FINAL REPORT: NEURAL NETWORK APPLICATION FOR ALZHEIMER'S DISEASE IDENTIFICATION USING EEG DATA

Garvish Bhutani
Student# 1008754279
garvish.bhutani@mail.utoronto.ca

Radhika Banerjea
Student# 1009068045
radhika.banerjea@mail.utoronto.ca

Sunny Zhang
Student# 1008739206
sunnyj.zhang@mail.utoronto.ca

ABSTRACT

This paper outlines the creation and performance of a multiclass classification model used in diagnosing patients as having Alzheimer's, Frontotemporal Dementia (FTD), or presenting as healthy. In this report, some background information for motivation for this project, a final architecture that was chosen for the deep learning model, and a thorough evaluation of the model's performance; taking two baseline models as reference is provided and examined. —-Total Pages: 9

1 INTRODUCTION

This project investigates the use of a neural network to identify patients with Alzheimer's disease using electroencephalogram (EEG) data. Alzheimer's is the most common form of dementia and causes the brain to shrink over time, leading to brain cell death alz (2023). The most common symptom of Alzheimer's is memory loss and a decline in social skills. Typically diagnosed in older adults, Alzheimer's is a condition that can significantly weaken an individual's ability to take care of themselves, and can also severely damage their relationships with those around them. Unfortunately, there isn't currently a cure for Alzheimer's and the condition only worsens with time, sometimes leading to death; but early diagnosis is key to preserving the quality of life for patients and their families. Certain treatments can also slow the progress of the disease and reduce some symptoms.

Current diagnostics are only able to pick up the signs of Alzheimer's in later stages, at which brain function is noticeably worsened, whether through memory loss, or a loss of basic functions such as reflex or difficulty speaking med (2024). This gives doctors limited options in treatment and can significantly hinder the quality of care. However, because Alzheimer's is a neurological disease, the changes in brain activity - measured via EEG - should be clear indicators of worsened brain health before symptoms start to set in. This project aims to build on previous attempts to use EEG data to diagnose Alzheimer's using advanced neural networks. Due to the complexity and volume of the data collected, neural networks are an adequate method through which to identify differentiating features in EEG data. Furthermore, deep learning can be advantageous over supervised learning algorithms due to their ability to handle large unlabeled datasets Saab et al. (2020) as well as automatic feature extraction of otherwise difficult-to-identify features in EEG data.

Alzheimer's is a more difficult condition to identify via EEG because it contains fewer easily distinguished features compared to other neurological conditions more commonly studied through EEG such as epilepsy. Rather than having shorter periods of abnormal signal behavior (such as during an epileptic seizure), patterns are more subtle, with changes in specific frequency bands of the signal Kopčanová et al. (2023). As such, we anticipated the diagnostic process to be much more complicated and expected deep learning architectures that require us to have some information about features to expect to not perform well. Our focus was instead on less defined networks, and optimizing the input data and parameters of those networks.

2 General Illustration of the Project

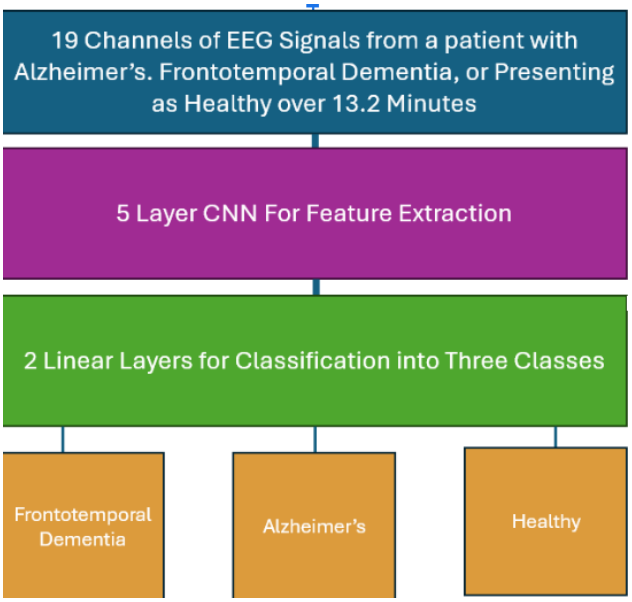


Figure 1: General Illustration of the Project

3 Background

Existing studies have attempted to accomplish the same diagnoses that this project aims to achieve with varying success. Some novel methods can achieve up to 97.6% accuracy (Safi & Mehdi, 2021), but most existing work typically uses supervised learning methods such as support vector machine (SVM) or k-nearest-neighbor (KNNs) classifiers. The following are some examples of past attempts at classifying different neurological conditions using EEG data and learning models:

