Identification of Alzheimer’s Disease using Deep Learning with the multi-layer perceptron (MLP) Classifier

Georgia Alexandrakis

Abstract

Throughout the past decade, the world cautiously predicts the potential impact of artificial intelligence (AI) on daily life from education to healthcare. This study examined the accuracy and efficiency of a deep learning model when detecting Alzheimer's disease in brain MRIs. Python code was used to create a model that would compute the accuracy and error loss with different hyper parameters. The multi-layer perceptron (MLP) Classifier was used to train the model and the classifier includes multiple hyper parameters that can be fine-tuned to improve the accuracy of the model. Next, we focused on optimizing the values of two hyper parameters: learning rate and hidden layer sizes. The model performed best with a learning rate of 0.0001 and an accuracy of 97.19%. For hidden layer sizes, the number of neurons per layer was optimized to compute the highest accuracy. The model performed best with one layer of 50 neurons with an accuracy of 95.94%. When both hyper parameters were changed in the same experiment, the accuracy decreased to 96.09%. The optimum model with a hidden layer size of two layers of 50 neurons and a learning rate of 0.0001 earned a sensitivity of 95.50% and a specificity of 99.02%. We observed that an accuracy above 90% can only be reached by optimizing with each hyper parameter. This study demonstrates the feasibility of using deep learning modeling with the MLP Classifier for successful identification of Alzheimer’s disease.

Introduction

Alzheimer’s disease is a serious health disorder in ages 65 or older that results in deterioration of thinking, memory, and reasoning (2). As shown in Figure 1, Alzheimer’s disease also known as dementia includes a variety of degrees of severity The disease targets the hippocampus of the brian which is responsible for learning and memory (9). Early stages of Alzheimer's disease causes toxic changes in the brain ultimately resulting in healthy neurons to die. Eventually, the brain shrinks resulting in changed behavior. In later stages of Alzheimers’s disease, more challenges arise when taking care of older individuals.

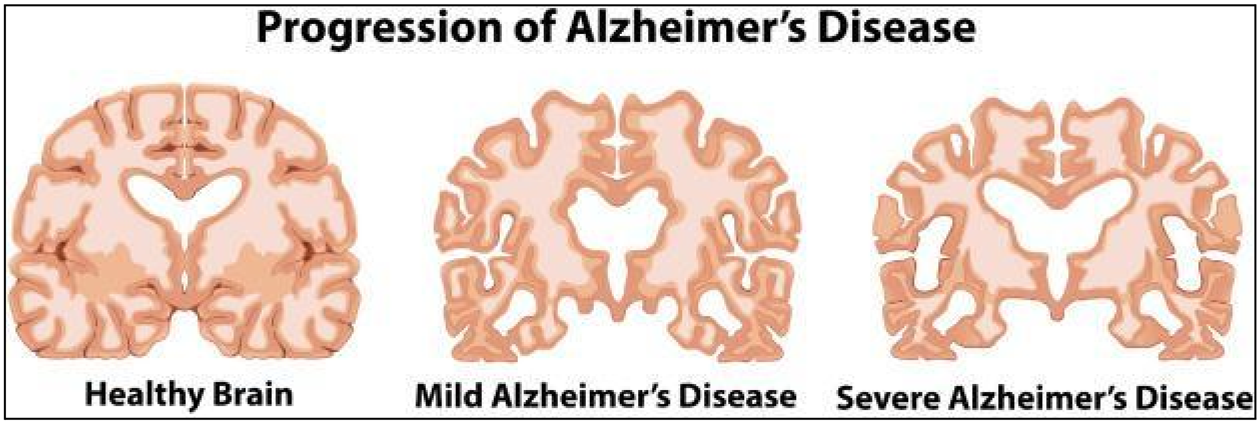

Figure 1. **Overview of different severities of Alzheimer's disease**

When training our model, we assured that it was able to accurately identify and classify the degree of dementia, not only stating if Alzheimer's disease is present. In other words, the model will perform beyond the scope of binary classification (5).

Alzheimer's disease can be detected with an MRI scan. The medical field is increasingly seeing the impact of AI by improvements in healthcare access and automating medical information. Applications of AI in radiology have included improvements to the radiology turnaround time (rTAT) for x-rays, CTs, and particularly Magnetic Resonance Imaging (MRI). Reading and diagnosing radiological images can take several days at best especially for brain MRI even in subspecialized centers. This reveals the question of whether a type of artificial intelligence such as machine learning (ML) and deep learning (DL) is capable of precisely detecting medical abnormalities in the brain. ML is able to solve problems faster than a human and can process large amounts of data. DL is a sub-field of ML that uses neutral networking to train the model. DL extracts features from non-linear hidden layers allowing for its classification to be better than ML classification. Feature extraction and classification are common and necessary steps for the model to successfully output the correct disease from the inputted brain MRI (6). After feature extraction occurs, classifiers are used to label the data. Figure 2 provides examples of the different classifiers that are able to detect brain MRIs. These classifiers include hidden layers, a neural network, and a hierarchical structure to compute an output.

