Integrating Brain Imaging for Outcome Prediction in Alzheimer's Disease

Jeffrey Qu (SEAS 2023)[1,2,3], Mansu Kim[1,2], Kefei Liu[1,2], Dr. Li Shen[1,2]

Department of Biostatistics, Epidemiology, and Informatics, University of Pennsylvania Perelman School of Medicine[1]; Institute for Biomedical Informatics[2]; University of Pennsylvania School of Engineering and Applied Sciences[3]

Introduction

Background Information

- Alzheimer's is a progressive disease that causes memory loss and interferes with cognitive function.
- Extensive research has been conducted to identify causes and possible treatments.
- One particular area of focus is early detection of the disease, from which brain imaging has emerged as a powerful diagnostic tool.

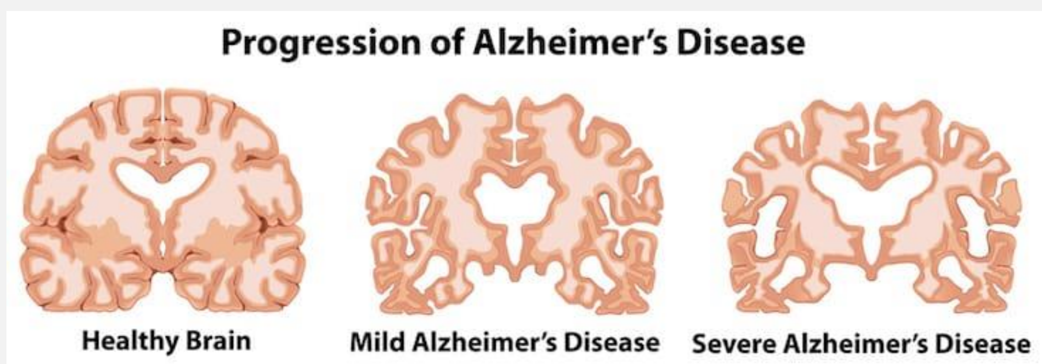


Figure 1. *Side-by-side comparison illustrating shrinking of brain tissue as a result of Alzheimer's.*

Objectives

- Develop and implement machine learning models and apply those to the analysis of brain imaging and outcome data from landmark Alzheimer's disease studies.

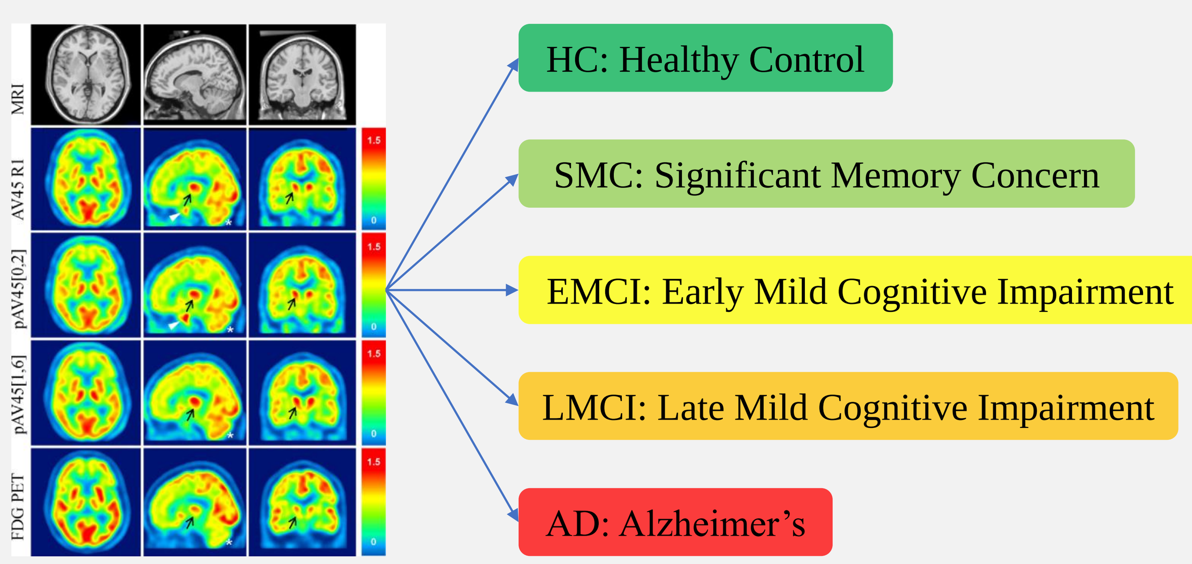


Figure 2. *Labels used for classification, corresponding to different stages of Alzheimer's.*