- *An Approach toward Artificial Intelligence Alzheimer's Disease Diagnosis Using Brain Signals* (Sadegh-Zadeh et al., 2023): This study identifies a potential solution to issues stemming from insufficient data and unbalanced data sets for Alzheimer's disease, using various EEG data augmentation techniques. It then uses a support vector machine classifier on the augmented data, which achieves a higher accuracy with the augmented data than without.
- *EEG Classification of Alzheimer's Disease, Frontotemporal Dementia and Control Normal Subjects using Supervised Machine Learning Algorithms on various EEG Frequency Bands* (Parihar et al.): This study uses the same dataset we are working with to implement supervised learning techniques such as k-nearest neighbors and support vector machine classifiers. This work can be used as a benchmark for the performance of our unsupervised model and illustrates the relative success of existing supervised models.
- *Early detection of Alzheimer's disease from EEG signals using Hjorth parameters* (Safi & Mehdi, 2021): This study adds on to previous Alzheimer's diagnostics work using Hjorth parameters, which are nonlinear features determined through statistical calculations used in other studies to describe characteristics of EEG signals. The highest accuracy achieved by this model is 97.64% and shows the substantial potential for improvement and the feasibility of machine learning techniques in Alzheimer's diagnosis.
- *An Investigation of Deep Learning Models for EEG-Based Emotion Recognition* (Zhang et al., 2020): This study attempts to use deep learning models to distinguish emotions through EEG data. They were able to achieve accuracy up to 94.17%. Although this study doesn't relate directly in application with our project, it serves as a useful proof-of-concept for using deep learning to solve classification problems with EEG.
- *Analyzing EEG Data with Machine and Deep Learning: A Benchmark* (Avola et al., 2022): This study provides a benchmark for using deep learning for EEG signal classification. The accuracy achieved in this study was comparatively lower, with the best model achieving 90.4% accuracy and most others significantly

lower than that. However, the information it provides on the feasibility and need for deep learning models in EEG analysis are valuable to understanding the implementation and requirements of our project.

These sources are selected from a larger collection of past work informing our project but were crucial in identifying the strengths and shortcomings of existing methods, and how deep learning models such as the one implemented in the project can improve on them. Through past work, we identified the following key points:

- A major barrier to Alzheimer's diagnosis is a lack of large balanced and labeled datasets necessary for supervised learning, creating a need for data augmentation or unsupervised learning
- Existing learning models have found great success in diagnosing Alzheimer's and other neurological conditions, indicating that building a working model is achievable
- Deep learning models in EEG analysis is a proven method and their implementation is both necessary and effective

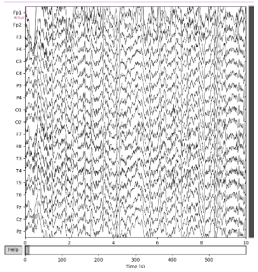
With these takeaways in mind, we attempt to build on this existing work and achieve comparable results to existing supervised training models using our own unsupervised learning model.

4 Data Processing

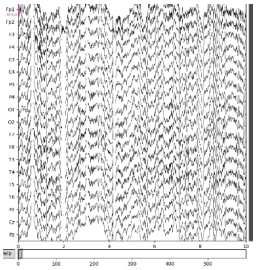
The dataset used is publicly available on OpenNeuro dat (2024) and contains 88 subjects with 19 channels of EEG data for an average of 13.2 minutes. The data is stored in BIDS accepted .set format. Each .set file corresponding to the subjects were parsed into Python using the MNE library. There were two sets of data for each subject available in this dataset, one preprocessed already and one with raw data. We plan to use their preprocessed data as well as preprocess the raw data on our own to determine which produces better results. The data that's already preprocessed may also be unusable on our architecture because it is at times discontinuous and isn't consistently size matched.

