Journal of Theoretical and Applied Information Technology
15th March 2021. Vol.99. No 5

ISSN: 1992-8645 www.jatit.org E-ISSN: 1817-3195

DETECTING DEPRESSION IN ALZHEIMER'S DISEASE AND MCI BY SPEECH ANALYSIS

[1]BASHAR ABDALLAH-QASAIMEH, [1]SYLVIE RATTÉ
[1]Department of Software and IT Engineering, École de technology supérieure, Canada
E-mail: bashar.abdallah.1@ens.etsmtl.ca, sylvie.ratte@etsmtl.ca

ABSTRACT

It is estimated that 30% of Alzheimer's patients suffer from depression. Since this condition can lead to further cognitive decline and suffering, its detection is essential to alleviate MCI (mild cognitive impairment) or AD symptoms in patients. This paper presents a machine learning method aimed at identifying MCI and Alzheimer's disease (AD) patients suffering from depression, using different features extracted from their speech. 276 participants (mean age 70.9 years) are selected from DementiaBank's Pitt Corpus for this research. The interviewer's voice and the silences are removed from the audio records as a preprocessing task. Several audio features are extracted from the patient's speech to achieve this task. For instance, MFCC's, Spectral Centroid, Spectral Roll-Off Point, and others. We trained and compared three families of classifiers (SVM, Random Tree, and Random Forest) through two experiments, one using Spectral feature variants and MFCC, and the other using only MFCC features. A third experiment is conducted for comparison with the literature review. In all cases, we used a bootstrapping method to solve the sampling bias, i.e., 70% of the patients not suffering from depression. From the results, the MFCC feature set was more appropriate for tree-based classifiers than SVM, in which the Random Tree classifier reported the highest classification performance (91.3%). Meanwhile, the other feature sets were more appropriate for SVM than the tree-based classifiers, where SVM reported an 89.1 classification accuracy, with 91.1% recall.

Keywords: *Alzheimer's; MCI; depression; automatic detection; speech analysis.*

1. INTRODUCTION

1.1 Alzheimer's disease

While a certain amount of cognitive decline is part of a healthy aging process, many older seniors worldwide suffer from a sharp cognitive decline caused by Alzheimer's disease (AD) and other types of dementia. AD, in particular, poses many challenges both at the individual and societal levels. First, it has a direct impact on patients, as well as their loved ones, because as the disease progresses, a greater and greater burden is put upon the circle of relatives, friends, and acquaintances. Second, the high number of patients exerts economic and social pressures on society since related medical care becomes more expensive. Indeed, in 2017, the estimated total costs associated with all individuals with AD or other dementias was $259 billion [1]. Furthermore, the percentage of total deaths due to Alzheimer's disease is greater than those for both breast cancer and prostate cancer combined [1].

1 out of 3 seniors dies of AD or other dementia types [1]. Alzheimer's disease usually progresses gradually, and in three general stages. At the Mild (early) stage, patients may still be able to accomplish familiar tasks, and might notice mild cognitive changes such as difficulty remembering the names of new people [1].

The disease then reaches the Moderate stage, which can last many years. During this stage, symptoms evolve, and worsen; the patient feels moody, is more forgetful, and has more and more difficulty accomplishing familiar tasks. Overall, the patient is losing the ability to perform regular daily tasks, including those involving several mental skills such as planning and keeping track of recent events. The disease gradually progresses until it reaches a severe stage, in which patients lose the ability to speak and respond to the environment.

Diagnosing AD is notoriously tricky, with the assessment failure percentage reaching as high as 50% [2]. One of the challenges is identifying the early signs of AD among patients that are believed to be in the Mild Cognitive Impairment (MCI) stage. According to [3], several studies suggest that the percentage of MCI among people 65 and older ranges between 3 and 19%. With MCI, changes are not severe enough to affect daily life functions. However, people diagnosed with MCI are more likely to develop Alzheimer's disease than people without MCI, and the rate at which MCI will progress to Alzheimer's disease lies between 10% and 15% annually [4].

1.2 Depression

While "feeling sad" is a typical transitional emotion, depression is a severe medical illness since it profoundly affects how patients feel, the way they think, and how they act. This illness can lead patients to lose their motivation to carry out normally enjoyable activities or to lose the ability to function at work or home. One out of fifteen adults can be affected by depression in any given year, and one in six people will experience depression at some time in their life [5].

A patient can be diagnosed with depression after two weeks following the appearance of symptoms. Symptoms can include the following: changes in appetite, thoughts of death or suicide, slowed movements and speech, and difficulty thinking, concentrating, or making decisions [5]. Moreover, in cases of major depression, feelings of worthlessness and self-loathing are common [5]. Fortunately, depression is a treatable disease, and between 80 and 90% of patients respond to treatment [5].