Dataset

- Data collected by the Alzheimer's Disease Neuroimaging Initiative (ADNI) was utilized, with the following breakdown:
  - **805** samples total:
    - **196 HC** (healthy controls)
    - **78 SMC** (significant memory concern)
    - **235 EMCI** (early mild cognitive impairment)
    - **162 LMCI** (late mild cognitive impairment)
    - **134 AD** (Alzheimer's disease)
  - **3** imaging modalities for each sample:
    - **VBM** (voxel-based morphometry)
    - **FDG** (fluorodeoxyglucose positron emission tomography)
    - **AV45** (florbetapir-fluorine-18 positron emission tomography)
  - **116** measurements taken by each imaging modality, corresponding to different brain regions: amygdala, angular gyrus, calcarine sulcus, cerebellar vermis, etc.

Machine Learning Models

- The following support vector machines (SVMs), ensemble classifiers, and deep neural networks (DNNs) were utilized:
  - Support vector machines (provided by scikit-learn):
    - **SVC** (Support Vector Classification)
    - **LinearSVC** (Linear Support Vector Classification)
    - **NuSVC** (Nu-Support Vector Classification)
  - Ensemble classifiers (provided by scikit-learn):
    - **Voting** (Soft Voting/Majority Rule Classifier)
  - Deep neural networks (constructed using tensorflow):
    - **Model 1**: 2 hidden layers, 64 nodes
    - **Model 2**: 1 hidden layer, 10 nodes
    - **Model 3**: 4 hidden layers, 512 nodes
    - **Model 4**: 4 hidden layers, 100 nodes

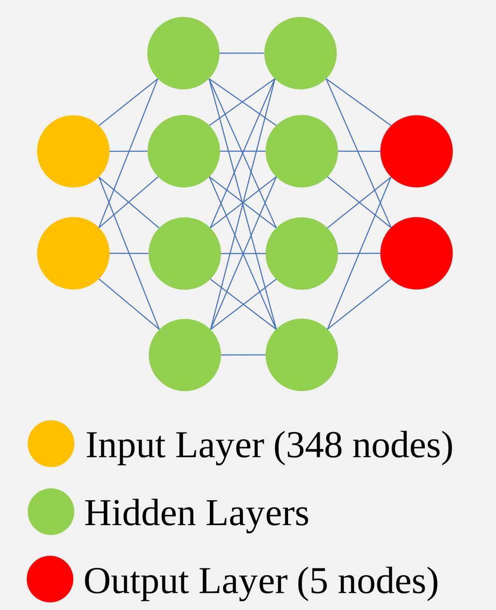
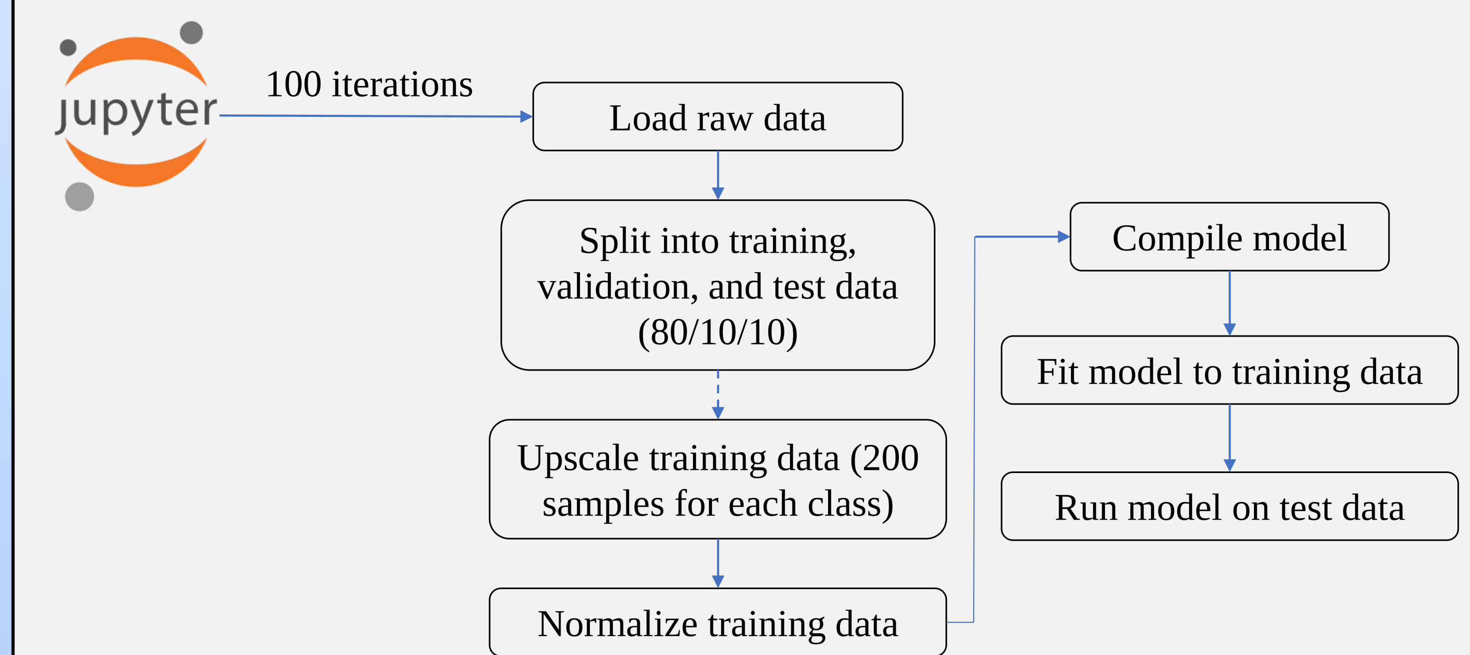


