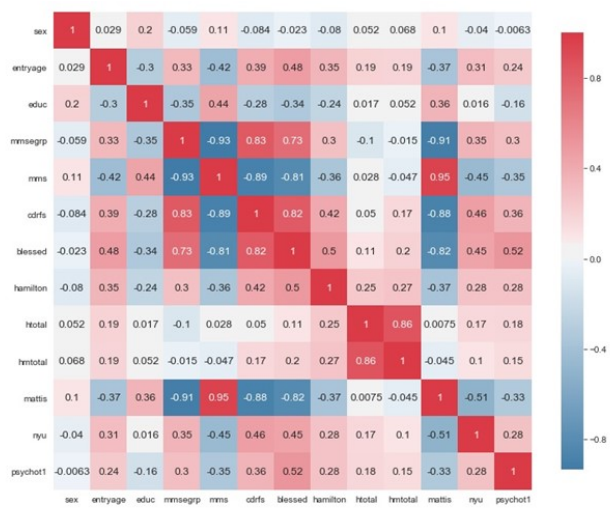

Fig. 2. Heat map of dependencies

Inter-feature correlation was calculated and is shown in Figure 3. A strong positive correlation is observed between coordinating sentences and responses

(0,49), and coordinating sentences and Reduced sentences (0,49). However, for most of these features the heat map shows low inter-feature correlation, meaning that they provide unique information with regards to predicting the linguistic disorder score.

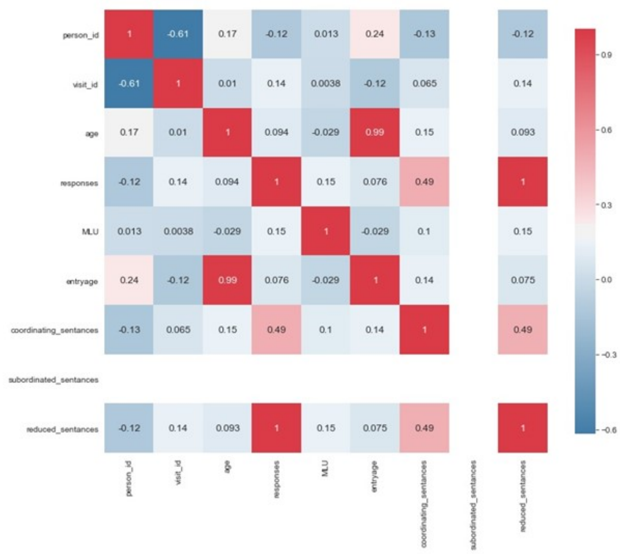

Fig. 3. Heat map of Lexical features and Syntactic features with patient's variables

However, statistically relevant lexical characteristics from the group with probable AD, with the exception of the MLU parameter, had a greater value than the control group. Due to existing pauses and syntactic faults, the group with probable AD had a higher number of utterances. Additionally, this group has talks that are typically longer and contain more phrases. Patients with probable AD might be differentiated from the control group by distinct and significant lexical characteristics. On average, the group of patients with probable AD is five years older than the control group. There is a significant 20 percent chance of reducing the sentences for description of a given picture by patients with probable AD in contrast to the healthy control group. Finally, the MLU coefficient is 35percent higher for prediction in the group with probable AD than it is in the control group. A SHAP library, which was implemented in terms of feature analysis, shows that the two most important features: Blessed (Figure 4) - which measures the degree of intellectual and personal deterioration, and MLU from NLP (Figure 5)- which shows the structural organization of the sentence and the linguistic disorder.