According to [6], around 30-50% of AD patients suffer from depression. This further complicates the task of diagnosing AD since both illnesses share some symptoms. For example, both induce poor concentration, impaired attention [7], apathy [6], and changes in eating and sleeping patterns [8]. Suffering from both MCI and depression can augment the risk of developing Alzheimer's disease by a factor of two (or more) when compared to non-depressed MCI patients [9]. Consequently, cognitive decline is more significant among Alzheimer's patients who are suffering from depression than for patients without depression [10] [11] [12]. Detecting and treating depression in AD patients helps diagnose AD more accurately, and can further improve patients' prognosis.

1.3 Related work

Many machine learning models have been proposed to distinguish depressed from non-depressed patients [13] [14] [15], [16] on the one hand, and AD from non-AD [17], [18], [19] on the other. Most models, particularly those for depression, are based on the use of several audio features, such as speech segments and fundamental frequency. For AD detection, textual features, such as lexical richness, are mostly preferred.

In [14], the authors used a non-linear analysis of EEG signals as a feature to discriminate 45 depressive patients from 45 control subjects; none of the participants was suffering from AD. The classification model had a 90% classification accuracy using the Regression classifier. [15] Used several audio features and some textual features to detect depression in non-AD subjects. Their best SVM classifier achieved a 74% accuracy. The authors in [16] used hybrid classifiers to detect depression from the spontaneous speech of 60 participants (again non-AD). The hybrid classifier consisted of SVM and GMM; this hybrid classifier were fed by several audio features such as MFCC, Pitch, Intensity, etc. The model had a 91.6 % classification accuracy. Fraser et al.

[17] achieved an 81% classification accuracy in distinguishing AD patients from healthy people using both audio and linguistic features. A logistic regression classifier was used to perform this task with several linguistic features, such as parts of speech. In a recent work on depression detection in AD patients, the same authors [13] used both acoustic and linguistic features to study the impact of depression on the detection of Alzheimer's disease, by classifying participants into depressive and non-depressive groups. In the experiment, the best model achieved a 65.8% classification accuracy using the Logistic Regression classifier. As can be seen, both detection tasks (AD vs non-AD, depressive vs non-depressive) are tricky, and highly dependent on the context of the data used.

Our research aimed to assist the medical community in detecting depression in Alzheimer's patients and in patients with MCI (the same task of [13]). Therefore, we used speech data recorded during the Picture Description Task (same data used in [13]), a task widely used to detect signs of AD. Our approach was also based on machine learning techniques, but involves three steps: the cleaning of the audio feed, the feature extraction process, and the classification phase.

ISSN: 1992-8645 www.jatit.org E-ISSN: 1817-3195

Since it is generally agreed that many signs of depression are hidden in the speech signal, while AD signs are more present in the textual content, we added a cleaning phase to reduce the audio to the patients' speech moments. The classification phase separates the patients into two sets, depressed and non-depressed. Our exploration aims to answer two questions that are limited to AD and MCI patients: 1) which classification technique performs best in distinguishing depressed patients from non-depressed patients? 2) Which features subset is the most appropriate for this task?

2. METHODS

2.1 Dataset

This section describes the dataset used in this paper. This dataset is part of the DementiaBank's (DB) Pitt Corpus, [20] a shared database consisting of multimedia data that contains 309 participants suffering from various types of dementia, including AD and MCI. It also contains 218 Control participants. 518 for Alzheimer's patients, MCI, and Control participants.

To answer our two questions, we extracted 276 MCI or AD patients (165 females and 194 males) from the DB, and excluded other types of dementia. The average ages of our participants was 70.9. Within that subset, 194 of the participants has AD, while the rest were suffering from both depression and AD. All the participants performed the Cookie Theft picture description task, a test that is part of the Boston Diagnostic Aphasia Examination [21].

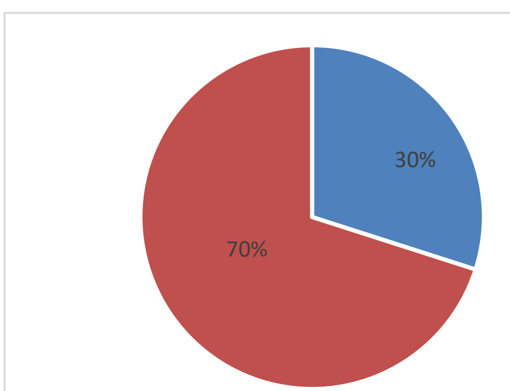

Figure 1: The dataset

The Hamilton depression rating scale (HAM-D) is used in DB to rate the depression level in participants [22]. If the total score for a participant is between 0 and 7, the participant is non-depressed; a score equal to or greater than 8 is an indicator that the participant is depressed [23]. For our AD patients, their HAM-D ranged between 0 (normal) and 20 (severe depression) with a 5.9 average for the non-depressed patients (70%, 194 patients), and an average of 10.6 for our depressed patients (30%, 82 patients) – see figure 1.