Figure 3. *Deep neural network structure.*

Processing Pipeline



Results

	Overall Acc	HC Sens	HC Spec	HC Acc	SMC Sens	SMC Spec	SMC Acc	EMCI Sens	EMC Spec	EMCI Acc	LMCI Sens	LMCI Spec	AD Sens	AD Spec	AD Acc
LinearSVC	0.63396	0.44472	0.7894	0.61706	0.33202	0.86718	0.5996	0.34151	0.83973	0.59062	0.252	0.86436	0.71567	0.89305	0.80436
Voting	0.63363	0.43277	0.79728	0.61502	0.3386	0.86092	0.59976	0.34446	0.83467	0.58957	0.25814	0.86003	0.7097	0.8997	0.8047
SVC	0.63221	0.43028	0.79169	0.61099	0.38735	0.84702	0.61719	0.31319	0.84548	0.57934	0.20713	0.87286	0.74105	0.8861	0.81357
NuSVC	0.62857	0.41036	0.7912	0.60078	0.35705	0.871	0.61402	0.31446	0.83589	0.57517	0.2225	0.86467	0.74627	0.87234	0.8093
model1	0.62419	0.37251	0.80072	0.58662	0.33992	0.83027	0.5851	0.2967	0.84513	0.57092	0.30363	0.84134	0.69851	0.91317	0.80584
model4	0.61223	0.38546	0.77315	0.5793	0.33202	0.82823	0.58012	0.27134	0.84897	0.56016	0.28334	0.83949	0.64328	0.91701	0.78015
model3	0.60161	0.38446	0.76773	0.5761	0.28722	0.84811	0.56767	0.30347	0.80781	0.55564	0.24155	0.84829	0.61119	0.91627	0.76373
model2	0.59629	0.2749	0.83831	0.55661	0.42293	0.74867	0.5858	0.22485	0.85891	0.54188	0.22741	0.86683	0.64254	0.85754	0.75004

Figure 4. *Sensitivity, specificity, and balanced accuracy for each class for all machine learning models.*

	SMC Sens	SMC Spec	SMC Acc
SVC	0.387352	0.847024	0.61719
NuSVC	0.357049	0.870999	0.61402
Voting	0.338603	0.860918	0.59976
LinearSVC	0.332016	0.867184	0.5996
model2	0.422925	0.748672	0.5858
model1	0.339921	0.830268	0.5851
model4	0.332016	0.828225	0.58012
model3	0.28722	0.848113	0.56767

	SMC Sens	SMC Spec	SMC Acc
model3	0.047558	0.975826	0.51169
Voting	0.028278	0.994947	0.51161
NuSVC	0.021851	0.996176	0.50901
LinearSVC	0.020566	0.996176	0.50837
model4	0.015424	0.994537	0.50498
model1	0.007653	0.998107	0.50288
SVC	0.006427	0.998225	0.50233
model2	0.005102	0.998783	0.50194

Figure 5. *SMC classification performance with upscaling (left) vs. no upscaling (right).*

	HC	SMC	EMCI	LMCI	AD
HC	893	418	393	231	73
SMC	266	252	146	67	28
EMCI	627	388	808	368	175
LMCI	315	146	309	410	447
AD	75	23	71	212	959