In our own preprocessing, there were two main steps: bandpass filtering and artifact repair (includes artifact removal and correction). The reason we weren't able to do re-referencing was because the dataset didn't include data from electrodes A1-A2 which are usually used as reference electrodes for EEG. Bandpass filtering and artifact repair were implemented using MNE libraries. A butter bandpass filter with a lower frequency of 0.5 Hz and a higher frequency of 45 Hz. These values were chosen because they are the limits chosen by the creators of the dataset. Artifacts were then identified and labeled using independent component analysis and then those corresponding to eye blink, line noise, and heartbeat were excluded.

A butterworth bandpass filter fil (2023) is a unique type of bandpass filter that uses something called a butterworth transfer function, which is a pre-defined mathematical function upon which the frequency response of the signal is adjusted. The objective of the filter is to get the data maximally flat. This filter was selected because we were trying to replicate the preprocessing process taken by the creators of the dataset ourselves and then compare the two. The reason why we chose not to immediately use the preprocessed data that was provided was because of some discontinuities in the data that we anticipated could either cause issues during the training process or require excessive further preprocessing to deal with. Initial plots appeared to show that the graphs of our preprocessed data were quite similar to that provided by OpenNeuro. As shown in Figures 2a and 2b, most of the graph's shapes are preserved in



(a) Manually preprocessed data



(b) Preprocessed data from OpenNeuro

our preprocessing, but don't include the discontinuities that exist in the preprocessed data provided by Open Neuro

After the bandpass filter, independent component analysis was used to identify the channel activity that corresponded to irrelevant stimuli such as eye movement or blinking or heartbeat. Independent component analysis takes EEG signals and divides them into groupings based on electrical activity indicative of similar stimuli Kim (2002). For

example, one common grouping is for detecting eye movements, which are usually quite noisy and can interfere with other EEG analysis. ICA isolates eye movement artifacts and allows the user to choose to exclude signal portions that are included in this group, effectively removing the eye movement artifacts from the signal.

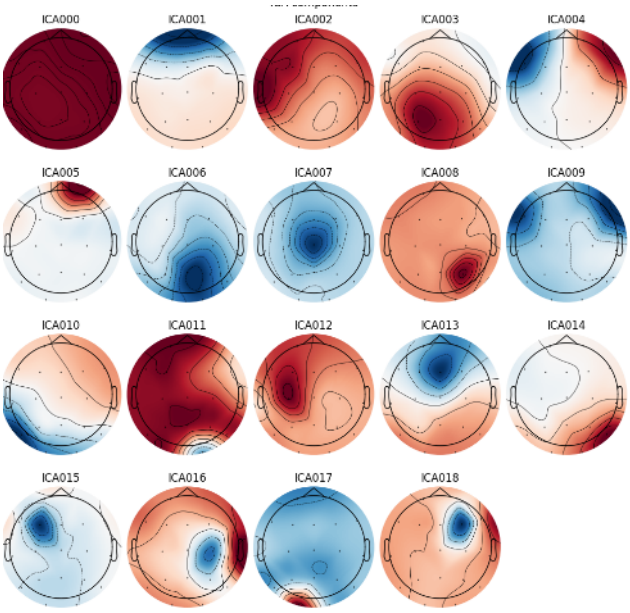


Figure 3: Component plots produced by the ICA python libraries. Note the plots of ICA001 and ICA004, which clearly correspond to eye blink and eye movements respectively eye (2020)

Using the ICA related libraries on mne mne (2024), we were able to conduct ICA analysis and then remove artifacts related to line noise, heartbeats, and eye movement or blinking. ICA is commonly used in EEG analyses and was more reputable than methods such as signal space projection so it was chosen. The fact that libraries were easily accessible also played a role in our selection.

5 Baseline Model

There were two main baseline models that we are examining to compare with our deep learning network: a simple ANN that we developed, and a highly complex graph signal processing based approach developed by a group of researchers in British Columbia.

The first baseline mostly uses conventional machine learning techniques to classify the data, but does extensive processing and re-representation of the data value using graph signal processing. This baseline is selected as an example of a unique unconventional approach to classifying using EEG data. Our goal is primarily to perform similarly or better than this model.