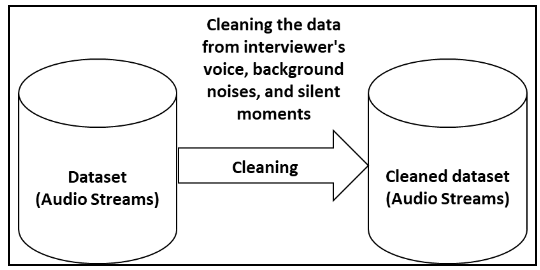

Figure 2: Cleaning the dataset

Since many signs of depression are hidden in the speech signal, we preprocessed the audio stream of each patient (see figure 2). First, we removed the parts where the interviewer was speaking. Second, we removed pauses – the intervals during which patients stop talking – irrespective of whether these pauses contained silence or audio noises. We thus concentrated our analysis on the patients' voices.

2.2 Feature extraction

This section describes the process of extracting audio features and the features used. The following figure (figure 3), describes the process of feature extraction.

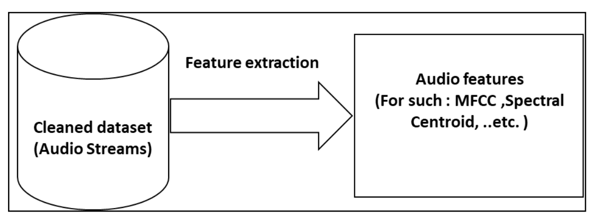

Figure 3: Feature Extraction

ISSN: 1992-8645 www.jatit.org E-ISSN: 1817-3195

We extracted two groups of features using the jAudio framework in both cases [24]. Since Standard low-level features (SLL) are more likely to achieve high-performance rates in classifying the speech audio [25], we hypothesize that they could perform well in detecting depression. Our first group, group 1, contains the following features:

- Spectral Centroid: "a short-time Fourier transform performed frame-by-frame along the time axis [26]".
- Spectral Roll-Off Point: " the frequency below which 85% of the energy in the spectrum is located. It is often used as an indicator of the skewness of the frequencies present in a window [25] ".
- Root Mean Square: " the amplitude of a window [25] ".
- Zero Crossings: " the number of time-domain zero crossings within the processing window [26] ".
- Mel-Scale Frequency Cepstral Coefficient (MFCC): " describes a spectrum window [25] ".
- Method of Moments: " This feature consists of the first five statistical moments of the spectrograph: the area (zeroth-order), the mean (first-order), the Power Spectrum Density (second-order), the Spectral Skew (third-order), and the Spectral Kurtosis (fourth-order) [25] ".

We did not extract the entire SLL feature set because some of the features cannot be extracted using jAudio, in addition to which we wanted to keep the feature extraction process as simple as possible. The method of moments feature is extracted to study the impact of its addition to our feature set; moreover, it is a step allowing the discovery of some other features that can improve the performance of the classification.

Since MFCC is one of the most used features in speech recognition [27], our second group, group 2, is composed solely of features derived from it; it contains the mean, the standard deviation, the kurtosis, and the skewness.

2.3 Classification

This section describes the machine learning tools and the classification techniques used in this paper (see figure 4). To quickly test a variety of classifiers, we used the Waikato Environment for Knowledge Analysis [28]. After some preliminary testing, we choose to compare three classifiers: Support vector machine (SVM) [29], Random Forest (RF) [30], and Random Tree (RT) [31].

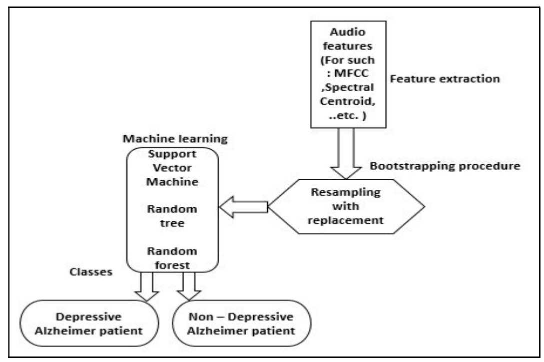

Figure 4: The classification process

To solve the imbalance of the data (and therefore, the sampling bias) – 70% of the dataset is non-depressive, while 30% is depressive (see figure 1) – we applied a Bootstrapping procedure. Contrary to other methods, bootstrapping is "a nonparametric resampling procedure that does not assume a normal distribution. It consists of repeatedly sampling the data and estimating the effect of each resampled dataset [32]". For our evaluation, we used 10-fold cross-validation. Our 276 patients were classified as Depressed (194 with Hamilton score equal to or greater than 8) and Non-depressed (Hamilton score between 0 and 7) classes [23].