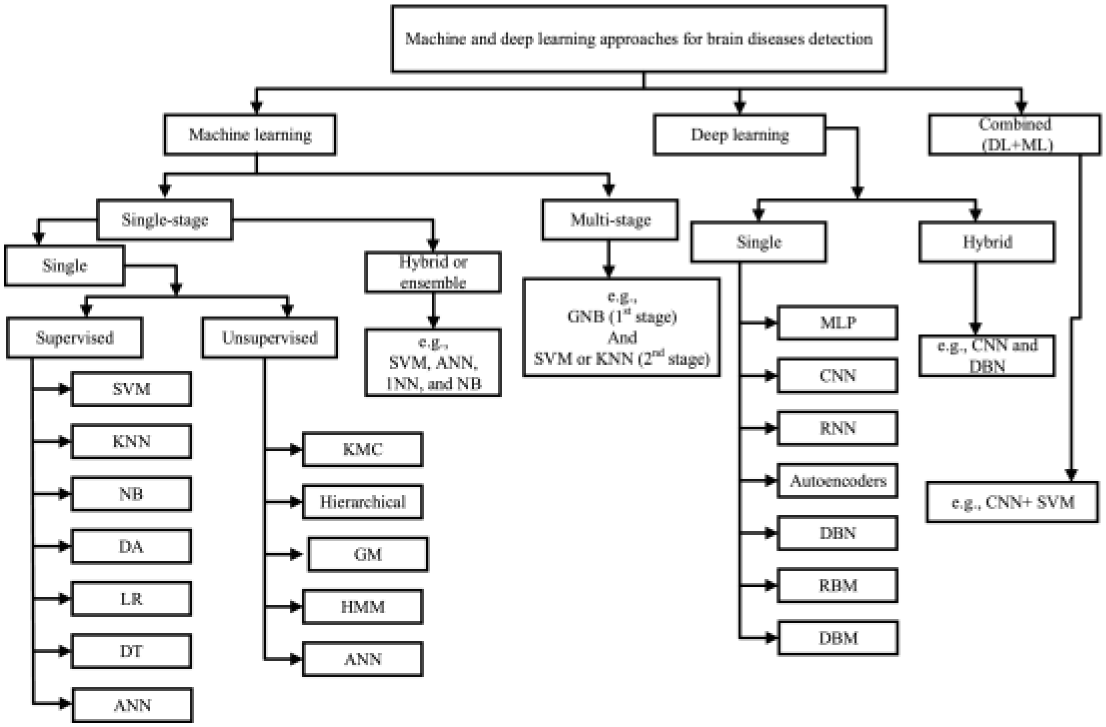

Figure 2. **Classifiers of ML and DL that can detect brain diseases in MRIs** (6)

Recent studies discovered that Convolutional Neural Neural (CNN) is particularly successful in detecting abnormalities such as tumors in brain MRIs. CNN computed an accuracy of 97.2% when detecting tumors in brain MRIs. (1). However, a scientific report by *Nature* concluded that the MLPClassifier is a high performing model that was successful during MRI classification of lymphoma compared to CNN (12). Another study suggested that the MLPClassifier was also more superior when recalling mild cognitive impairment (MCI) in brain MRIs (11). Therefore, in this study, the MLPClassifier, a similar neural network was used to compute the accuracy of detecting Alzheimer's disease in brain MRIs.

In order to compare the accuracy of DL with a human radiologist when detecting Alzheimer disease, the MLPClassifier was optimized with different hyper parameters including hidden layer sizes and learning rate to identify the best performing architecture. Brain MRI scans from a dataset were used to train and set the model. The results of our model were later analyzed to compute its accuracy, sensitivity, and specificity.

Data Processing and Materials

The dataset used in this study was from kaggle and included four categories of the severity of dementia in the brain: mildly demented, moderately demented, non demented, and very mildly demented (3). Figure 3 illustrates the four categories of brain MRIs. In total, the dataset included 3200 images. 80% of those images were used to train the model and while 20% images were used to test the model. Out of 640 tested images, 307 images were non demented, 215 images were very mildly demented, 109 images were mildly demented, and 9 images were moderately demented. The different diagnosis images were stored in separate files allowing us to create a path while training our model. Np.arrays or matrices were created from our dataset to more easily train our model by flattening the images. Additionally, the images and diagnosis were stored separately into two separate np.arrays called inputs and targets. Inputs and targets were processed into variables called inputs_train, inputs_test, targets_train, and targets_test.

