Utilizing AI for Alzheimer's Disease Classification

Trinity McKenny
trinity.mckenny@gmail.com

Abstract

Alzheimer's is a serious disease and its early diagnosis is essential to properly understand the disease, access information, and provide sufficient care and support [1]. This study hopes to assess the potential of AI to help early diagnoses. When properly used, AI could potentially diagnose Alzheimer's with more efficiency and accuracy. Many classifier models were used: the MLP classifier had the least accuracy of 32%, the K Nearest Neighbors classifier had an accuracy of 60%, the Decision Tree Model had an accuracy of 72%, the Logistic Regression model had an accuracy of 83%, and the Random Forest classifier did the best with a 91.5% accuracy. There are many factors that the AI algorithm takes into consideration when making the diagnosis, and each factor is weighed differently. Factors like hand dominance had a smaller weight in the diagnosis than the mini mental state evaluation, and all of the analysis is covered in the Logistic Regression Feature Importance to shed light on the factor importance.

1. Introduction

Alzheimer's disease is a form of dementia and accounts for 60-80% of all dementia cases [3]. It can be identified with shrinkage of the brain area, and the use of AI algorithms can potentially make this diagnosis faster and more effective. An early diagnosis can be beneficial for many reasons. First, it can allow the family to better prepare for the future. Legal documents can be updated, care preferences discussed, safety discussed, etc. Second, it can allow the individual to possibly make some lifestyle changes that can preserve cognitive function, like changing diet or increasing exercise. Third, it can save costs for the medical and long-term care for the individual [2]. Utilizing AI for an early diagnosis can help the individual and their family mentally, emotionally, and financially.

2. Background

Much research has been done to try and find a cause and cure for this disease, but our aim is to try and increase the accuracy and precision of the diagnosis. Especially with the new potential treatments coming into effect, false positive patients should not be taking the new medicine, and false negative patients should not be delaying their treatment.

The presence of Alzheimer's can be characterized by shrinkage in the cortex, shrinkage of the hippocampus, and enlarged ventricles (see Figure 1). AI is the perfect tool for this data classification because it can make expedited diagnoses without the risk of human error [12].

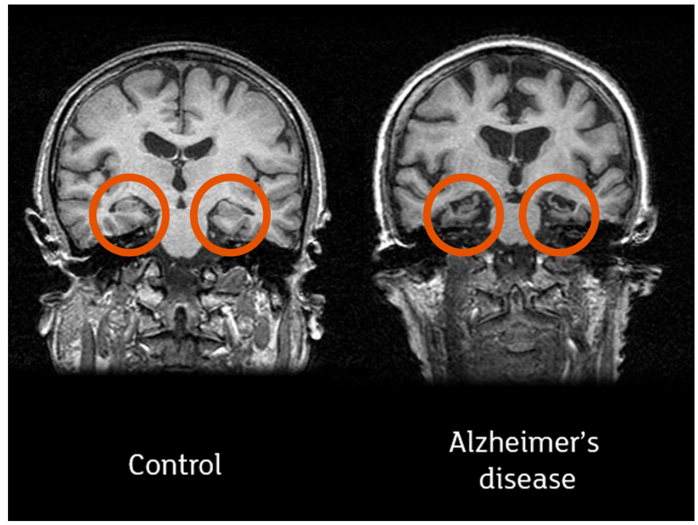

Figure 1. Healthy Brain v. Brain Affected By Alzheimer's MRI Scan Notice how the brain affected by Alzheimer's has shrinkage in the cortex, shrinkage of the hippocampus, and enlarged ventricles [10] [11] .

3. Dataset

The Open Access Series of Imaging Studies (OASIS) is a project aimed at making MRI data sets readily available to the scientific community. It compiles and distributes MRI data sets and hopes to facilitate future discoveries in neuroscience. OASIS is made available by the Washington University Alzheimer's Disease Research Center, Dr. Randy Buckner at the Howard Hughes Medical Institute, the Neuroinformatics Research Group at Washington University School of Medicine, and the Biomedical Informatics Research Network.

This data set consists of a cross-sectional collection of 416 subjects aged 18-96. For each subject, 3 or 4 individual T1-weighted MRI scans obtained in single scan sessions are included. The data set includes both male and female subjects. 100 of the subjects over the age of 60 have been clinically diagnosed with Alzheimer's disease, and an additional 20 nondemented subjects have been included as a control group.