3. RESULTS

To study the effects of the audio feature extractions on the machine learning classifiers, we constructed two (3) experiments. Experiment (A) used subsets of the features in group 1, while group 2 used in Experiment B and C (See figure 5).

ISSN: 1992-8645 www.jatit.org E-ISSN: 1817-3195

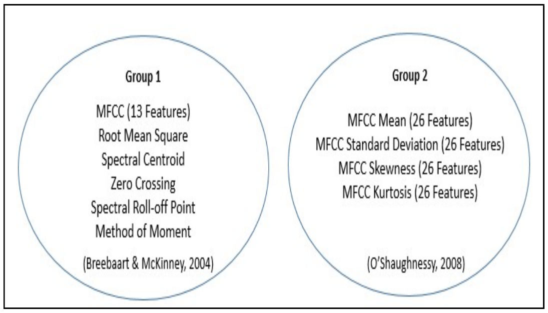

Figure 5: Feature groups

3.1 Experiment A

This experiment was executed on two (2) subsets, which both used the features of group 1.

3.1.1 Subset 1

This subset contained the following features: Spectral Centroid, Root Mean Square, Zero Crossings, and MFCC. The results are shown in Figure 6.

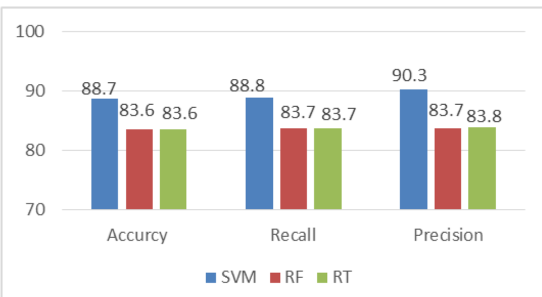

Figure 6: Classification Results Subset 1

From Figure 6, the SVM classifier reported the highest classification performance rates as follows: accuracy (88.7%), recall (88.8%), and precision (90.3%). Therefore, SVM is ranked first, ahead of the other two (2) classifiers (i.e., Random Tree and Random Forest).

3.1.2 Subset 2

This feature set contained the same features of the first subset to which we added the Spectral Roll-Off point and Method of the moment. Figure 7 shows the results.

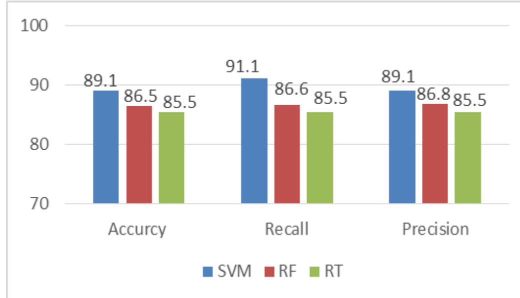

Figure 7 Classification Results subset 2

From Figure 7, the SVM classifier once again reported the highest classification performance rates as follows: accuracy (89.1%), recall (91.1%), and precision (89.1%). Therefore, SVM is ranked first, ahead of the other two (2) classifiers. We can observe a general improvement in the performance of all classifiers, and especially with the Random Tree and Random Forest.

3.2 Experiment B

This experiment was executed on three (3) subsets of features. Each subset consisted of different combinations of (26 MFCC) features (Average, Standard deviation, Kurtosis, and Skewness).

3.2.1 Subset 1

This subset contained the Average, Standard Deviation, Kurtosis, and Skewness of MFCC.

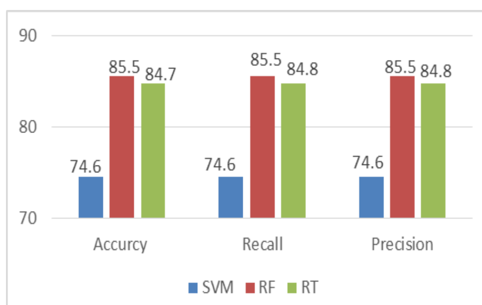

Figure 8: Classification Results Subset 1, Exp B

From Figure 8, the Random Forest classifier reported the highest classification performance rates as follows: accuracy (85.5%), recall (85.5%), and precision (85.5%). Therefore, Random Forest is ranked first, ahead of the other two (2) classifiers.

ISSN: 1992-8645 www.jatit.org E-ISSN: 1817-3195

3.2.2 Subset 2

This Subset contained the Standard Deviation, Kurtosis, and Average of MFCC.

