Section 1: Research Problem Formulation and Critical Analysis

Step1: Finding a phenomenon

Phenomenon	What	Why	How
Alzheimer's disease affects speech center	Brain cells die and the brain shrinks affecting cognitive abilities	Memory forming locations-hippocampus gets deteriorated due to plaque formation	genetics/old age/decreased blood flow
Changes in speech patterns	syntactic, semantic, and information impairments	Alzheimer's disease	Linguistic skills breakdown

Goal/Question:

Predicting the risk of Alzheimer's disease by inferring the language of the subject

Metric:

- Accuracy = (TP+TN / TP+FP+TN+FN) * 100
- F1 = 2*(Precision*Recall)/(Precision + Recall)
- Stopping condition: monitoring validation accuracy, stop when validation starts to decrease or loss starts to increase

TP = True Positive, TN = True Negative, FP = False Positive, FN = False Negative

Step2: Understanding the State of the Art

Literature Review:

- Papers have used LSTM, CNN-LSTM models, LSTM stacked on GRU with attention mechanism using GloVe word embedding as input
- Minimal feature engineering and Part Of Speech tagging to every word in the corpus is done
- Used data augmentation methods such as Synonym Substitution, Word Embedding Substitution, Contextual Augmentation to make up for low training data. For example - "What a beautiful car!", "beautiful" is replaced by synonyms like "nice".
- Recent papers have used pre-trained embedders like BERT, XLNet, XLM along with masking and then passed them to Bi-LSTM for classification
- Some papers used feature banks such as ComParE, MRCG functionals which focused on extracting acoustic features

Step3: Determining the basic ingredients

Input:

1) Recording of subjects speaking and narrating details of pictures
2) speech to text transcripts of subjects with Alzheimer's and without Alzheimer's
3) pre-trained word embedding like GloVe, Word2Vec, fastText

Output:

Prediction of "Alzheimers" or "control" labels for Alzheimer or not respectively

Hypothesis:

Linguistic_Features = {

Noun rate, Pronoun rate, adjective rate, verb rate, filler words rate, **repetitiveness**, part of speech tags, Story recall/consistency, pause duration

}

The hypothesis function will be a function of the above given linguistic features

Step4: Formulating Specific Mathematical Definitions

- The model will yield P(AD=1/transcript) or P(AD=0/transcript) whether the person has Alzheimer disease or not given the transcript of his speech.
- h(x) = F(Linguistic_features, hyperparameters)

  Where h(x) is the estimator for predicting Alzheimer or not

Step 5: Selecting the toolkit

The implementation was done using

1) TensorFlow
2) Scikit-Learn
3) Pandas
4) NumPy
5) Matplotlib

Step6: planning the model

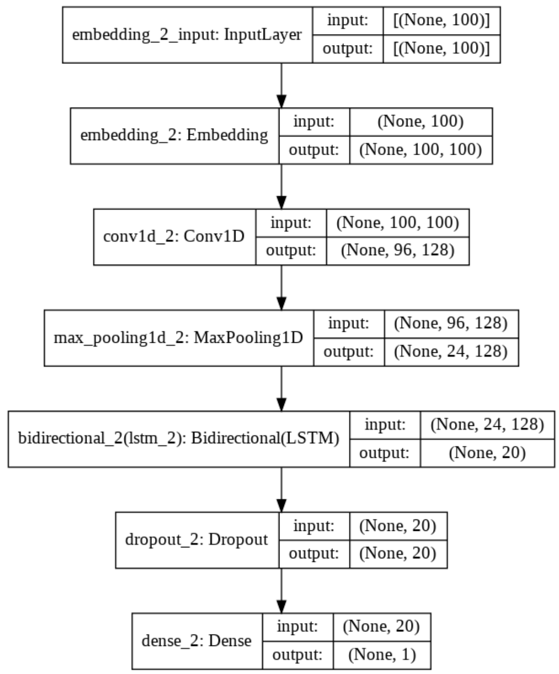
I follow a language-based model using CNN-Bi-LSTM architecture

In CNN-LSTM, convolution is used to extract features and LSTM is used to learn the patterns and these both collectively bias the model towards the solution.

"What" model

- We use this model to find an equation that describes the data and helps us reach our well-defined goals, We want an equation that describes the data and maps the data to an output using existing NLP methods. We also want to understand why are things the way it is, Which features in the linguistic patterns are optimized and which are different.