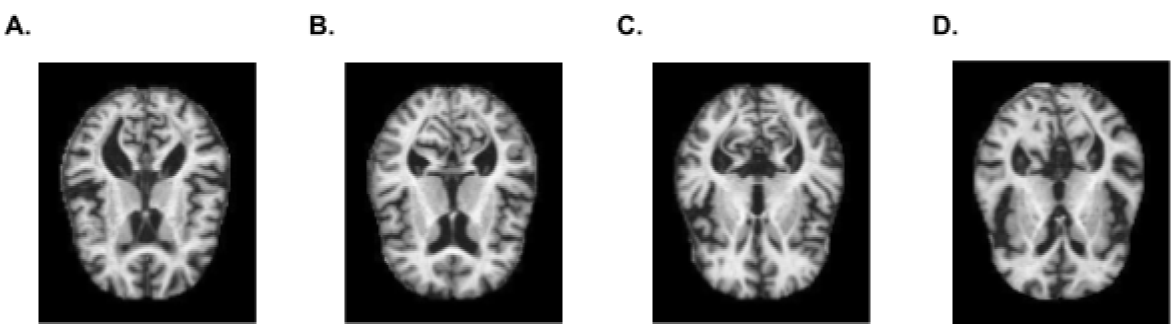

Figure 3. **MRIs with different severities from the dataset that were inputted in the model** (A) Very mild demented brain (B) Non demented brain (C) Moderate demented brain (D) Mild demented brain

Methodology

As shown in Figure 4, the MLPClassifier, a specific neural network that performs classification, was used when training our model. This classifier includes multiple hyperparameters such as hidden_layer_sizes, batch_size, learning_rate_init, max_iter, and verbose to improve the accuracy of the model. In this study, learning_rate_init and hidden_layer_sizes were fine tuned to create the most accurate model. The default for hidden_layers sizes is a single layer with 100 neurons while the default for the learning_rate_init is 0.01. However, after this study, the default may not always provide the most accurate model. The learning rate determines how much the model updates itself after seeing the data. Therefore, a high learning rate will allow for the model to quickly learn the new data. However, if the learning rate is too high, the updates will become unpredictable, preventing good learning. If the learning rate is too low, the model will take much longer to learn.

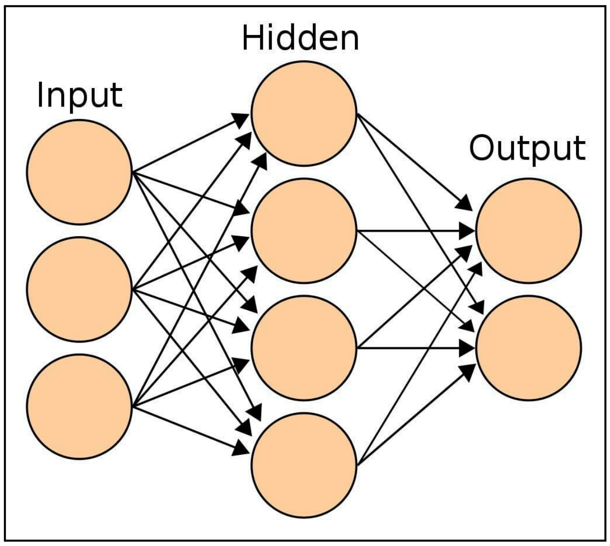

Figure 4. **Visual representation of the MLPClassifier**

Hidden layers make up the MLPclassifier. Adding more hidden layers with different amounts of neurons in each layer can affect the accuracy of the model. In most cases, adding more layers improves the accuracy. However, it may also result in the underfitting of the training data (10).

Results

We first optimized each major component involved in classification of images and performance of the DL model as described below.