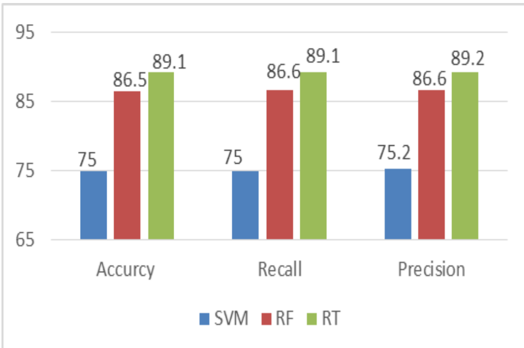

Figure 9: Classification Results, Subset 2, Exp A

From Figure 9, the Random Tree classifier reported the highest classification performance rates as follows: accuracy (89.1%), recall (89.1%), and precision (89.2%). Therefore, Random Tree is ranked first, ahead of the other two (2) classifiers. The changes in this subset (excluding MFCC Skewness features) did improve the classification performance, especially the Random Tree classifier.

3.2.3 Subset 3

This subset consists of only two features, Standard Deviation, and Kurtosis of MFCC.

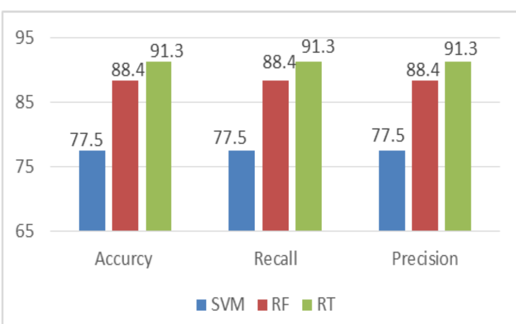

Figure 10: Classification Results Subset 3, Exp B

From figure 10, the Random Tree classifier reported the highest classification performance rates as follows: accuracy (91.3%), recall (91.3%), and precision (91.3%). Therefore, the Random Tree is ranked first, ahead of the other two (2) classifiers. This improvement is the result of excluding the MFCC Average features from this subset.

3.2.4 Polynomial kernel function for SVM

Du to the results of Support Vector Machine (SVM) in experiment B (74.6 – 77.5% accuracy). This experiment aimed to improve the performance rates (accuracy, recall, and precision) of SVM. In all experiments above (experiment A and B), the Support Vector Machine kernel function is the Radial Bias Function (RBF). In this experiment, the kernel function of the Support Vector Machine is the Polynomial function. The experiment employed the second feature group (see figure 5).

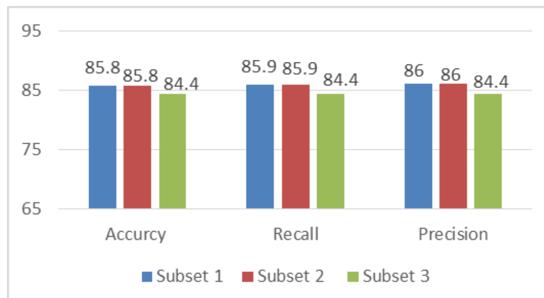

Figure 11: Classification Results Subset 3, Poly function

From figure 11, the SVM classifier reported an improvement in performance rates using the polynomial kernel function in comparison with the radial bias function (RBF) in 3.2.1 – 3.2.3. The highest performance rates are reported using subsets 1 and 2 as following: accuracy (85.8%), recall (85.9%), and precision (86%).

3.3 Experiment C

This experiment aimed to compare our research methods with [13], by using their classifier LR (logistic regression). The features group used in this experiment is the same group used in section 3.2 (subset 3). This is because it shows the highest classification performance in this research. Three feature subsets are used to achieve this experiment, the first subset contained the Average, Standard Deviation, Kurtosis, and Skewness of MFCC.

The second subset contained the Standard Deviation, Kurtosis, and Average of MFCC. While the third subset contained the Standard Deviation and Kurtosis. The following figure (12) shows the classification performance using the second feature group with logistic regression classifier.

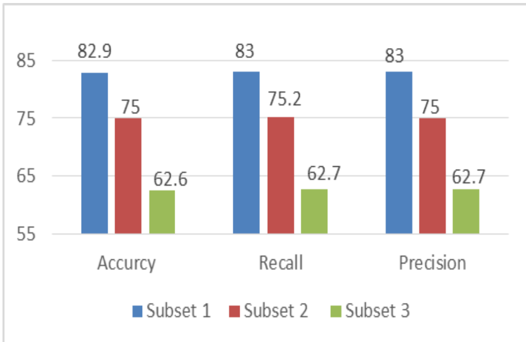

Figure 12: Classification Results Experiment C

From figure 12, the first feature subset performed the highest classification performance (82.9% classification accuracy 83% recall and precision) using logistic regression classifier. While the third subset achieved the lowest classification performance (62.6% accuracy, 62.7% recall, and precision).