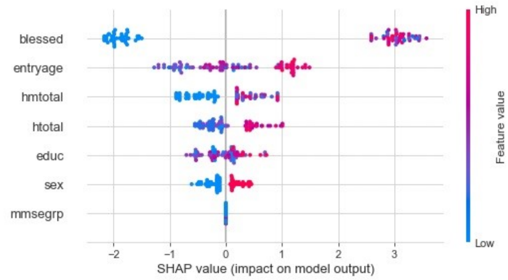

Fig. 4. Feature analysis in SHAP library

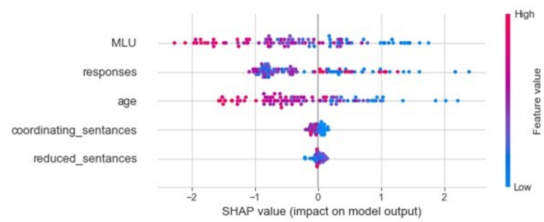

Fig. 5. Feature analysis in SHAP library

4.3 Classification

According to the Figure10, it becomes obvious that SVM not only reaches the decision limit, but also finds the most optimal decision boundaries. Such an optimal limit is one that has a maximum margin of the nearest points of all classes. The nearest deciding boundary points that maximize the distance between the boundary and the points are the support vectors. Here, the deciding boundary is the maximum margin classifier or the maximum margin hyperplane. The SVM classifier was evaluated using following metrics: Precision, Recall, F1-Score and Support (Table 1 and Figure 6). In order to improve prediction performance, it is important to choose the right prediction algorithm. In the Table 1 and Table 2, we compare the f1 score. From the results, we can conclude that as a diagnostic model in this case, the Gaussian NB model shows better performance then SVC(Figure 7 and Table 2).

Table 1. Confusion matrix - *SVC*

	Precision	Recall	f1-score	support
control	0.82	0.70	0.76	20
demented	0.87	0.93	0.90	42
micro svg	0.85	0.85	0.85	62
macro svg	0.85	0.81	0.83	62
weighted svg	0.85	0.85	0.85	62

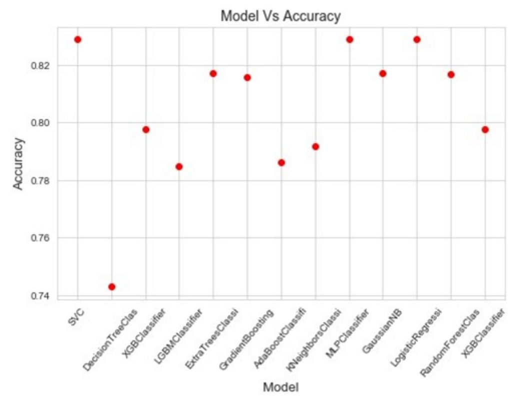

Fig. 6. Fig.10. Model accuracy

Table 2. Confusion matrix - *GaussianNB*

	Precision	Recall	f1-score	support
control	0.90	0.90	0.90	20
demented	0.95	0.95	0.95	42
micro svg	0.94	0.94	0.94	62
macro svg	0.93	0.93	0.93	62
weighted svg	0.94	0.94	0.94	62

4.4 Future Work

In the future we would like to expand upon our work. Firstly, we would like to increase the number of metrics in the feature vector by adding new more complex linguistic features. We would also like to explore and take advantage upon the complex links between the words labeled in the Pitt Corpus. Finally, prosodic features are shown to be effective in the detection of Alzheimer's Disease in conversational speech and should prove to give additional information in our task as well.

5 CONCLUSION

Our study uses a diagnostic model that is effectively generated and based on verbal statements to predict probable AD, in terms of differentiating between patients with probable AD and healthy control group.The potential to predict the phenotype of probable AD, which may have progressed beyond the stage of MCI, is our study's clinical benefit. The pathogenetic pathway of AD has a gap