Graph signal processing is a method of data analysis that can be done on time series data which represents signals as the vertices of a graph. The graph is represented mathematically by a set of vertices and a set of edges. The paper uses graph fourier transform to decompose the signal into frequency components. They then used these components to define a set of features of each signal, ranging from graph energy, spectral entropy, to heat trace. Features of each subject were then passed into conventional machine learning methods such as Naive Bayes or Random Forest to make predictions about the subject. Unfortunately, this model didn't have great success, finishing with a test accuracy of only 44%.

The confusion matrix of that model is shown in Figure 4. The labels are as follows: A = Alzheimer's Detection, F = Frontotemporal Dementia, and C = Healthy Controls.

We will use the ANN as a representation of how close we can get using conventional machine learning methods without deep learning and the complex GSP methods. Unsurprisingly, the ANN we used, which only has two linear layers, wasn't able to improve accuracy substantially past a guess. Its final test accuracy was 33%.

```
Validation ROC AUC for class A (Class 0): 0.6181818181818182
Test ROC AUC for class A (Class 0): 0.5631868131868132
Validation ROC AUC for class C (Class 1): 0.7008547008547008
Test ROC AUC for class C (Class 1): 0.9239130434782609
Validation ROC AUC for class F (Class 2): 0.7559523809523809
Test ROC AUC for class F (Class 2): 0.6941176470588236

Validation Precision, Recall, F1-score:
              precision    recall  f1-score   support

           A       0.50      0.09      0.15        11
           C       0.55      0.92      0.69        13
           F       0.57      0.57      0.57         7

    accuracy                           0.55        31
   macro avg       0.54      0.53      0.47        31
weighted avg       0.54      0.55      0.47        31


Test Precision, Recall, F1-score:
              precision    recall  f1-score   support

           A       0.50      0.15      0.24        13
           C       0.27      1.00      0.42         4
           F       0.75      0.60      0.67        10

    accuracy                           0.44        27
   macro avg       0.51      0.58      0.44        27
weighted avg       0.56      0.44      0.42        27
```