A. Hidden Layer Sizes

After processing the data, experiments were created in the same function that processed the data. For each stimulation, all four processed lists containing the images were used to train and test the model along with a specific hyperparameter of the MLPClassifier. The hyperparameter, hidden_layer_sizes, includes a maximum of three layers with various amounts of neurons per layer. For the purpose of this study, the number of neurons per layer kept constant to make it easier to graph and interpret our findings. To compute the accuracy of each experiment, a new function created a list called Preds that could count how many test targets the model correctly classified. The function would then divide preds by the total number of test targets to calculate the accuracy of the model. As shown in Figure 5, the first experiment ran included layers of 100 neurons per layer and the parameter value of hidden_layer_sizes was set to [(100), (100,100), (100,100,100)]. After 80 iterations, the default of one layer with 100 neurons terminated with an accuracy of 80% because the training loss was not improving. Two layers with 100 neurons terminated after 71 iterations with an accuracy of 88% and three layers with 100 neurons terminated after 87 iterations with an accuracy of 90.16%. This was repeated with different parameter values such as layers with 10, 25, 50, and 200 neurons. Two layers of 200 neurons output an accuracy of 75.94% while one layer of 200 neurons performed an accuracy of 90%. However, when the model had 10 neurons per hidden layer, the accuracy increased at a linear rate as more layers were added. Three layers of 10 neurons displayed an accuracy of 90.94% while one layer of 10 neurons of 60.16%. The model performed best with one layer of 50 neurons and an accuracy of 95.94% and performed the worst with one layer of 25 neurons. This demonstrates that the amount of neurons per layer largely contributes to the model's accuracy.

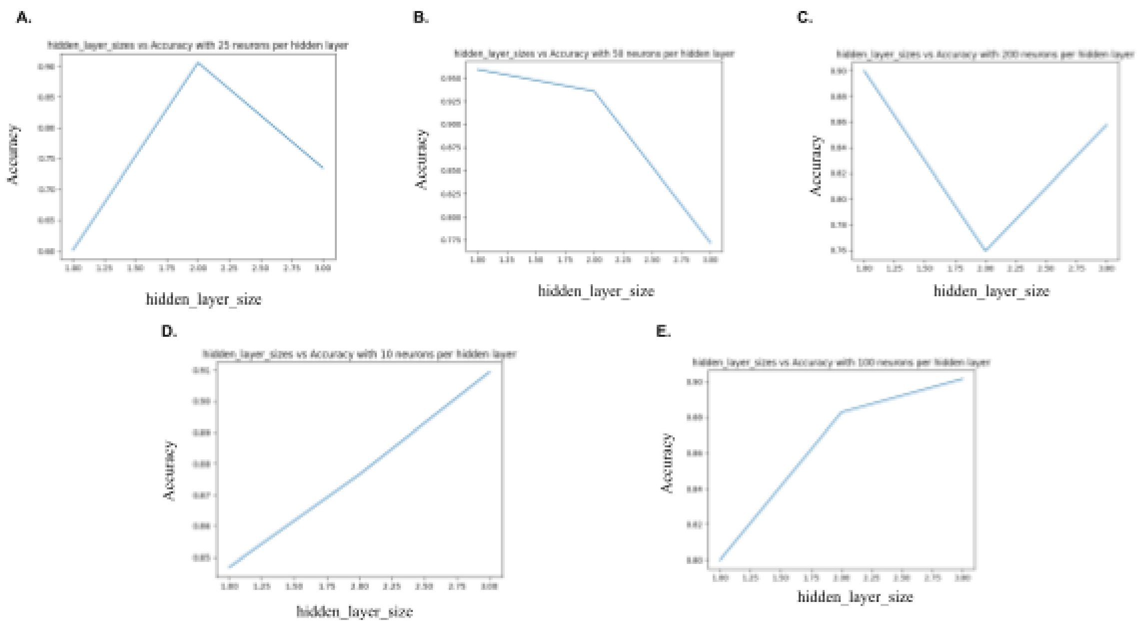

Figure 5. **Line graphs representing hidden layer size and accuracy of the model**
Each graph represents different amounts of neurons per hidden layer in the model. (A) 25 neurons per hidden layer (B) 50 neurons per hidden layer (C) 200 neurons per hidden layer (D) 10 neurons per hidden layer (E) 100 neurons per hidden layer

B. Learning Rate

A separate simulation was run to compute the best learning rate for the model. The hyperparameter, learning_rate_init, was changed based on the array described as [0.0000000001, 0.000000001, 0.00000001, 0.0000001, 0.000001, 0.00001, 0.0001, 0.001, 0.01, 0.1]. As the learning rate increased, the model improved its accuracy from 33.13% to 96.87% until it passed the learning rate of 0.0001. A learning rate greater than 0.0001 resulted in a smaller accuracy. Eventually, the model plateaued at 40.7% when the learning rate was too large. With 200 iterations and the lowest loss rate of 0.0374, Figure 6 illustrates that the model performed best with a learning rate of 0.0001.