4. DISCUSSION

In this section, we are going to discuss the results of our experiments and show how these results could be improved. As shown in the previous section, the recommended features (see section 2.2) achieved generally high classification performances. From Figures 6 and 7 (Results of Experiment A), SVM (RBF kernel function) achieved the highest classification performance, while the Random Tree achieved the lowest.

The second part of experiment (A) (see 3.1.2) shows some improvement in the classification performance. This improvement is the result of adding the Method of moment and the Spectral Roll-Off point features, which seem to be more compatible with SVM than with tree-based classifiers. The results are in agreement with [25]; these features are appropriate for speech classification, but experiment B (section 3.2) shows that MFCC is more appropriate to the specific task of depression classification.

From Figures 8, 9, 10, and 11, the classifiers can provide at least a 74.6% - and up to 91.3% - classification accuracy. These results are compatible with the discussion in [27] to the extent that the features of MFCC are the most used in speech recognition systems because they achieve high performances.

The first part of experiment B (see 3.2.1) showed excellent classification performances for the tree-based classifiers, while SVM (RBF kernel function) achieved the lowest. We eliminated some features (Skewness and Average) in the second and third parts of experiment B to evaluate the effectiveness of these features (see 3.2.2 and 3.2.3). After this elimination (see Figures 8-10), the classification accuracy reached up to 91.3% for the Random Tree classifier. From these results, the features of MFCC would appear to be more appropriate for the tree-based classifiers than for SVM (RBF kernel).

The last part of experiment B (see 3.2.4), is an experiment to improve the classification performance rates of Support Vector Machine, due to the low-performance rates in experiment B (3.2.1 – 3.2.3). To improve the performance rates of the Support Vector Machine, the Polynomial kernel function is used in 3.2.4 instead of the Radial Bias Function (RBF). As a result of this experiment, the Support Vector Machine showed performance improvement rates when we used the Polynomial kernel function (see figure 11).

From this experiment (3.2.4), SVM (RBF kernel) reporting better classification performance with linear feature set (see experiment A), while SVM (Polynomial kernel) reporting better classification performance with non – linear feature set (see experiment B). Further, tree-based classifier reporting acceptable classification performance using linear features (see experiment A), while reporting better classification performance using non-linear features (26 MFCC) – see experiment B.

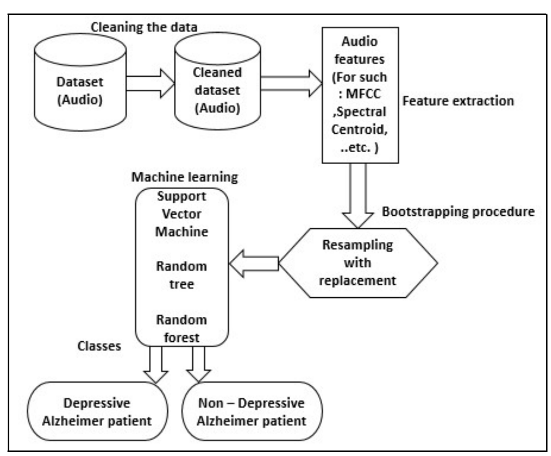

Figure 13: Research methodology

The audio features in this research are selected based on the literature review (not selected randomly). Although results appear better when using a subset of MFCC (see section 3.2.3), it remains to be seen if the results are consistent across datasets and languages. After solving the sampling bias (see figure 4), several classifiers are tested, we mentioned the top three classifiers in this research based on the classification performance (accuracy, recall, and precision) – see figure 13.

The reference [13] performed the same task (depression detection in Alzheimer's patients) using a subset of 130 AD patients (we used 276) from the Pitt Corpus (the same used in this research). The authors achieved a 65.8% classification accuracy, using a Logistic Regression classifier (LR). They introduced many features, including MFCCs and linguistics features. Experiment C shows the significance of our specific methodology (see figure 13). In this experiment, we used the Logistic Regression classifier - used in [13] - with our features and cleaned data.

The research methodology (see figure 13) of this research starting from cleaning the audio streams, and select the feature sets based on the literature review - the feature set used in this research are recommended by the literature review - makes their classifier (Logistic Regression) perform better classification performance using the same dataset.

Our results show that removing the interviewer's voice, the various noises, and the silent moments from the data is critical to depression classification (focus the analysis on the patients' speech). In fact, while silences can be useful for AD detection since they are indicators of hesitations, they are not useful for evaluating depression.