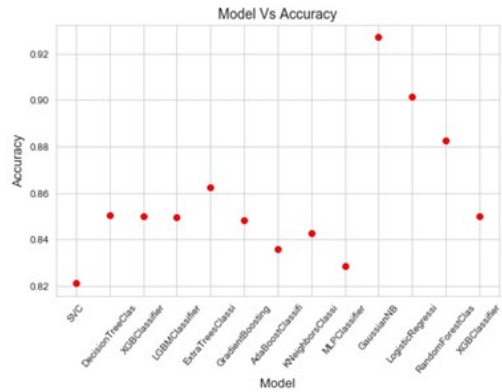

Fig. 7. Model accuracy

between the prodromal stage and the stage at which all of its symptoms will manifest, and this fact has to be taken into consideration. It is also possible to predict pathological processes in the brain that occur as a consequence of AD, before clinically visible symptoms begin. An effective technique for automating the diagnosis of diseases in the large population using only transcripts has been proposed. This is part of the effort to update an automated tele-diagnosis tool that can help disease screening among the large population, where manual neuropsychological examination may be limited by a number of factors. The findings in our study utilizing ML algorithms and statistical analysis indicate that using such learning algorithms as evaluation measurements of the tests carried out on each patient separately, over many years, can aid in the a effective diagnosis of AD.According to the results of our research, apart from the evaluation of the tests, the experiment that combines lexical and syntactic features gives a slightly lower, but still sufficiently high result. Both evaluate the data correctly. The predictive diagnostic model is able to register cognitive deficits. In addition, the predictive diagnostic model can assess the patient for the required degree of rehabilitation in terms of linguistic dysfunction. Longitudinal studies are recommended as a way to monitor the disease over time. At the moment, we are testing the multimodal approaches, that include consideration of other datasets, such as inclusion of MR imaging.We are strengthening our work by diversifying our processing techniques, applications, and other algorithms. Based on the related research in the field, we are extending the initial set of features, such as those related to different medical tests that are expected to yield statistically significant performance in the predictive model. Furthermore, we are considering alternative approaches that could help determine the degree of degenerative brain changes in the obtained post-mortem samples.

6 REFERENCES

[1] C. Ballard, S. Gauthier, A. Corbett, C. Brayne, D. Aarsland, and E. Jones, "Alzheimer's disease," en, *Lancet*, vol. 377, no. 9770, pp. 1019–1031, Mar. 2011. DOI: 101931.doi:10.1016/S0140-6736(10)61349-9..

[2] Alzheimer's Association Report, "Alzheimer's disease facts and figures," *Alzheimer's &amp Dementia*, vol. 16, no. 3, pp. 391–460, Mar. 2020. DOI: 10.1002/alz.12068.

[3] R. Armstrong, "β-amyloid plaques: Stages in life history or independent origin?" *Dementia and Geriatric Cognitive Disorders*, vol. 9, no. 4, pp. 227–238, 1998. DOI: 10.1159/000017051.

[4] G. M. McKhann, D. S. Knopman, H. Chertkow, *et al.*, "The diagnosis of dementia due to alzheimer's disease: Recommendations from the national institute on aging-alzheimer's association workgroups on diagnostic guidelines for alzheimer's disease," *Alzheimer's &amp Dementia*, vol. 7, no. 3, pp. 263–269, Apr. 2011. DOI: 10.1016/j.jalz.2011.03.005.

[5] A. Tillas, "Language as grist to the mill of cognition," *Cognitive Processing*, vol. 16, no. 3, pp. 219–243, May 2015. DOI: 10.1007/s10339-015-0656-2.

[6] J. Reilly, A. D. Rodriguez, M. Lamy, and J. Neils-Strunjas, "Cognition, language, and clinical pathological features of non-alzheimer's dementias: An overview," *Journal of Communication Disorders*, vol. 43, no. 5, pp. 438–452, Sep. 2010. DOI: 10.1016/j.jcomdis.2010.04.011.

[7] M. Verma and R. J. Howard, "Semantic memory and language dysfunction in early alzheimer's disease: A review," *International Journal of Geriatric Psychiatry*, vol. 27, no. 12, pp. 1209–1217, Feb. 2012. DOI: 10.1002/gps.3766.

[8] M. J. Ball, M. R. Perkins, N. Müller, and S. Howard, *The handbook of clinical linguistics*. Wiley Online Library, vol. 56.

[9] J. L. Locke, "A theory of neurolinguistic development," *Brain and Language*, vol. 58, no. 2, pp. 265–326, Jun. 1997. DOI: 10.1006/brln.1997.1791.