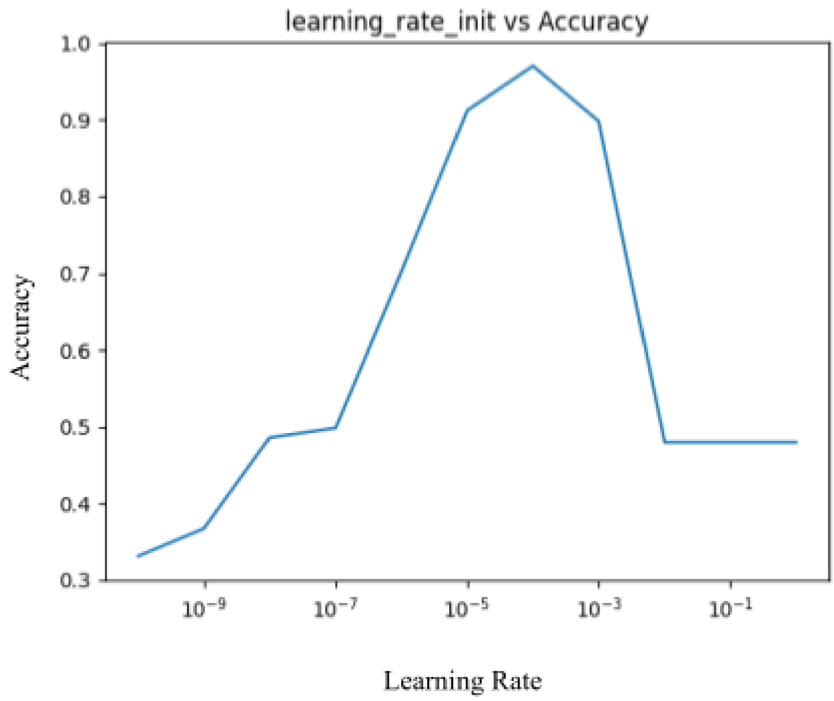

Figure 6. **Line graph representing changes in the accuracy as a function of learning rate**

C. Hidden Layer Sizes and Learning Rate Combined

Results from the previous experiments were combined to improve the model. Since the model performed best with a learning rate of 0.0001 and a hidden layer size of one layer with 50 neurons, these hyperparameter values were combined to compute a new accuracy. However, when combined, the model performed worst with a training accuracy of 99.79% and a test accuracy of 96.09%. When a learning rate of 0.0001 and a hidden layer size of two layers with 50 neurons were combined, the model performed better with a training accuracy of 99.98%, test accuracy of 97.19%, and its lowest loss of 0.0667. A learning rate of 0.0001 and a hidden layer size of two layers of 25 neurons also performed a training state of 99.79% and a test accuracy of 96.09%. After many trials, the model produced the best test accuracy of 97.19% with a learning rate of 0.0001 and a hidden layer size of two layers of 50 neurons.

D. Efficiency of Model

In order to calculate the efficiency and how quickly the model is learning, our study created a graph that depicts the relationship between each iteration and the loss rate. A lower loss rate means that the model is more accurate in diagnosing brain MRIs. The faster the loss rate decreases, the more efficient the model as it takes less time for the model to improve its accuracy. Different hyperparameters values of the learning rate and hidden layer sizes were tested to compute the model with the best efficiency. As shown in Figure 7, it seemed that the model with a learning rate of 0.0001 and a hidden layer size of two layers of 10 neurons decreased its loss rate the fastest. However, after 3 iterations, the loss rate improved slower at a linear rate. Other experiments such as a learning rate of 0.0001 and a hidden layer size of either two layers of 50 or 25 neurons followed a faster exponential decay.