Figure 6. *Confusion matrix for LinearSVC model.*

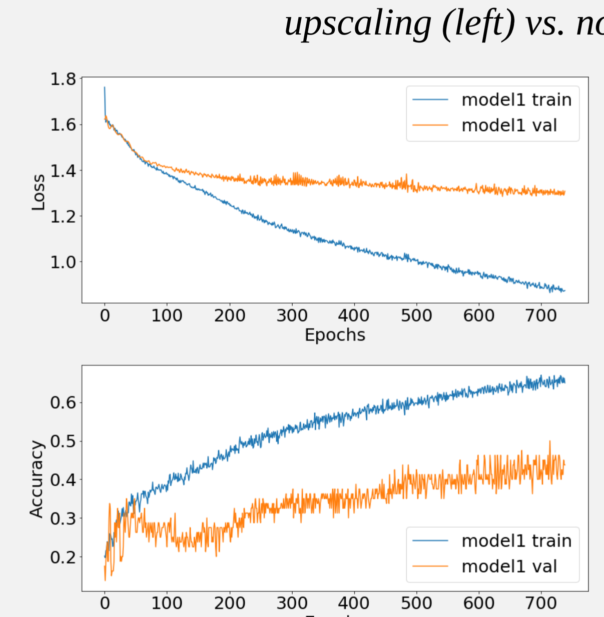


Figure 7. *Learning curve for Model 1.*

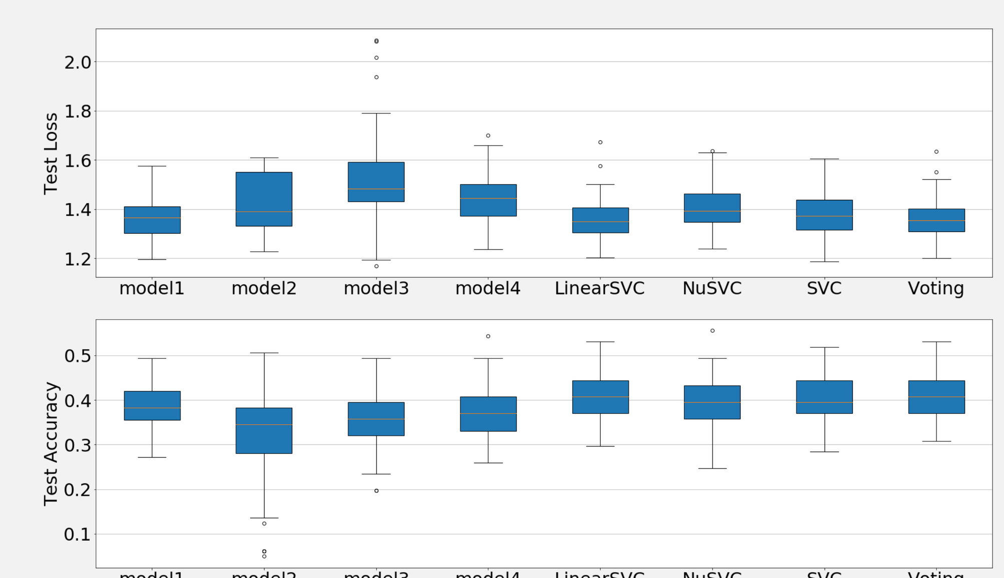


Figure 8. *Test loss and accuracy for all machine learning models.*

Discussion

- Upscaling training data balances sensitivity and specificity. Models were able to better detect underrepresented classes (i.e. SMC).
- Because of the imbalanced classes, balanced accuracy proved to be a more appropriate metric for analyzing model performance.
- For multiclass classification, models generally had little trouble distinguishing between HC and AD. However, there was frequent confusion between SMC, EMCI, and LMCI. This can be attributed to the proximity of the three classes, which could result in incorrect classification of borderline subjects for a particular class.
- SVMs generally outperformed deep learning models. Among the deep neural networks, smaller networks generally outperformed larger ones. This makes sense in the context of this study, since the dataset was relatively small, and smaller models are less susceptible to overfitting.
- Future work could involve adding gene data as another modularity. Incorporating brain networks is also a possibility through the use of graph convolutional networks (GCNs).

Acknowledgements

This study was supported by the Penn Undergraduate Research Mentoring (PURM) program, the Shen Lab at the Perelman School of Medicine at the University of Pennsylvania, and funding from NIH R01 EB022574, R01 LM013463, RF1 AG063481, and NSF IIS 1837964. The complete ADNI Acknowledgement is available at http://adni.loni.usc.edu/wpcontent/uploads/how_to_apply/ADNI_Acknowledgement_List.pdf.