

Dementia Detection with Phonetic and Phonological Features

M. Zakaria Kurdi

University of Lynchburg, Lynchburg, VA
Kurdi_m@lynchburg.edu

Abstract

In this paper the ADDR_eSS challenge dataset was used for training and testing a binary classifier designed to diagnose AD. This dataset consists of transcripts of descriptions of the Cookie Theft picture, produced by 54 subjects in the training part and 24 subjects in the test part. Two machine learning experiments were conducted on the task of classifying transcribed speech samples with text samples that were produced by people with AD from those produced by normal subjects. The first experiment showed that, among all the subtypes of phonetic and phonological features covered in this paper, vowels provided the best classification performance. The second experiment that used four feature selection techniques showed that the adopted phonetic and phonological features provided about 0.87 F1 score. This performance is close to the best performance reported in the address challenge, by systems using multiple linguistic levels and machine learning techniques. This result confirms the importance of the covered features as indicators of dementia.

Introduction

Alzheimer's disease is a neurological condition characterized by a decline in cognitive function, including memory loss, impaired reasoning, and a degradation in language abilities. In individuals with dementia, including conditions such as Alzheimer's Disease (AD), the cognitive decline can affect various aspects of language and communication, including phonology. Dementia can lead to changes in the muscles used for speech, affecting articulation and pronunciation. This may result in alterations in the production of sounds such as changes in the rhythm and intonation of speech (prosody). In other words, AD can affect the overall phonological quality of communication.

Several previous works focused on lexicon, syntax, and acoustic features of spoken language produced by people with dementia (Kurdi, 2023). However, little work has been done about the study of the impact of AD on phonetic and phonological aspects of the language and on using features of these aspects to automatically diagnose AD based on text

and speech produced by patients. Hence, this work aims to fill this gap by conducting a systematic study of the phonetic and phonological features using transcribed samples of the speech produced by AD patients and healthy control subjects.

Set of Features

Consonant Features

Consonant Place of Articulation (CPA) is about the location of the restriction in the vocal tract, where the production of the consonant occurs. Eighteen features about CPA have been used in this work (For more details about the phonetic and phonological features used in this paper, please refer to (Kurdi, 2017)).

Consonant Manner of Articulation (CMA) is about the way the airflow is obstructed or modified while producing a consonant sound. As we see in table 2, fricatives and friAffLiq are the only CMA features with a systematic high ranking.

Results and Discussion

Seven Machine-Learning algorithms have been used in the experiments: Logistic Regression (LR), Support Vector Machine (SVM), Adaptive Boosting (AB) with 100 estimators, bagging (Breiman, 1994), and Random Forest (RF) with 2 as maximum of depth (Ho, 1995) and eXtreme Gradient Boosting (XG). The following parameters have been used with XG: `learning_rate = 0.001`, `n_estimators=5600`, `max_depth = 5`, `min_child_weight=1`, `gamma=0`, `subsample=0.8`, `colsample_bytree=0.9`, `objective= binary:logistic`, `seed = 25`. In addition, a Multilayer Perceptron (MLP) was used with the following parameters: `max_iter=200`, `hidden_layer_sizes= 50`, `activation function: tanh`, `solver = adam`, `alpha=1e-8`. All the machine learning parameters

have been selected empirically, after having tried multiple combinations the ones that gave the optimal results were adopted. The seven adopted machine learning algorithms were selected for their better performance after having done some experiments with other algorithms, such as Decision Trees and Naïve Bayes. F1-score are reported as they have been used in the ADreSS challenge and they are the result of the combination of recall and precision.

In this evaluation, the data is split into two parts: one for training and one for testing. The splitting is the same as the one proposed by the ADDRess challenge. Using the same splitting allows us to use the papers involved in the ADDRess Challenge as baseline that can be compared to the models proposed in this paper.

Evaluation per Linguistic Types

Here are some observations from figure 1. Except for Miscellaneous, all the linguistic types perform relatively well. Nonetheless, those performances are lower than the best reported in the ADDRess Challenge or in (Kurdi, 2023). Vowel features provide the top classification results with the LR machine learning algorithm, with F1 of 0.81. These results confirm the high impact of the vowel features, as shown in table 3. The number of features per type doesn't seem to systematically impact the results. For example, although CMA has about three times more features than vowels, it gives a slightly lower performance. This result could be because CMA has bigger internal redundancies. The LR algorithm provides the best classification results in four out five types.