Figure 4: Confusion Matrix for Yale BMJ Naive Bayes' Model

6 ARCHITECTURE

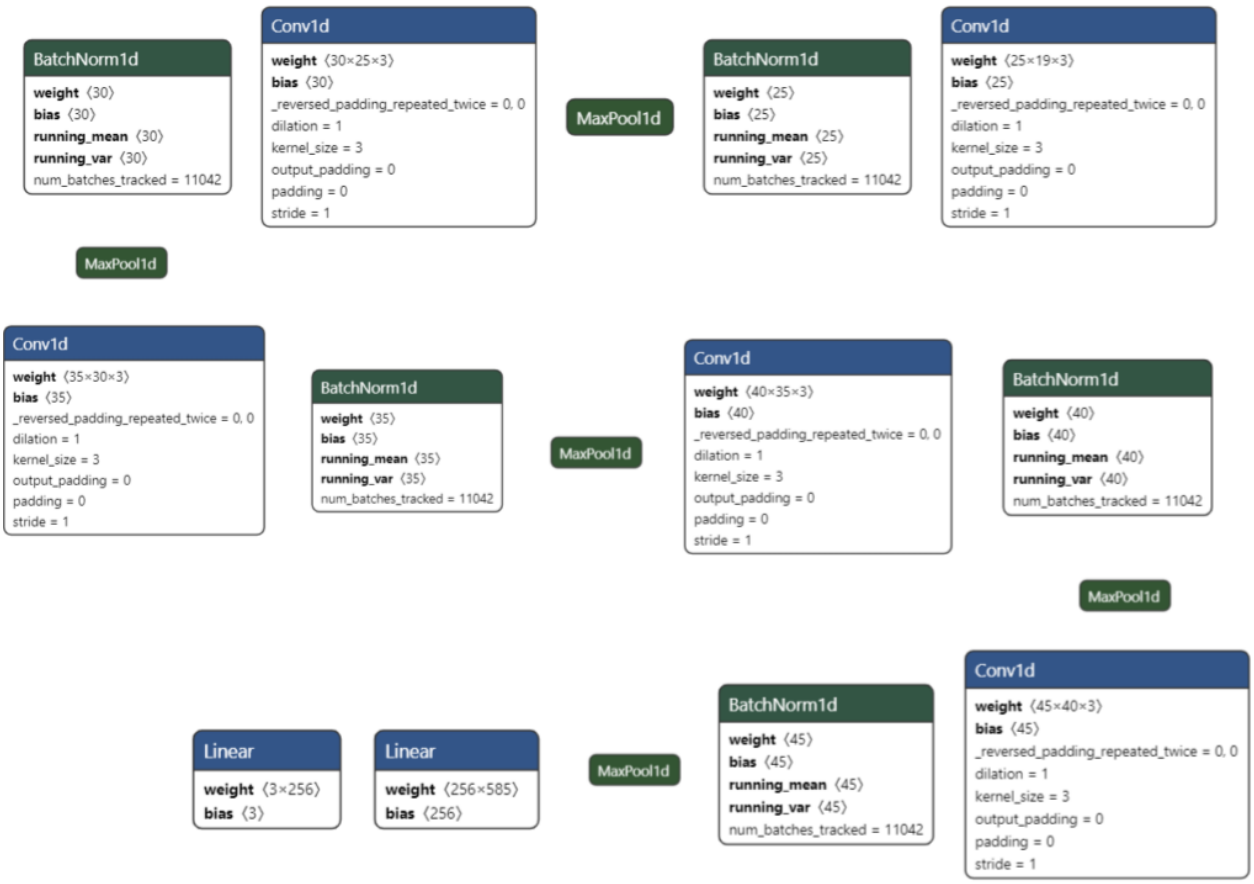


Figure 5: Architecture of the highest performing model (70% test accuracy)

The best model (figure 5) includes 5 CNNs (1 Dimensional) which upsample the number of channels from $19 \rightarrow 25 \rightarrow 30 \rightarrow 35 \rightarrow 40 \rightarrow 50$. Each convolution layer is followed by a batch normalization layer to make the training more stable and a maximum pooling layer of stride = 4 and kernel size = 4 to filter out the important features. None of the layers have padding, and the kernel size and stride for all the convolution layers are set to 3 and 1 respectively. There are 2 linear layers in the end to map the output to 3 classes. Softmax function is applied on the last linear layer to find the probabilities of each class. ReLU activation function is applied to every other layer. For training, stochastic gradient descent (SGD) optimizer along with Cross Entropy Loss and an exponentially decaying learning rate is used. An initial learning rate of 0.1 gave the best results.

7 Quantitative Results

We selected four different methods of analyzing our results as we trained our models. First, we looked at the training curve of the model showing the training and validation accuracy to look for under or overfitting of the model. We then tested the model on a partition of data that the model hadn't been introduced to, our testing dataset. Using these results, we generated a confusion matrix to gain a better understanding of which classes were being predicted well and which were being predicted incorrectly to observe any patterns that might give a better idea of how best to improve the model. Finally, we used the confusion matrix to generate values for recall, precision, and F1 to further assess points for improvement. Examples of our quantitative results can be found in the Evaluation section.

8 Qualitative Results

	Predicted	Actual
0	Healthy	Healthy
1	Healthy	FTD
2	Healthy	Healthy
3	Alzheimer's	Alzheimer's
4	FTD	FTD
5	Healthy	FTD
6	Alzheimer's	Alzheimer's
7	Healthy	FTD
8	Healthy	Healthy
9	Healthy	Healthy

Figure 6: Outputs of the model

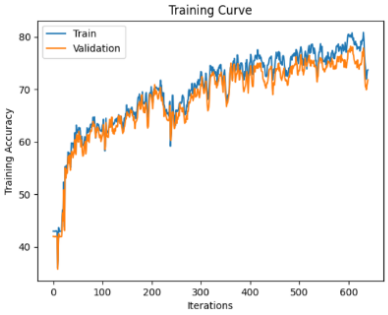
The outputs of the model correspond to the predicted diagnosis for the given subject (figure 6). Our results indicated that it was better suited to handle a two-class problem predicting between healthy and Alzheimer's group patients, rather than between all three classes. This is shown in the confusion matrix above in Quantitative Results. We were also quite selective with our models, which had a high variance in performance, the range sometimes spanning upwards of 50% test accuracy. The most likely explanation is that the model picks up on the incorrect features because the indicators of Alzheimer's are so subtle, causing it to miss the actual features. However, more research into the actual neurological impacts of Alzheimer's would have to be done to improve the consistency of the models.