The values given are sex, hand dominance, age, education level, SES (socioeconomic status), MMSE (mini-mental state evaluation), CDR (clinical dementia rating), eTIV (intracranial volume), nWBV (normalized whole brain volume), ASF (atlas scaling factor), and delay. MMSE is ranked on a scale of 1-30, where a value of 23 or lower indicates dementia [6]. CDR is ranked on a scale of 0.0-2.0, where any value over 0.0 indicates dementia. eTIV represents the sum of brain, ventricular, and extraventricular CSF, averaging 1469 $\pm$ 102 cm^3 in men and 1289 $\pm$ 111 cm3 in women [8]. ASF is ranked from 0.88-1.56, allow-

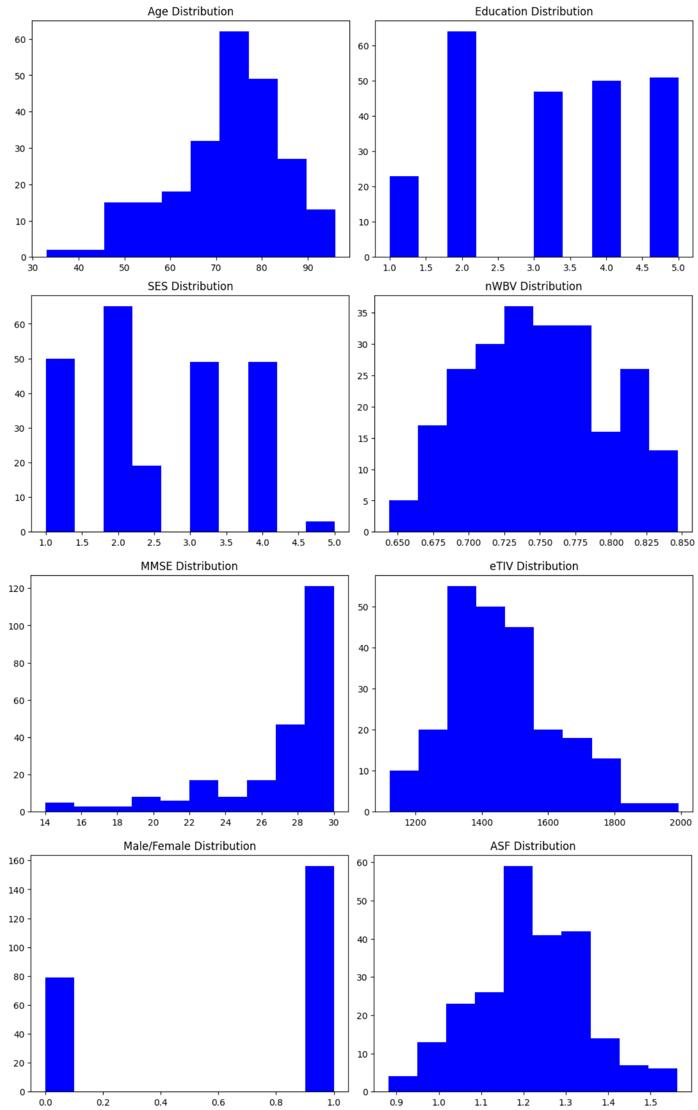

Figure 2. Feature Distributions

ing the comparison of eTIV to total body anatomy [5]. Delay is the number of days since the last visit to the doctor. Although there is no known cure yet, identifying Alzheimer's in the early stages may allow the patient to delay its onset or even fully cure the disease in the future.

4. Methodology

4.1. Performance Metrics

Accuracy is defined as the amount of correctness with the following equation [9]:

$$\frac{TP+TN}{FP+FN+TP+TN}$$

Precision is the fraction of relevant instances among the retrieved instances defined as follows [9]:

$$\frac{TP}{FP+TP}$$

Recall measures how often the machine learning model correctly identifies the true positives from all positive samples in the data defined as follows [9]:

$$\frac{TP}{FN+TP}$$

F1 combines the precision and recall values to achieve a balance with the formula [9]:

$$\frac{2TP}{2TP+FP+FN}$$

Finally, the ROC curve visualizes the trade-off between true positive and false positive rates at various thresholds.