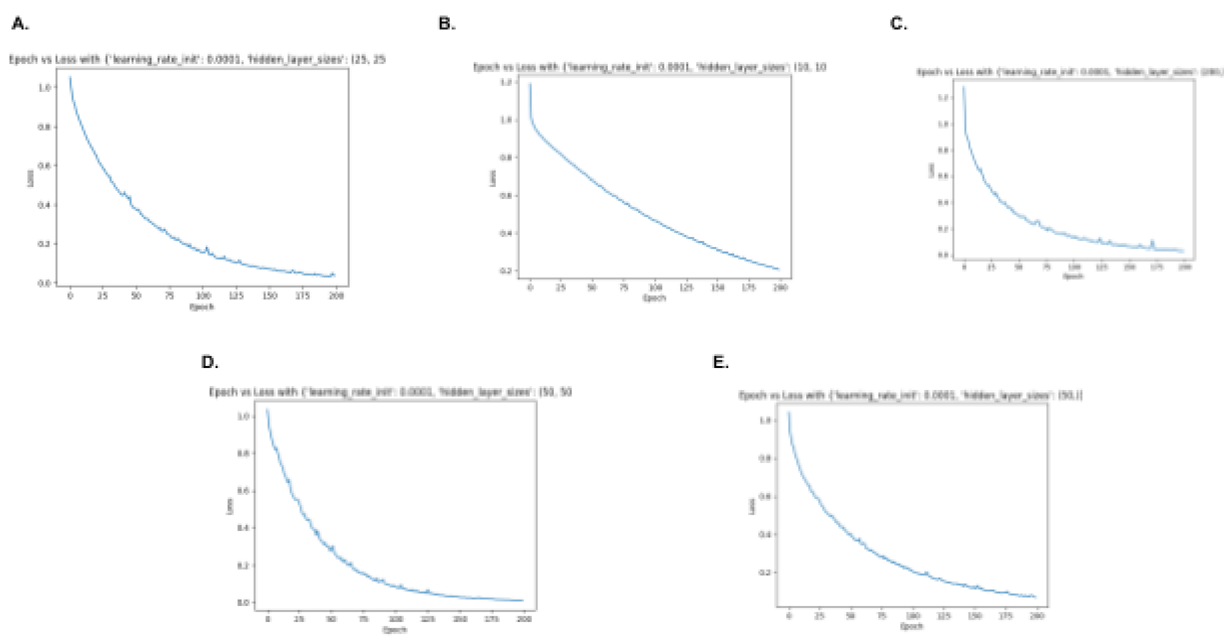

Figure 7. **Line graph representing epoch and loss rate with different hyper parameters**
Each graph has a different number for learning rate, the number of neurons in each hidden layer, and the number of layers. The steeper the slope the more efficient the model. (A) learning rate of 0.0001 with two 25 neuron layers (B) learning rate of 0.0001 with two 10 neuron layers (C) learning rate of 0.0001 with one 200 neuron layer (D) learning rate of 0.0001 with two 50 neuron layers (E) learning rate of 0.0001 with one 50 neuron layer

E. Calculations

As shown in Figure 8, the results from our experiments were later displayed in multiple confusion matrices to analyze which categories of Alzheimer severity our model performed the best and the weakest. These matrices would help identify whether our model is stating any false positives or false negatives. When the learning rate was 0.0001 and a hidden layer size was two layers of 50 neurons, the model reached its highest sensitivity of 95.50% and a specificity of 99.02%. Table 1 depicts the sensitivity and specificity for the four trials. Our data and calculations suggests that the model is more likely to detect true negatives than true positives. Therefore, the model again performed the best with a learning rate of 0.0001 and a hidden layer size of two layers of 50 neurons.