[10] J. Williams, A. Weakley, D. Cook, and M. Schmitter-Edgecombe, "Machine learning techniques for diagnostic differentiation of mild cognitive impairment and dementia," *AAAI Workshop - Technical Report*, pp. 71–76, Jan. 2013.

[11] P. Johnson, L. Vandewater, W. Wilson, *et al.*, "Genetic algorithm with logistic regression for prediction of progression to alzheimer's disease," *BMC Bioinformatics*, vol. 15, no. S16, Dec. 2014. DOI: 10.1186/1471-2105-15-s16-s11.

[12] M. Mitolo, S. Gardini, P. Caffarra, L. Ronconi, A. Venneri, and F. Pazzaglia, "Relationship between spatial ability, visuospatial working memory and self-assessed spatial orientation ability: A study in older adults," *Cognitive Processing*, vol. 16, no. 2, pp. 165–176, Mar. 2015. DOI: 10.1007/s10339-015-0647-3.

[13] M. Goryawala, Q. Zhou, W. Barker, D. A. Loewenstein, R. Duara, and M. Adjouadi, "Inclusion of neuropsychological scores in atrophy models

improves diagnostic classification of alzheimer's disease and mild cognitive impairment," *Computational Intelligence and Neuroscience*, vol. 2015, pp. 1–14, 2015. DOI: 10.1155/2015/865265.

[14] R. C. Petersen, "Mild cognitive impairment," *New England Journal of Medicine*, vol. 364, no. 23, pp. 2227–2234, Jun. 2011. DOI: 10.1056/nejmcp0910237.

[15] D. Beltrami, G. Gagliardi, R. R. Favretti, E. Ghidoni, F. Tamburini, and L. Calzà, "Speech analysis by natural language processing techniques: A possible tool for very early detection of cognitive decline?" *Frontiers in Aging Neuroscience*, vol. 10, Nov. 2018. DOI: 10.3389/fnagi.2018.00369.

[16] A. Shiromaru-Sugimoto, H. Murakami, A. Futamura, *et al.*, "The subjective perception of past, present, and future time in patients with alzheimer's disease: A qualitative study.," *Neuropsychiatric Disease and Treatment*, vol. Volume 14, pp. 3185–3192, Nov. 2018. DOI: 10.2147/ndt.s186081.

[17] T. Sugimoto, T. Sakurai, R. Ono, *et al.*, "Epidemiological and clinical significance of cognitive frailty: A mini review," *Ageing Research Reviews*, vol. 44, pp. 1–7, Jul. 2018. DOI: 10.1016/j.arr.2018.03.002.

[18] 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. 2011. DOI: 10.1109/tasl.2011.2112351.

[19] R. S. Bucks, S. Singh, J. M. Cuerden, and G. K. Wilcock, "Analysis of spontaneous, conversational speech in dementia of alzheimer type: Evaluation of an objective technique for analysing lexical performance," *Aphasiology*, vol. 14, no. 1, pp. 71–91, Jan. 2000. DOI: 10.1080/026870300401603.

[20] S. Wankerl, E. Nöth, and S. Evert, "An n-gram based approach to the automatic diagnosis of alzheimer's disease from spoken language," Aug. 2017. DOI: 10.21437/interspeech.2017-1572.

[21] F. Boller, J. Becker, and P. Corpus. "Dementiabank english pitt corpus." (), [Online]. Available: https://dementia.talkbank.org/access/English/Pitt.html. (accessed: 31.01.2022).

[22] "Dementiabank." ((1994)), [Online]. Available: https://dementia.talkbank.org/access/.

[23] E. Kaplan, H. Goodglass, S. Weintraub, *et al.*, *Boston naming test*. Pro-ed Austin, 2001.

[24] G. H. Kaplan EF Weintraub S, *Boston naming test*, 2nd ed. Lea Febiger, Philadelphia, 1983.