9 Evaluation of Model on New Data

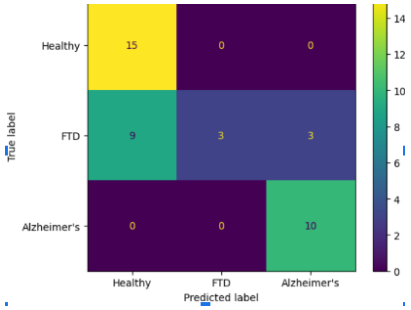
A major issue with EEG data is the lack of well labeled data with consistent electrode placements. This means that it's difficult to find data with the same formatting and quality as our training set, and particularly difficult to find data pertaining to our classification problem involving Alzheimer's. Thus, we instead used a partition of our data to set aside for testing later on. In order to confirm our theories about this, one of our team member's spoke with the Director of Neuroscience at Cove Neuroscience, a startup focusing on diagnostics of neurological conditions using similar methods as those in our project, albeit more complex. They immediately recognized the dataset we were using and noted that while the dataset was unusually good for EEG data, it would be difficult to find new data and extrapolate our findings to them. Results from the assessment are shown in figures 7a and 7b.

As indicated by the confusion matrix, the model is proficient in diagnosing between healthy and Alzheimer's group patients, but struggles to diagnose FTD. The model tends to predict FTD patients to be healthy patients, which could indicate that the features associated with the condition are too small to be detected with our current model.

The accuracy, precision and F1 scores for each class is noted in Table 1



(a) Training curve of the final model



(b) Confusion matrix for final model on testing set (40 samples)

	Precision	Recall	F1
Healthy	62.5%	100%	76.9%
FTD	100%	20%	33.3%
Alzheimer's	76.9%	100%	86.9%

Table 1: Precision, Recall and F1 of the three classes

These results showed that the model performed well on data that it hadn't been introduced yet. We consider its performance to be highly successful and further information about its comparison to other existing models can be found in Discussion. Further evaluation of the model on new data is possible as long as the data has at least 19 channels of EEG data and more than 30 seconds of data.

10 Discussion

Diagnosing neurological conditions using EEG is a relatively new method, and hasn't been well proven. This problem is a difficult challenge for many researchers to tackle over longer periods, much less our group in the last four months. That said, our model performed strongly against others that had been used on this specific dataset as well as with other Alzheimer's datasets. There was a limited selection of studies for this dataset because it was released in 2023Parihar & Swami (2023), but our model outperformed almost every other in terms of testing accuracy. We were able to achieve 26% higher accuracy when compared to the GSP based baseline model we had hoped to match the performance of. Thus, we believe our model has been incredibly successful and has exceeded our expectations. However, these results only came from an extensive process of optimizations, which could be continued to further improve the model.

The best model was found after a lot of reiterations with different hyperparameters intelligently chosen because grid search wasn't always feasible with our compute power and time limit. Some parts of the tuning process are outlined as follows:

10.1 CNN Size

The first parameter we adjusted was the CNN size. Beginning with the basic 3 layer CNN from our progress report, we adjusted the size of the CNN slightly to determine how the complexity of the model affects its performance. Because of the significant increase in training time, we were hesitant in increasing model complexity too much, particularly because we had already spent substantial time trying other model architectures and didn't have access to a lot of compute power.

10.2 Linear Layers

The linear layers are used to map the features to 3 classes which are being detected. Initially, the number of linear layers was decided based on the number of data points the model has after flattening. A lot of models were tested with only a few pooling layers resulting in a lot of data points input into the linear layers. Thus a balance was achieved between the number of linear layers and the difference between output and input number of points in the linear layer. For instance, if the numbers of points in the beginning are 15000, 4 linear layers were used to reduce slowly. However, the results were inconsistent. Some models provided better results with less pooling layers and more linear layers while

the best model has 5 pooling layers and only 2 linear layers (more pooling and less linear). The pooling layers cause the size to shrink thus requiring only 2 linear layers.

10.3 Batch Size

Although the model we trained for the progress report was trained with a batch size of 1, having thousands more data points meant we needed to increase the batch size to improve efficiency. We experimented with batch sizes of 64 and 127, but neither produced substantially different results. Decreasing batch size further later on actually yielded worse results, which led us to stay at the batch size of 64 for the remainder of optimization.

