Using pauses in speech as a feature to classify dementia patients

Anne Krakers

October 24, 2022

Research Proposal

Anne Krakers, S4075927

1 Introduction

Dementia is a neurodegenerative disorder in which cognitive functioning deteriorates. Combined with memory impairments and impairments in executive functioning, the progressive decline in language ability is a hallmark of dementia that places a significant burden on both people with dementia and their caregivers [7]. Language disabilities can occur early in the course of dementia and become more frequent as the disease progresses [3]. One of these language disabilities is the use of pauses in speech. Healthy adults tend to follow predictable patterns of pausing, whereas people with dementia have atypical patterns of pausing in speech [7]. Earlier studies have observed that people with dementia have an increased difficulty in finding words and therefore produce more pauses [5]. These pauses often have an increased duration than pauses in normal speech [5]. When a model can be used to classify if someone has dementia or not based on the amount and duration of pauses the person produces in their speech, this can be very useful in diagnosing people with dementia, especially for early diagnosis. Early diagnosis is very important because the it helps in better management for both patients and caretakers [9] . The goal of this thesis is to examine if pauses in speech can be used as a feature to classify dementia patients, the research question is: "Can pauses be used as a feature to classify who has dementia and who doesn't by training a model on a dataset that includes both dementia and normal speech?".

2 Background

In recent times, there have been various research going around in the task of using pauses to diagnose dementia. Aharon Satt et al. [6] describes a study of a protocol and system for automatic detection and status determination of early-stage dementia and Mild Cognitive Impairment from speech and voice recordings. The 89 participants in the recordings were eldery people that can be divided into three groups according to their medical profile: people with the early stage of Alzheimer disease (AD), people with MCI, and a control group of healthy people with similar age range and demographic attributes. The participants were recorded while performing several vocal tasks. The recordings were used to extract vocal features,

these features were used to train and test a classifier. This classifier aimed at distinguishing between three states: early AD, MCI and healthy. They used several features from different recordings, which came from different tasks. They used amount of pauses, duration of spauses, voiced and unvoiced segments, response time, speech rate, correctness of spoken sentences repetition, irregularity in the pronunciation rate and pronunciation errors. A Support Vector Machine binary classifier with Gaussian radial basis function kernel was used to classify. The study revealed that vocal features enable quite accurate classification of healthy and demented subjects. The classification accuracy was relatively high and the results encourage moving towards an automatic system for dementia assessment. This thesis examines whether it will also be possible to only use pauses as a feature, and reach this high classification accuracy as well.
Rachel A. Sluis et al. [7] examined the number and duration of pausing in speech for people with dementia compared to healthy controls and identified how such measures are impacted by disease severity. The data used is the same data as will be used in this thesis, namely data from the DementiaBank. In their research they conducted a series of one-way analyses of covariance (ANCO-VAs). The study found that there is a progressive increase in the duration of pausing between healthy controls, the mild dementia group and the moderate dementia group respectively. The findings show support for the analysis of pausing as an indication of speech dysfluency using an automated approach. These findings are very relevant for this thesis, as it will be examined if the pauses can be used as a feature to classify dementia patients.

3 General approach

In this thesis, data from TalkBank will be used. TalkBank is the world's largest open acces integrated repository for spoken language data [1]. We will use data from the DementiaBank to be exact, which is a component of TalkBank.
As the goal of this thesis is to create a model that uses pauses (both silent and filled pauses) as a feature to classify, all other things that are in the data will be excluded. This means that all text and punctuation will be removed, except the brackets with dots that represent silent pauses and the ampersand hyphen marks followed by stopwords that represent filled pauses. Silent pauses are presented as (.), (..), (...) in the data, which means short pause, long pause and very long pause respectively. Filled pauses are pauses filled with words like ah, eh, hm and uhm. These pauses are preceded by '&-', so for example 'uhm' will look like this in the data: &-uhm. Since everything else is removed from the data, only the silent and filled pauses will be used by the classifier. Both the number of pauses and the length of the pauses will be taken in account. The goal is to use three models, as then they can be compared to see which scores best. The three models that will be used are a SVM classifier, Naïve Bayes and the KNN algorithm.
SVM is a representation of the training data as points in space separated into categories by a clear gap that is as wide as possible. New examples are then mapped into that same space and the classifier will predict which category they belong to, based on which side of the gap they fall [4]. The main advantage why SVM will be used in this thesis is because it is robust to outliers [8]. There are some cases in the data where people with dementia do not use any pauses, SVM can propably handle this well.
Naive Bayes algorithm is based on Bayes' theorem with the assumption of independence between every pair of features. Naive Bayes classifiers work well in many real-world situations such as document classification [8], this is why it will be used in this thesis.

The k-nearest neighbours algorithm is a supervised learning classifier, which uses proximity to make classifications or predictions about the grouping of an individual data point. It assumes that similar points can be found near each another. A class label is assigned based on a majority vote, so the label most represented around a given data point is used [8]. KNN is used because it is robust to noisy data. KNN will be used because it is known for it's highly accurate predictions [2].

4 Expected output and evaluation

The output of the models will be that a certain person does or does not have dementia. The results will be evaluated by calculating the precision, recall, accuracy and F1-score for each model. The three models will then be compared and a conclusion if pauses can be used as a feature to classify if a person has dementia or not will be drawn from this.

References

[1] The TalkBank CLARIN Knowledge Centre | CLARIN ERIC. https://www.clarin.eu/blog/talkbank-clarin-knowledge-centre.

[2] Usage of k-Nearest Neighbors (KNN). https://www.ibm.com/docs/en/ias?topic=knn-usage.

[3] Paulo Caramelli, Leticia Lessa Mansur, and Ricardo Nitrini. Chapter 32 - language and communication disorders in dementia of the alzheimer type. In Brigitte Stemmer and Harry A. Whitaker, editors, *Handbook of Neurolinguistics*, pages 463–473. Academic Press, San Diego, 1998.

[4] Rohit Garg. 7 Types of Classification Algorithms. https://analyticsindiamag.com/7-types-classification-algorithms/, 10 2021.

[5] A. Pistono, J. Pariente, C. Bézy, B. Lemesle, J. Le Men, and M. Jucla. What happens when nothing happens? an investigation of pauses as a compensatory mechanism in early alzheimer's disease. *Neuropsychologia*, 124:133–143, 2019.

[6] Aharon Satt, Alexander Sorin, Orith Toledo-Ronen, Oren Barkan, Ioannis Kompatsiaris, Athina Kokonozi, and Magdalini Tsolaki. Evaluation of speech-based protocol for detection of early-stage dementia. *Proceedings of the Annual Conference of the International Speech Communication Association, INTERSPEECH*, 08 2013.

[7] Rachel A. Sluis, Daniel Angus, Janet Wiles, Andrew Back, Tingting (Amy) Gibson, Jacki Liddle, Peter Worthy, David Copland, and Anthony J Angwin. An automated approach to examining pausing in the speech of people with dementia. *American Journal of Alzheimer's Disease & Other Dementias®*, 35:1533317520939773, 2020. PMID: 32648470.

[8] Ray Sunil. Understanding Support Vector Machine(SVM) algorithm from examples (along with code). https://www.analyticsvidhya.com/blog/2017/09/understaing-support-vector-machine-example-code/, 8 2021.

[9] S. Thapa, U. Naseem, P. Singh, H. Huo, G. Bharathy, and M. Prasad. Exploiting linguistic information from nepali transcripts for early detection of alzheimer's disease using natural

language processing and machine learning techniques. *International Journal of Human-Computer Studies*, 2022.