- we need to capture complex patterns and learn intricate features of linguistics and potentially, the speech of the subject



Step7: Model Implementation

The model followed a sequential CNN-LSTM architecture

1) Loss = binary crossentropy
2) Dropout = 0.4
3) Optimizer = Adam
4) Conv filters = 128, strides = 1
5) Embedding dimensions = vocab_len*100

Step8: Completing the model

- On preliminary testing, the model has an accuracy of 85.52 and an F1 score of 84.42 which is in the good range of current papers.
- The model can be made more complex by adding more dense layers, dropouts, LSTM layers.

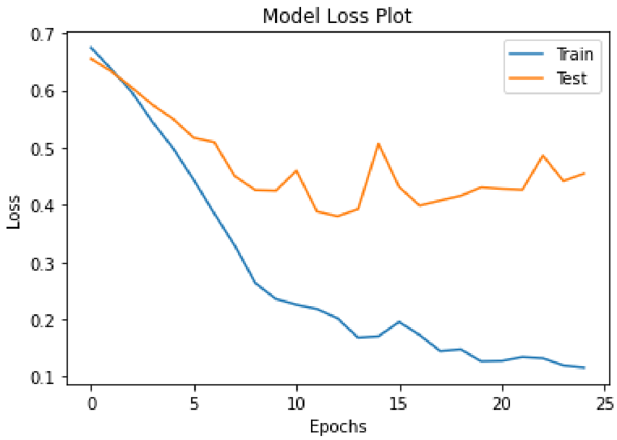
- Early stopping has been used to break the training process once the model accuracy decreases for some steps
- Grid search was implemented to find the optimal hyperparameters which yield the maximum test accuracy

Step9: Testing of the model

We use the following parameters while testing the model

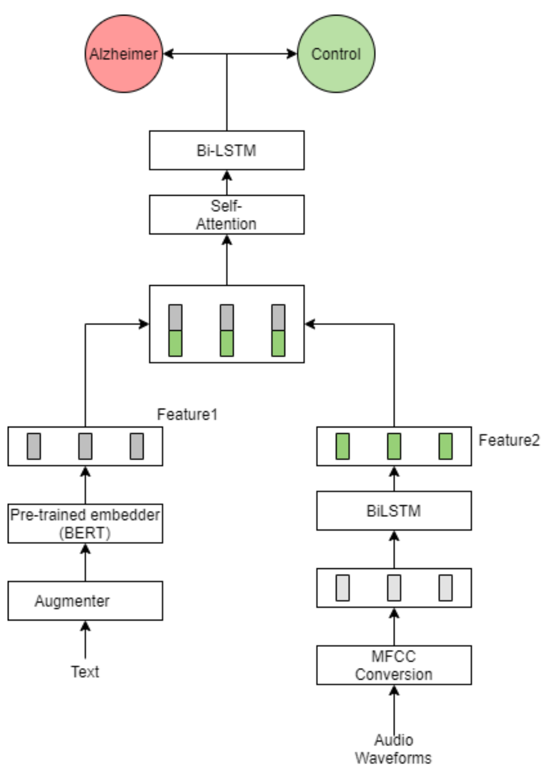
1)Test accuracy

2)Test loss



Step10: Future Work

1) A multimodal approach to this goal can be used in the future for a better feature representation which comprises of a speech and language model in the architecture with audio extraction tools like Praat to use acoustic features along with linguistic features will help us determine the severity of the symptoms.



2) Due to less training data and bias towards English words/ semantics, a multi-language model needs to be proposed which might also have scope for transfer learning - pre-training on English language and fine-tuning on other languages.

3) Physical features of Alzheimer's and controls subjects such as brain size, gender, age, family history for Alzheimer's could be used as well.
4) adding to linguistic features, we can also add psycholinguistic and sentiment features

Annexure
Code: https://github.com/Shreyas-Bhat/Alzheimer_prediction

References

1. Early detection of Alzheimer's dementia through spontaneous speech(Frontiers Research-Address)(2021), Pranav Mahajan and Veeky Baths
2. Transformer-based deep neural network language models for Alzheimer's disease risk assessment from targeted speech(2021), Roshanzamir et al.
3. Classifying Alzheimer's Disease Using Audio and Text-Based Representations of Speech, Haulcy et al.