4.2. Data Pre-Processing

To ensure compatibility with machine learning libraries, we pre-processed the data set. We made male (0) and female (1) and right (1) and left (0) handed binary. We also replaced any null values with the average of that section. This does not change or skew the data in any way since it is replaced with the mean. All null CDR values were dropped, and any CDR value of 0.5 was changed into 1.0 to keep demented and non-demented distinct. CDR was then removed from the data in order to test the algorithm. Delay ended up being dropped after all null CDR values were removed because all of the remaining delay values were null. The data was then split 80% train to 20% test. A random state of 42 was used to allow the testing data to be shuffled consistently.

4.3. Machine Learning Models

- A Decision Tree model was created first, and it resembles a tree. The base of the tree is the root node. From the root node flows a series of decision nodes that depict decisions to be made. From the decision nodes are leaf nodes that represent the consequences of those decisions. Each decision node represents a question or split point, and the leaf nodes that stem from a decision node represent the possible answers. How we determine the rules for each node is done by recursively splitting the nodes based on loss.
  Hyperparameters: Random state = 1; Max depth = 5

- A Logistic Regression model was created next, and it is a data analysis technique that uses mathematics to find the relationships between two data factors. It then uses this relationship to predict the value of one of those factors based on the other. The prediction usually has a finite number of outcomes, like yes or no.
  Hyperparameters: Random state = 0; Max-Iterations = 1000000

- A K Nearest Neighbors model operates by finding the k nearest neighbors to a given data point, and it takes the majority vote to classify the data point. The value of k is crucial, and one needs to choose it wisely to prevent overfitting or underfitting the model.
  Hyperparameters: Nearest Neighbors = 9

- A Multilayer Perceptron model is a neural network made by combining several neurons. They are organized into three layers: one input layer, which distributes the unit features to the first hidden layer, some number of hidden layers, and an output layer that corresponds to the desired model response.

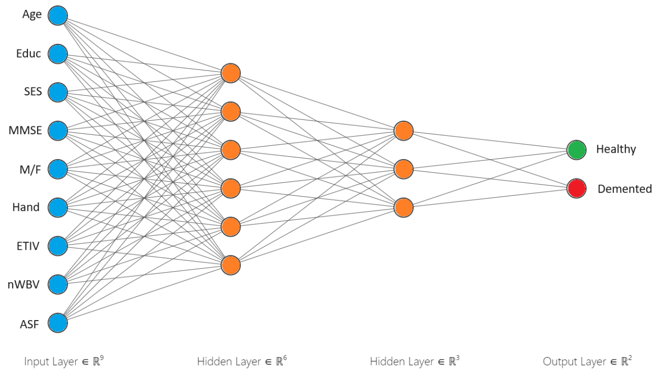

Figure 3. Architectural set up for the MLP Model.

- A Random Forest model grows multiple decision trees which are merged together for a more accurate prediction. The Random Forest model contains multiple uncorrelated models (the individual decision trees) that are able to perform more efficiently and effectively as a group rather than as individual branches.
  Hyperparameters: Random State = 2

The Logistic Regression model and the K Nearest Neighbor models were used as a baseline, while the Multilayer Perceptron model and the Random Forest model were the main models that were analyzed.

5. Results

5.1. Logistic Regression Feature Importance

Each factor that is imputed is weighted in order to determine the best diagnosis for each individual. A greater positive value means that the increased value of the factor in the actual simulation causes a greater chance for the individual to develop Alzheiemer's. A greater negative value means that the increased value of the factor causes a greater chance for the individual to not develop Alzheimer's. Lastly, a value close to zero means that the presence or absence of the factor will likely have little to no effect on the Alzheimer's diagnosis.

Table 1. Weights of Factors

Factor	Weight
Age	0.21
Educ.	0.18
SES	-0.073
MMSE	-1.1
eTIV	0.0015
nWBV	-0.00079
ASF	-0.0080
M/F	-0.047
Hand	-0.0026

Age seems to be a big factor that is present in individuals who develop Alzheimer's due to the relatively high positive weight. This makes sense since dementia usually develops more in older individuals than younger ones. Education seems to be a big factor, and this also could make sense. Younger individuals, including 18-year olds, were present in the study but these individuals