[25] J. D. Rohrer, M. N. Rossor, and J. D. Warren, "Syndromes of nonfluent primary progressive aphasia: A clinical and neurolinguistic analysis," *Neurology*, vol. 75, no. 7, pp. 603–610, Aug. 2010. DOI: 10.1212/wnl.0b013e3181ed9c6b.

[26] M. C. Evans, J. Barnes, C. Nielsen, L. G. Kim, S. L. Clegg, *et al.*, "Volume changes in alzheimer's disease and mild cognitive impairment: Cognitive associations," *European Radiology*, vol. 20, no. 3, pp. 674–682, Sep. 2009. DOI: 10.1007/s00330-009-1581-5.

[27] B. MacWhinney, *The Childes Project.* Psychology Press, Jan. 2000, vol. 2. DOI: 10.4324/9781315805641.

[28] S. Bird, E. Klein, and E. Loper, *Natural language processing with Python*, en. Sebastopol, CA: O'Reilly Media, Jul. 2009.

[29] D. Klein and C. D. Manning, "Accurate unlexicalized parsing," in *Proceedings of the 41st Annual Meeting of the Association for Computational Linguistics*, Sapporo, Japan: Association for Computational Linguistics, Jul. 2003, pp. 423–430. DOI: 10.3115/1075096.1075150.

[30] J. C. Morris, "Clinical dementia rating: A reliable and valid diagnostic and staging measure for dementia of the alzheimer type," *International Psychogeriatrics*, vol. 9, no. S1, pp. 173–176, Dec. 1997. DOI: 10.1017/s1041610297004870.

[31] K. S. Shpanskaya, K. R. Choudhury, C. Hostage, K. R. Murphy, J. R. Petrella, and P. M. Doraiswamy, "Educational attainment and hippocampal atrophy in the alzheimer's disease neuroimaging initiative cohort," *Journal of Neuroradiology*, vol. 41, no. 5, pp. 350–357, Dec. 2014. DOI: 10.1016/j.neurad.2013.11.004.

[32] M. F. Folstein, S. E. Folstein, and P. R. McHugh, ""mini-mental state"," *Journal of Psychiatric Research*, vol. 12, no. 3, pp. 189–198, Nov. 1975. DOI: 10.1016/0022-3956(75)90026-6.

[33] D. Wolf, F. U. Fischer, and A. Fellgiebel, "A methodological approach to studying resilience mechanisms: Demonstration of utility in age and alzheimer's disease-related brain pathology," *Brain Imaging and Behavior*, vol. 13, no. 1, pp. 162–171, May 2018. DOI: 10.1007/s11682-018-9870-8.

[34] E. Matteau, N. Dupré, M. Langlois, *et al.*, "Mattis dementia rating scale 2," *American Journal of Alzheimer's Disease &amp Other Dementiasr*, vol. 26, no. 5, pp. 389–398, Jun. 2011. DOI: 10.1177/1533317511412046.

[35] C. M. Bauer, "Differentiating between normal aging, mild cognitive impairment, and alzheimer's disease with FDG-PET: Effects of normalization region and partial volume correction method," *Journal of Alzheimer's Disease &amp Parkinsonism*, vol. 03, no. 01, 2013. DOI: 10.4172/2161-0460.1000113.

[36] A. J. Keller, E. M. S. Sherman, and E. Strauss, "Blessed dementia scale," in *Encyclopedia of Clinical Neuropsychology*, Springer International Publishing, 2018, pp. 1–4. DOI: 10.1007/978-3-319-56782-2_531-2.

[37] I. W. Miller, S. Bishop, W. H. Norman, and H. Maddever, "The modified hamilton rating scale for depression: Reliability and validity," *Psychiatry Research*, vol. 14, no. 2, pp. 131–142, Feb. 1985. DOI: 10.1016/0165-1781(85)90057-5.

[38] E. E. Sundermann, P. M. Maki, L. H. Rubin, R. B. Lipton, S. Landau, and A. B. and, "Female advantage in verbal memory," *Neurology*, vol. 87, no. 18, pp. 1916–1924, Oct. 2016.