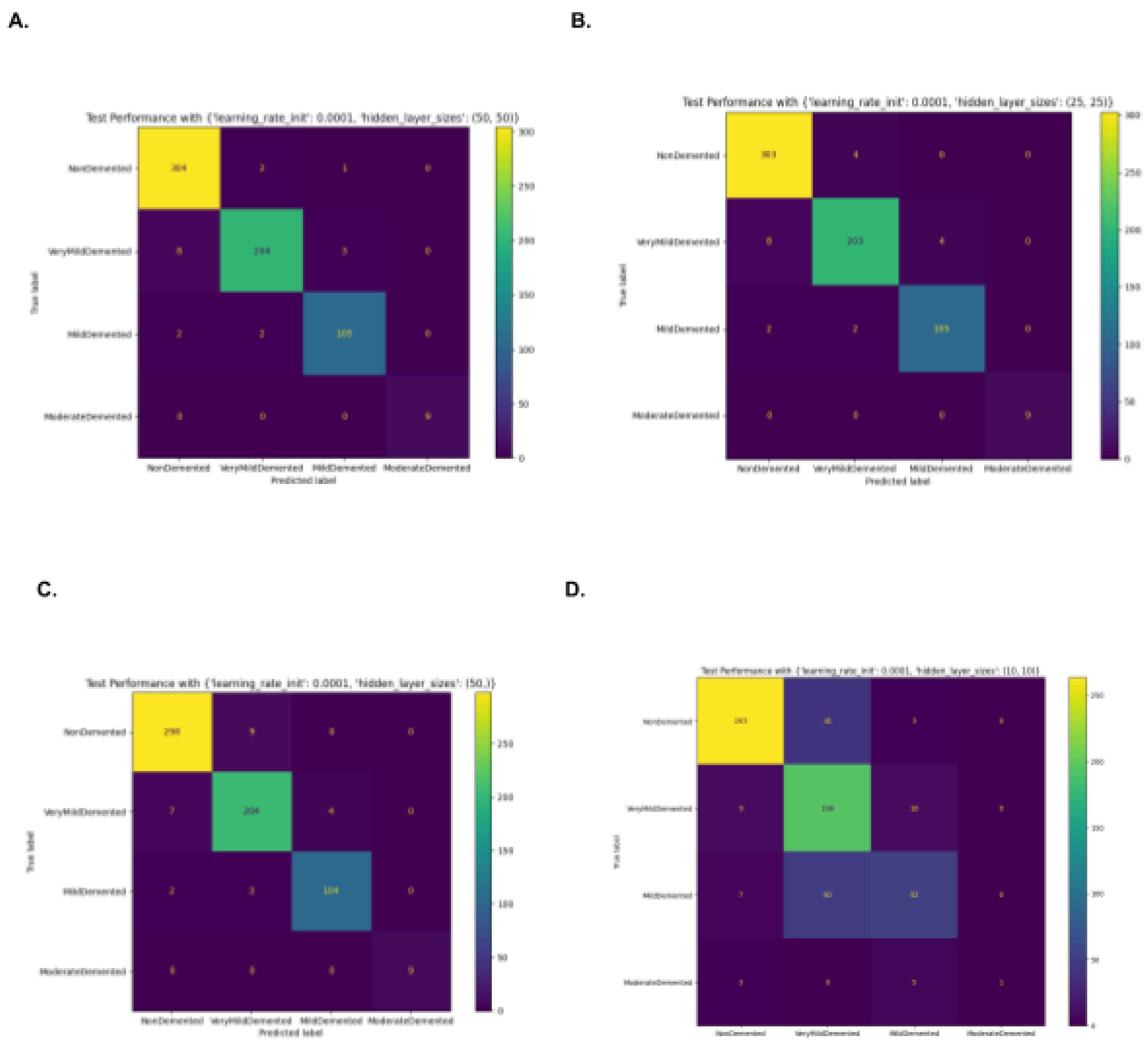

Figure 8. **Confusion matrices with different hidden layer sizes**

The confusion matrices represent the final results. (A) When hidden layer sizes had two layers of 50 neurons, the model had an accuracy of 99.02% for non demented, 95.33% for very mildly demented, 96.33% for mildly demented, and 100% for moderately demented. (B) When hidden layer sizes had two layers of 25 neurons, the model had an accuracy of 98.70% for non demented, 94.42% for very mildly demented, 96.33% for mildly demented, and 100% for moderately demented. (C) When hidden layer sizes had one layer of 50 neurons, the model had an accuracy of 97.07% for non demented, 94.88% for very mildly demented, 95.41% for mildly demented, and 100% for moderately demented. (D) When hidden layer sizes had two layers of

10 neurons, the model had an accuracy of 86.97% for non demented, 91.16% for very mildly demented, 47.71% for mildly demented, and 11.11% for moderately demented.

Hidden layer size	Sensitivity	Specificity
Two layers with 50 neurons	95.50%	99.02%
Two layers with 25 neurons	95.20%	98.70%
One layer of 50 neurons	95.20%	97.07%
Two layers of 10 neurons	74.77%	85.67%

Table 1. **Table of the sensitivity and specificity vs hidden layer size**
The table reveals the data calculated from each of the four trials with the same learning rate of 0.0001 and different hidden layer size values

Discussion

One limitation in the model was its small sampling size. The model was trained with only 3200 images most of which were patients without dementia. There was a small majority of patients with moderate dementia making it difficult for the model to improve its accuracy in this group. Another limitation in the experiment was keeping the number of neurons per layer constant. Changing the number of neurons per layer would result in many possible inputs and a larger experiment. More inputs would contribute to the possibility of creating a model with a greater accuracy. However, different neurons per layer can result in difficulty graphing and comparing the data. Additionally, only two of the parameters of the MLPClassifier were changed throughout the experiment. The MLPClassifer has numerous hyper parameters and changing the values of other hyper parameters could result in a more accurate model.

Other studies conducted the accuracy of AI classifying tumors and the possibility of a stroke in brain MRIS. The National Scientific Ethical Board of Denmark revealed that their AI model earned a sensitivity of 89% and a specificity of 90% when locating areas of ischemia on brain MRI scans (8). Their study found that the AI sensitivity declines with smaller lesions. Similarly, in our study, the accuracy of the model declined when the brain MRIs were less demented. Both studies suggest that there is still a small percentage that AI incorrectly detects abnormalities in brain MRIs. Furthermore, *Nature*, a peer-reviewed journal, found that AI was capable of detecting Alzeihmer's disease of more than a 90% accuracy (7). Their findings verify and provide more confidence with our results.