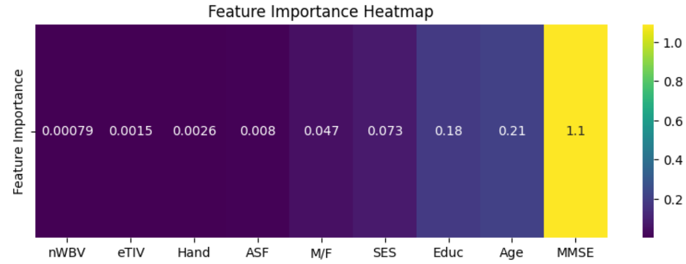

Figure 4. Absolute Values of the Logistic Regression Weights for Relative Feature Importance.

probably have not graduated college or possibly even high school. Older adults have a greater tendency to develop Alzheimer's and they would have graduated from these higher educations. Socioeconomic status is a small factor for an individual's Alzheimer's diagnosis. Someone with a higher socioeconomic status is less likely to develop Alzheimer's, and this is probably because these individuals have more money to spend on their physical and mental health. The mini mental state evaluation is a very big factor in determining Alzheimer's diagnosis. Since this is a professional medical evaluation, it makes sense that the two results line up. The intracranial volume, normalized whole brain volume, atlas scaling factor, sex, and hand dominance do have a weight in the diagnosis, although it is very small.

5.2. Model Metrics and Performance

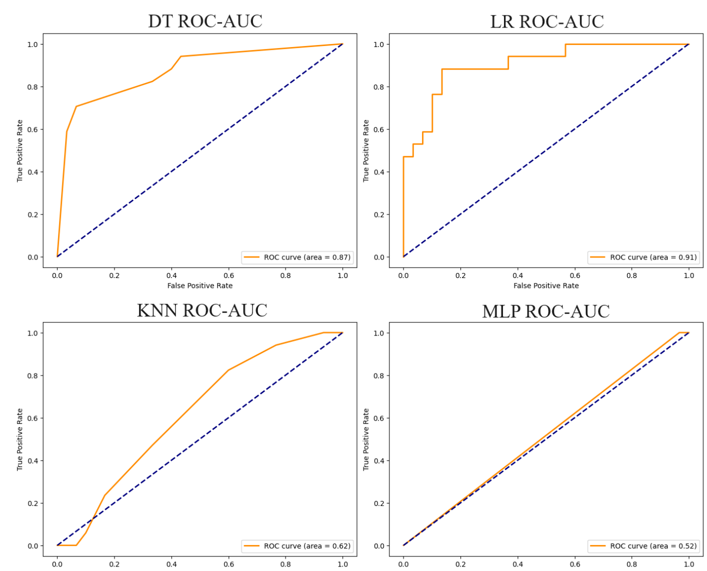

Figure 5. ROC Curve Models

Table 2. Model Results

Model	Acc.	Prec.	Recall	F1	ROC-AUC
K-Nearest Neighbors	0.596	0.470	0.444	0.457	0.620
Decision Tree	0.720	0.820	0.580	0.680	0.870
Logistic Regression	0.830	0.765	0.765	0.765	0.910
Multi-Layer Perceptron	0.362	1.000	0.362	0.531	0.430
Random Forest Generator	0.915	0.882	0.882	0.882	0.940

Looking at all of the tested models overall, the Random For-

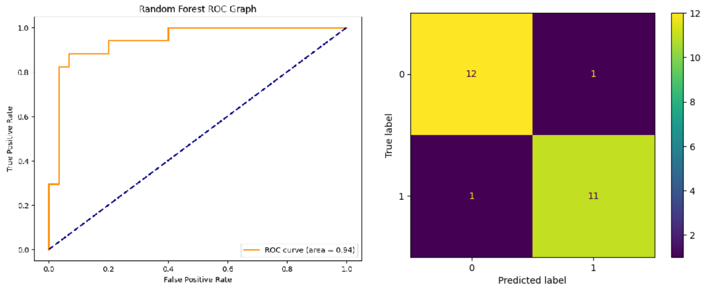

Figure 6. Model Metrics for Random Forest Classifier (Best Model)

est Classifier was able to determine Alzheimer's disease with the greatest accuracy. It was able to perform with 91.5% accuracy in addition to the 88.2% precision, 88.2% recall, 88.2% f1, and had an ROC curve of 94% (Figure 5). In addition, the Random Forest Classifier was plotted in a Confusion Matrix (Figure 6) and only incorrectly categorized two out of twenty-five patients with one false positive and one false negative [7].

The MLP Classifier performed with the lowest accuracy for many reasons. The MLP Classifiers form a dense network of nodes which are susceptible towards redundancy, inefficiency, and overfitting. The MLP Classifiers also do not provide the most accuracy during image processing because spatial information is lost when the image is changed into an MLP [4].