5. CONCLUSION

As mentioned before, Alzheimer's poses many challenges to individuals and societal levels. Moreover, the high cost of medical care exerts the economy. Alzheimer's is not just forgetting the memories, according to [1], Alzheimer's is the sixth-leading cause of death in the United States. Moreover, one in three people (seniors) dies with Alzheimer's or other dementia. Around 30-50% - according to [6] - of Alzheimer's patients are suffering from depression, this could lead to more cognitive decline. Both illnesses share some symptoms such as impaired attention [7], which makes the task of detecting depression in Alzheimer's patients more challenging.

The objective of this paper was to establish a starting point for increasing the performance of machine learning models using speech analysis to classify depressive Alzheimer's patients from their non-depressive counterparts.

The feature sets are selected based on the literature review, while several classifiers are tested, we mentioned the top three classifiers in this paper (based on the performance). The data (audio streams) is cleaned from the interviewer's voice, background noises, and silent moments. Around 30% of the dataset are depressive Alzheimer patients, this makes the data imbalanced (sampling bias) –see figure1.

To solve the sampling bias, Bootstrapping (resampling with replacement) is used. We compared the performances of three classification techniques (Support Vector Machine, Random Tree, and Random Forest) using two different groups of features, one having two subsets of Standard Low-Level (SLL) features, and the other using three subsets of MFCC features.

The results show that the Random Tree classifier achieved the best classification performance (91.3% classification accuracy) with the Standard Deviation and Kurtosis of MFCC features, while SVM achieved the best classification performance (89.1 classification accuracy) with the SLL features (see section 2.2). From the results, cleaning the data (audio streams) from noises including the silent moments, selecting the audio features based on the literature review, has a significant effect on classification results.

For future work, we recommend studying the behavior of tree classifiers (Random Tree and Random Forest) because many research papers suggest that they are unstable, and could be affected by slight modifications of the data used. Moreover, we recommend implementing the same task by employing a balanced dataset, and by testing the different features subsets and classifiers in another language. The feature sets can be extended with the Auditory Filter Bank Temporal Envelopes (AFTE) feature set, which was a feature set that ranked second in speech classification performance based on the results of [25].

ISSN: 1992-8645 www.jatit.org E-ISSN: 1817-3195

ACKNOWLEDGMENT

The research presented in this paper was financially supported by NSERC (Natural Sciences and Engineering Research Council of Canada) RGPIN-2018-05714 within the project "Pattern recognition for the detection and monitoring of verbal and non-verbal alterations in Alzheimer's disease".

REFERENCES:

[1] Alzheimer's Association. 2018 Alzheimer's disease facts and figures. Alzheimer's & Dementia. 2018 Mar 1, 14(3):367-429.

[2] Roark B, Mitchell M, Hosom JP, Hollingshead K, Kaye J. Spoken language derived measures for detecting mild cognitive impairment. IEEE transactions on audio, speech, and language processing. 2011 Feb 7, 19(7):2081-90.

[3] Gauthier S, Reisberg B, Zaudig M, Petersen RC, Ritchie K, Broich K, Belleville S, Brodaty H, Bennett D, Chertkow H, Cummings JL. Mild cognitive impairment. The lancet. 2006 Apr 15, 367(9518):1262-70.

[4] Grundman M, Petersen RC, Ferris SH, Thomas RG, Aisen PS, Bennett DA, Foster NL, Jack Jr CR, Galasko DR, Doody R, Kaye J. Mild cognitive impairment can be distinguished from Alzheimer disease and normal aging for clinical trials. Archives of neurology. 2004 Jan 1, 61(1):59-66.

[5] American Psychiatric Association. Diagnostic and Statistical Manual of Mental Disorders. (2013). Washington. DC American Psychiatric Association.

[6] Lee HB, Lyketsos CG. Depression in Alzheimer's disease: heterogeneity and related issues. Biological psychiatry. 2003 Aug 1; 54(3):353-62.

[7] Korczyn AD, Halperin I. Depression, and dementia. Journal of the neurological sciences. 2009 Aug 15; 283(1-2):139-42.

[8] Thorpe L. Depression vs. dementia: How do we assess. The Canadian Review of Alzheimer's disease and Other Dementias. 2009 Sep; 12(3):17-21.

[9] Modrego PJ, Ferrández J. Depression in patients with mild cognitive impairment increases the risk of developing dementia of Alzheimer type: a prospective cohort study. Archives of neurology. 2004 Aug 1; 61(8):1290-3.

[10] Lyketsos CG, Steele C, Baker L, Galik E, Kopunek S, Steinberg M, Warren A. Major and minor depression in Alzheimer's disease: prevalence and impact. Journal of Neuropsychiatry and Clinical Neurosciences. 1997 Sep 1; 9(4):556-61.