10.4 Learning Rate

Three types of learning rates were tested on this model, constant learning rates, a linearly decreasing learning rate that decreased and equal amount per epoch from an initial value to a final value, and an exponentially decreasing learning rate. It was found that the exponentially decreasing rate fared the best, while decreasing the constant learning rate reduced accuracy and the linearly learning rate had marginal impacts.

10.5 Upsampling vs Downsampling

Upsampling was intensively tested on different number of convolution layers generally with increments of 5 channels each layer. Since the input number of channels is 19, downsampling could only be tried on smaller models (3CNN and 2CNN model). While keeping the parameters constant, upsampling gave better results than downsampling.

10.6 Pooling Type and Kernel Size

Average pooling and max pooling were tested with kernel size =2, 4 and stride = 2, 4. It was observed that max pooling gives better results than average pooling and kernel, stride = 4 (or higher in general) give a better result than kernel, stride = 2. To conclude, MaxPool with kernel, stride = 4 gave the best results.

10.7 Activation Function

Two activation functions were tested in this project: ReLU, and Leaky ReLU. Leaky ReLU was found to perform far more poorly than ReLU consistently across different models. For this reason, ReLU was chosen as the the activation function for the final model.

10.8 Optimization Function

Two optimization functions were tested; those being, Stochastic Gradient Descent (SGD) and Adam. The substantial testing of these two functions led us to believe that Adam did not perform well on our 3CNN model nor on our 5CNN model. Overall it was noticed that SGD produced the highest testing accuracy on both models, and was therefore used.

10.9 Different Types of Neural Networks

Other architectures were considered but we decided they weren't appropriate for the problem we were addressing. The alternative architecture we worked the most with was the LSTM. Although this architecture has been used to diagnose neurological conditions such as epilepsyXu et al. (2020) and schizophrenia Shoeibi et al. (2021), we determined that it wouldn't be viable when diagnosing Alzheimer's because LSTMs require features that are well encoded. These features are present in conditions with acute impacts such as epileptic seizures as opposed to long term debilitation such as in the case of Alzheimer's.

Contrary to other neurological conditions commonly diagnosed by EEG, Alzheimer's doesn't exhibit easily distinguishable signals. This raises two main problems: firstly, it's near impossible to identify the behaviour or frequency of distinguishing features for training and it's difficult to identify reasons why the model is underfitting to adjust it. These make this project extremely difficult, particularly when our group members didn't have extensive neuroscience background.

However, as shown throughout the report, our model has performed better than expectations especially while differentiating patients with Alzheimer's from healthy patients.

11 Ethical Considerations

Several ethical considerations need to be taken with respect to the ethics of the model that we have constructed for this project.

1. Proper advertising about the testing results of the model. As has been noted during class, just stating the testing accuracy of a model is not enough to actually represent its efficacy. As indicated earlier our model does tend to under diagnose Frontotemporal Dementia and mistake it for patients presenting as healthy or as having Alzheimer's. For people using this model to detect Frontotemporal Dementia, this can be dangerous.
2. Due to the limited amount of data that this set was trained on we cannot get a proper gauge on the types of biases that the model may have comprehensively. For example, if it has racial or gender-based biases that we were not able to get a sense of from our limited testing here, that would certainly be an ethical concern.

References

2020. URL https://eeglab.org/tutorials/06_RejectArtifacts/RunICA.html#inspecting-ica-components.

2023. URL https://www.mayoclinic.org/diseases-conditions/alzheimers-disease/symptoms-causes/syc-20350447#:~:text=Alzheimer's%20disease%20causes%20the%20brain,thinking%2C%20behavior%20and%20social%20skills.

2023. URL https://www.mathworks.com/help/signal/ref/butter.html.

2024. URL https://openneuro.org/datasets/ds004504/versions/1.0.6.

2024. URL https://www.alz.org/alzheimers-dementia/diagnosis/medical_tests.

Jan 2024. URL https://mne.tools/stable/auto_tutorials/preprocessing/40_artifact_correction_ica.html#what-is-ica.

Danilo Avola, Marco Cascio, Luigi Cinque, Alessio Fagioli, Gian Luca Foresti, Marco Raoul Marini, and Daniele Pannone. Analyzing eeg data with machine and deep learning: A benchmark, 2022. URL https://arxiv.org/abs/2203.10009.

Eung-Soo Kim. Independent component analysis of eeg and source position estimation, Feb 2002. URL https://www.researchgate.net/publication/245569161_Independent_Component_Analysis_of_EEG_and_Source_Position_Estimation.