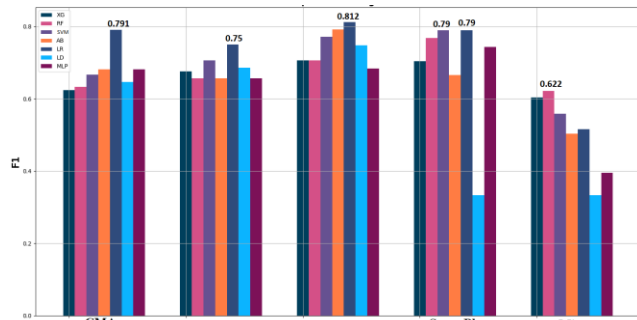


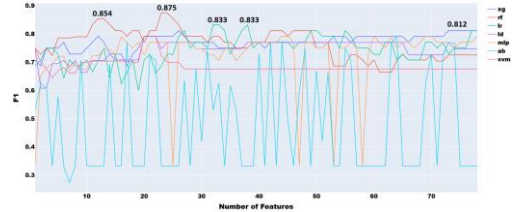
Figure 1. F1 results of the features grouped by their linguistic type

Evaluation with Feature Selection

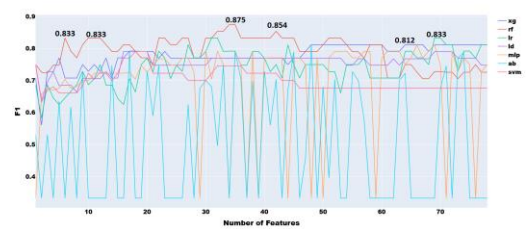
The F1 results of this evaluation are presented in figure 2. First, the results show that more features don't necessarily bring a classification improvement, probably because of the internal redundancies between the features. The results also show that the four feature selection techniques help provide the same performance of 0.875, with information Gain and Anova requiring less features to reach their top performance than the two other feature selection. Despite the small size

of the training data, this performance suggests that phonetic and phonological features alone help obtain a decent F1 score.

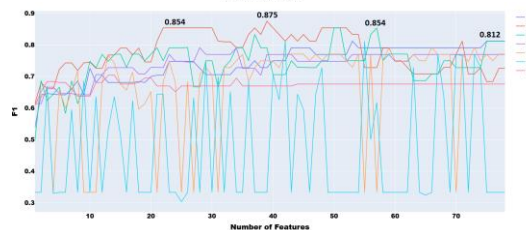
Anova



χ^2



ReliefF



Info. Gain

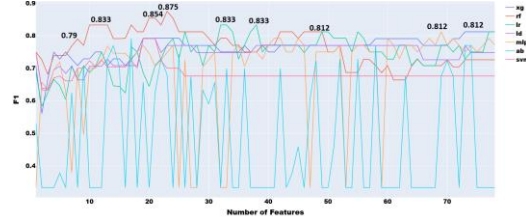


Figure 2. F1 score of the gradual feature lengths, with the four adopted feature selection techniques

Conclusion

This paper is about exploring the impact of AD on pronunciation and how phonetic and phonological features can help build an automatic classifier that can predict whether a transcribed speech sample is produced by someone with AD.

The results showed that some aspects of phonetics and phonology such as vowels, syllabic features are strongly impacted by AD. Besides, by using the phonetic and phonological features for detecting AD a near state-of-the-art performance was obtained. This result confirms the importance of those features.

References

Breiman, L. 1994. Bagging Predictors. Department of Statistics. University of California Berkeley. Technical Report No. 421.

<https://www.stat.berkeley.edu/~breiman/bagging.pdf> Retrieved Jan. 2024; 08-23.

Ho T K. 1995. Random Decision Forests. In: Proceedings of the 3rd International Conference on Document Analysis and Recognition, Montreal, QC, August 14–16.

Kurdi, M. Zakaria 2017, *Traitement Automatique des langues et linguistique informatique 1*. London : ISTE Editions.

Kurdi, M. Zakaria 2023. *Automatic Identification of Alzheimer's Disease using Lexical Features extracted from Language Samples*, arXiv:2307.08070v1,

<https://doi.org/10.48550/arXiv.2307.08070>, accepted in the Journal of Medical Artificial Intelligence.