The Random Forest Classifier performed with the highest accuracy for many reasons. The classifier is very intricate and complex, allowing it to be accurate and precise. The Random Forest model is known to prevent overfitting of the data and be robust against outliers, and it can easily perform with missing data and null values.

6. Conclusion

This finding underscores the potential of machine learning algorithms in assisting medical professionals in early diagnosis and intervention strategies for Alzheimer's patients. The high accuracy achieved by the Random Forest Classifier suggests its robustness in discerning intricate patterns within MRI data indicative of Alzheimer's pathology. Although the accuracy percentage was high, there is still a lot of room for error. The model at this point is not good enough for definitive diagnoses, but it does offer a possible tool for enhancing diagnostic accuracy and patient care in clinical settings.

In the future, this study could be enhanced if the program was exposed to more patient data. This can be achieved with regular check-ups with the patients for a set amount of time and an increased patient data pool using different patients from different hospitals.

7. Acknowledgements

The MRI data set available was made available by the OASIS project. Guidance was made available by the Inspirit 1:1 mentorship program.

8. Appendix

References

[1] Alzheimer's Association. Early detection and diagnosis. https://www.alz.org/professionals/public-health/public-health-topics/early-detection-diagnosis. [Online; accessed 20-June-2024]. 1

[2] Alzheimer's Association. Why get checked. https://www.alz.org/alzheimers-dementia/diagnosis/why-get-checked. [Online; accessed 20-June-2024]. 1

[3] CDC. About dementia. https://www.cdc.gov/aging/dementia/index.html, 2019. [Online; accessed 20-June-2024]. 1

[4] Dinesh. Cnn vs mlp for image classification. https://medium.com/analytics-vidhya/cnn-convolutional-neural-network-8d0a292b4498, 2019. [Online; accessed 20-June-2024]. 4

[5] Lawrence V Fulton, Diane Dolezel, Jordan Harrop, Yan Yan, and Christopher P Fulton. Classification of alzheimer's disease with and without imagery using gradient boosted machines and resnet-50. *Brain sciences*, 9(9):212, 2019. 2

[6] Shea Gluhm, Jody Goldstein, Kiet Loc, Alexandra Colt, Charles Van Liew, and Jody Corey-Bloom. Cognitive performance on the mini-mental state examination and the montreal cognitive assessment across the healthy adult lifespan. *Cognitive and Behavioral Neurology*, 26(1):1–5, 2013. 1

[7] Scikit Learn. Decision trees. https://scikit-learn.org/stable/modules/tree.html. [Online; accessed 20-June-2024]. 4

[8] Mitsunori Matsumae, Ron Kikinis, István A Mórocz, Antonio V Lorenzo, Tamás Sándor, Marilyn S Albert, Peter McL Black, and Ferenc A Jolesz. Age-related changes in intracranial compartment volumes in normal adults assessed by magnetic resonance imaging. *Journal of neurosurgery*, 84(6):982–991, 1996. 1

[9] F. Pedregosa, G. Varoquaux, A. Gramfort, V. Michel, B. Thirion, O. Grisel, M. Blondel, P. Prettenhofer, R. Weiss, V. Dubourg, J. Vanderplas, A. Passos, D. Cournapeau, M. Brucher, M. Perrot, and E. Duchesnay. Scikit-learn: Machine learning in Python. *Journal of Machine Learning Research*, 12:2825–2830, 2011. 2

[10] Li Su, Andrew M Blamire, Rosie Watson, Jiabao He, Benjamin Aribisala, John Brien, et al. Cortical and subcortical changes in alzheimer's disease: a longitudinal and quantitative mri study. *Current Alzheimer Research*, 13(5):534–544, 2016. 1

[11] Emma Taylor. All you need to know about brain scans and dementia. https://www.alzheimersresearchuk.org/news/all-you-need-to-know-about-brain-scans-and-dementia/, 2022. [Online; accessed 20-June-2024]. 1

[12] Liz Ticong. Mastering ai data classification: Ultimate guide. https://www.datamation.com/big-data/ai-data-classification/, 2024. 1

Code Base Maintained by Trinity McKenny: https://colab.research.google.com/drive/18Knzpm6CffDVAm07E11TkcVmmCBVXqed?usp=sharing