Conclusion

After combinations of the hyperparameters, the model performed the best with a learning rate of 0.0001 and a hidden layer size of two layers with 50 neurons. In total, the model correctly identified 606 images out of 640. The model's strongest category was moderately demented and its weakest category was mildly demented. This weakness would limit the model's utility in detecting the earliest stages of the disease. However, it is difficult to compare the accuracy of each category because the categories do not have the same number of images. Since the non demented category had 307 images, more incorrectly answered questions would result in a higher accuracy. Even though our model performed its best with an accuracy of 97.19%, there is a 2.81% chance of AI making an error in diagnosing Alzheimer's in brain MRIs. However, according to the Molecular Psychiatry, scientists conducted a similar experiment to compute the accuracy of human radiologists diagnosing Alzheimer's in brain MRIs. Their findings suggested a 90% sensitivity and 84% specificity (4). These results demonstrate that our model outperformed the accuracy, sensitivity, and specificity of a human radiologist when detecting Alzheimer's in brain MRIs suggesting the possibility for artificial intelligence replacing the radiologists that specialize in the brain. More studies are still needed to compute the accuracy of artificial intelligence when detecting abnormalities in other areas including joint MRIs and total body CT scans.

References

1. Abdusalomov, Akmalbek Bobomirzaevich, et al. "Brain Tumor Detection Based on Deep Learning Approaches and Magnetic Resonance Imaging." *Cancers*, vol. 15, no. 16, Multidisciplinary Digital Publishing Institute, Aug. 2023, pp. 4172–72, https://doi.org/10.3390/cancers15164172.

2. Alzheimer's Association. “What Is Alzheimer’s Disease?” *Alzheimer’s Association*, 2025, www.alz.org/alzheimers-dementia/what-is-alzheimers.
3. “Augmented Alzheimer MRI Dataset.” *www.kaggle.com*, www.kaggle.com/datasets/uraninjo/augmented-alzheimer-mri-dataset.
4. Chouliaras, Leonidas and O’Brien, John T. “The Use of Neuroimaging Techniques in the Early and Differential Diagnosis of Dementia.” Molecular Psychiatry, vol. 28, 22 Aug. 2023, https://doi.org/10.1038/s41380-023-02215-8.
5. Dunlap, Garrett. “Is Alzheimer’s Disease Genetic? | Decode Your DNA-Learn about Your Risks!” *Nebula Genomics Blog*, May 2022, nebula.org/blog/is-alzheimers-disease-genetic/.
6. Khan, Protima. “Machine Learning and Deep Learning Approaches for Brain Disease Diagnosis: Principles and Recent Advances.” *Ieeexplore.ieee.org*, ieeexplore.ieee.org/abstract/document/9363896.
7. Kozlov, Max. “AI That Reads Brain Scans Shows Promise for Finding Alzheimer’s Genes.” *Nature*, Nov. 2023, https://doi.org/10.1038/d41586-023-03482-9.
8. Krag, Christian H., et al. “Diagnostic Test Accuracy Study of a Commercially Available Deep Learning Algorithm for Ischemic Lesion Detection on Brain MRIs in Suspected Stroke Patients from a Non-Comprehensive Stroke Center.” *European Journal of Radiology*, vol. 168, Nov. 2023, p. 111126, https://doi.org/10.1016/j.ejrad.2023.111126.
9. National Institute on Aging. “Alzheimer’s Disease Fact Sheet.” National Institute on Aging, National Institutes of Health, 5 Apr. 2023, www.nia.nih.gov/health/alzheimers-and-dementia/alzheimers-disease-fact-sheet.
10. Solanki, Sunny. “Scikit-Learn - Neural Network.” *Coderzcolumn.com*, CoderzColumn, 2020, coderzcolumn.com/tutorials/machine-learning/scikit-learn-sklearn-neural-network.
11. Tibon, Roni. “A Machine Learning Approach for MRI-Based Classification of Individuals with Mild Cognitive Impairment.” *The Preprint Server of Biology*, 28 May 2021, https://doi.org/10.1101/2021.05.27.445930.
12. Yun, Jihye, et al. “Radiomic Features and Multilayer Perceptron Network Classifier: A Robust MRI Classification Strategy for Distinguishing Glioblastoma from Primary Central Nervous System Lymphoma.” *Scientific Reports*, vol. 9, no. 1, Apr. 2019, https://doi.org/10.1038/s41598-019-42276-w.