[11] Milwain EJ, Nagy Z. Depressive symptoms increase the likelihood of cognitive impairment in elderly people with subclinical Alzheimer pathology. Dementia and geriatric cognitive disorders. 2005; 19(1):46-50.

[12] Shim YS, Yang DW. Depression as prognostic factor: 6 months follow-up in a geriatric institution. Archives of gerontology and geriatrics. 2006 Sep 1; 43(2):277-83.

[13] Fraser KC, Rudzicz F, Hirst G. Detecting late-life depression in Alzheimer's disease through analysis of speech and language. InProceedings of the Third Workshop on Computational Linguistics and Clinical Psychology 2016 Jun (pp. 1-11).

[14] Hosseinifard B, Moradi MH, Rostami R. Classifying depression patients and normal subjects using machine learning techniques and nonlinear features from EEG signal. Computer methods and programs in biomedicine. 2013 Mar 1; 109(3):339-45.

[15] Asgari M, Shafran I, Sheeber LB. Inferring clinical depression from speech and spoken utterances. In2014 IEEE international workshop on Machine Learning for Signal Processing (MLSP) 2014 Sep 21 (pp. 1-5). IEEE.

[16] Alghowinem S, Goecke R, Wagner M, Epps J, Gedeon T, Breakspear M, Parker G. A comparative study of different classifiers for detecting depression from spontaneous speech. In2013 IEEE International Conference on Acoustics, Speech and Signal Processing 2013 May 26 (pp. 8022-8026). IEEE.

[17] Fraser KC, Meltzer JA, Rudzicz F. Linguistic features identify Alzheimer's disease in narrative speech. Journal of Alzheimer's disease. 2016 Jan 1; 49(2):407-22.

[18] Tanveer M, Richhariya B, Khan RU, Rashid AH, Khanna P, Prasad M, Lin CT. Machine learning techniques for the diagnosis of Alzheimer's disease: A review. ACM Transactions on Multimedia Computing, Communications, and Applications (TOMM). 2020 Apr 15; 16(1s):1-35.

[19] Syed MS, Syed ZS, Lech M, Pirogova E. Automated Screening for Alzheimer's Dementia through Spontaneous Speech. INTERSPEECH (to appear). 2020:1-5.

[20] Becker JT, Boiler F, Lopez OL, Saxton J, McGonigle KL. The natural history of

ISSN: 1992-8645 www.jatit.org E-ISSN: 1817-3195

Alzheimer's disease: description of study cohort and accuracy of diagnosis. Archives of Neurology. 1994 Jun 1; 51(6):585-94.

[21] Goodglass H, Kaplan E. The Assessment of Aphasia and Related Disorders, 2nd edn Lea & Febiger: Philadelphia. Dictionary of Biological Psychology. 1983; 230.

[22] Hamilton M. The Hamilton Depression Scale—accelerator or break on antidepressant drug discovery. Psychiatry. 1960; 23:56-62.

[23] Zimmerman M, Martinez JH, Young D, Chelminski I, Dalrymple K. Severity classification on the Hamilton depression rating scale. Journal of affective disorders. 2013 Sep 5; 150(2):384-8.

[24] McKay C, Fujinaga I, Depalle P. jAudio: A feature extraction library. InProceedings of the International Conference on Music Information Retrieval 2005 Sep (pp. 600-3).

[25] Breebaart J, McKinney MF. Features for audio classification. InAlgorithms in Ambient Intelligence 2004 (pp. 113-129). Springer, Dordrecht.

[26] Li D, Sethi IK, Dimitrova N, McGee T. Classification of general audio data for content-based retrieval. Pattern recognition letters. 2001 Apr 1; 22(5):533-44.

[27] O'Shaughnessy D. Automatic speech recognition: History, methods and challenges. Pattern Recognition. 2008 Oct 1; 41(10):2965-79.

[28] Hall M, Frank E, Holmes G, Pfahringer B, Reutemann P, Witten IH. The WEKA data mining software: an update. ACM SIGKDD explorations newsletter. 2009 Nov 16; 11(1):10-8.

[29] Smola AJ, Schölkopf B. A tutorial on support vector regression. Statistics and computing. 2004 Aug 1; 14(3):199-222.

[30] Breiman L. Random Forests Mach Learn. 2001; 45: 5–32.

[31] Drmota M. Random trees: an interplay between combinatorics and probability. Springer Science & Business Media; 2009 Apr 16.

[32] Preacher KJ, Hayes AF. Asymptotic and resampling strategies for assessing and comparing indirect effects in multiple mediator models. Behavior research methods. 2008 Aug 1; 40(3):879-91.