Martina Kopčanová, Luke Tait, Thomas Donoghue, George Stothart, Laura Smith, Aimee A Flores, Paula Davila-Perez, Stephanie Buss, Mouhsin M Shafi, Alvaro Pascual-Leone, Peter J Fried, and Christopher S Y Benwell. Resting-state eeg signatures of alzheimer's disease are driven by periodic but not aperiodic changes. *bioRxiv (Cold Spring Harbor Laboratory)*, Jun 2023. doi: https://doi.org/10.1101/2023.06.11.544491. URL https://www.ncbi.nlm.nih.gov/pmc/articles/PMC10312609//#:~:text=A%20large%20body%20of%20evidence,of%20the%20peak%20alpha%20frequency.

Akanksha Parihar and Preety D Swami. Eeg classification of alzheimer's disease, frontotemporal dementia and control normal subjects using..., Aug 2023. URL https://www.researchgate.net/publication/373302163_EEG_Classification_of_Alzheimer's_Disease_Frontotemporal_Dementia_and_Control_Normal_Subjects_using_Supervised_Machine_Learning_Algorithms_on_various_EEG_Frequency_Bands.

Akanksha Parihar, D Preety, and Swami. *EEG Classification of Alzheimer's Disease, Frontotemporal Dementia and Control Normal Subjects using Supervised Machine Learning Algorithms on various EEG Frequency Bands*. URL http://www.ijstm.com/images/short_pdf/1686303248_P-109.pdf.

Khaled Saab, Jared Dunnmon, Christopher Ré, Daniel L Rubin, and Christopher Lee-Messer. Weak supervision as an efficient approach for automated seizure detection in electroencephalography. *npj Digital Medicine*, 3(1), Apr 2020. doi: https://doi.org/10.1038/s41746-020-0264-0. URL https://www.ncbi.nlm.nih.gov/pmc/articles/PMC7170880/.

Seyed-Ali Sadegh-Zadeh, Elham Fakhri, Mahboobe Bahrami, Elnaz Bagheri, Razieh Khamsehashari, Maryam Noroozian, and Amir M Hajiyavand. An approach toward artificial intelligence alzheimer's disease diagnosis using brain signals. *Diagnostics*, 13(3):477–477, Jan 2023. doi: https://doi.org/10.3390/diagnostics13030477. URL https://www.ncbi.nlm.nih.gov/pmc/articles/PMC9913919/.

Mehrnoosh Sadat Safi and Mohammad Mehdi. Early detection of alzheimer's disease from eeg signals using hjorth parameters. *Biomedical Signal Processing and Control*, 65:102338–102338, Mar 2021. doi: https://doi.org/10.1016/j.bspc.2020.102338. URL https://www.sciencedirect.com/science/article/pii/S1746809420304523?via%3Dihub.

Afshin Shoeibi, Delaram Sadeghi, Parisa Moridian, Navid Ghassemi, Jónathan Heras, Roohallah Alizadehsani, Ali Khadem, Yinan Kong, Saeid Nahavandi, Yu-Dong Zhang, and Juan Manuel Gorriz. Automatic diagnosis of schizophrenia in eeg signals using cnn-lstm models. *Frontiers in Neuroinformatics*, 15, Nov 2021. doi: https://doi.org/10.3389/fninf.2021.777977.

Gaowei Xu, Tianhe Ren, Yu Chen, and Wenliang Che. A one-dimensional cnn-lstm model for epileptic seizure recognition using eeg signal analysis. *Frontiers in Neuroscience*, 14, Dec 2020. doi: https://doi.org/10.3389/fnins.2020.578126. URL https://www.frontiersin.org/journals/neuroscience/articles/10.3389/fnins.2020.578126/full.

Yaqing Zhang, Jinling Chen, Jen Hong Tan, Yuxuan Chen, Yunyi Chen, Dihan Li, Lei Yang, Jian Su, Xin Huang, and Wenliang Che. An investigation of deep learning models for eeg-based emotion recognition. *Frontiers in Neuroscience*, 14, Dec 2020. doi: https://doi.org/10.3389/fnins.2020.622759. URL https://www.frontiersin.org/journals/neuroscience/articles/10.3389/fnins.2020.622759/full.