Predicting Dementia Diagnosis with Neuroimaging Data

Boya Jiang,Chen Chen,Tiankai Xie,Yuan Zhong

12/15/2021

Introduction

Alzheimer's is a dementia degenerative disease starting with mild memory impairment in the early stages and progressing to a complete loss of the mental and physical faculties. Definitive Alzheimer's Disease (AD) diagnosis relies on a magnetic resonance image (MRI) study. Brain MRI scans are detailed three-dimensional anatomical images, and changes in the hippocampus, frontal and parietal regions are evidential markers in the progress of AD. The ability to diagnose and classify AD at an early stage allows clinicians to make more knowledgeable decisions regarding clinical interventions. In this project, we apply multiple representative methods to predict Alzheimer's status.

Data and Preprocessing

OASIS (Open Access Series of Imaging Studies) is a well-known initiative that is publicly available for study and analysis. The present MRI dataset, OASIS-I (presented in 2007), is a cross-sectional collection of data for 416 participants aged 18-96 yrs, 316 non-demented and 100 at various stages of AD. Subjects were characterized by the Clinical Dementia Rating (CDR) scale from cognitive normal (CDR = 0), very-mild dementia (CDR = 0.5) to mild dementia (CDR = 1). The data set also contains the following demographics information: male/female, age, education (Educ), and socioeconomic status (SES). Data set of MRI scans of axial plane contains 176 slices/images of 176 x 208 pixel size.

Since AD is more prevalent among older adults, we selected individuals between 60 and 96 years old. The CDR was dichotomized to 0 for cognitive normal (CN), and 1 for any level of dementia. Our final samples composed of 198 individuals, which were randomly split into two groups in 80:20 ratios as training and testing sets, respectively. In our variable of interest, there does not exist data formatting problem, but there is a missing data problem in SES, with 18 missing records. By using predictive mean matching as imputation method from mice package, missing data problem is solved. Table 1 shows a summary of demographic variables collected for CN and dementia groups. The training and testing sets had balanced CDR distribution. Figure 1 shows a comparison of MRI scan between a CN patient and a dementia patient at the same slice.

Characteristic	**Overall**, N = 198	**0**, N = 98	**1**, N = 100
M.F			
F	131 (66%)	72 (73%)	59 (59%)
M	67 (34%)	26 (27%)	41 (41%)
Age	76 (71, 82)	74 (69, 83)	77 (72, 81)
Educ			
1	23 (12%)	8 (8.2%)	15 (15%)
2	60 (30%)	25 (26%)	35 (35%)
3	40 (20%)	23 (23%)	17 (17%)

Characteristic	**Overall**, N = 198	**0**, N = 98	**1**, N = 100
4	36 (18%)	17 (17%)	19 (19%)
5	39 (20%)	25 (26%)	14 (14%)
SES			
1	37 (19%)	20 (20%)	17 (17%)
2	52 (26%)	31 (32%)	21 (21%)
3	49 (25%)	23 (23%)	26 (26%)
4	57 (29%)	23 (23%)	34 (34%)
5	3 (1.5%)	1 (1.0%)	2 (2.0%)

Methods and Results

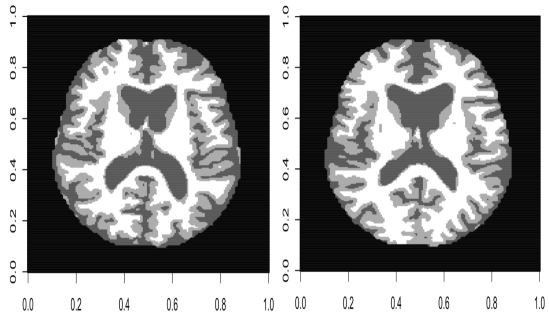

Figure 1: MRI scan of CN (left) and dementia (right) patients

In order to classify dementia from CN, we applied the following four methods: penalized logistic regression (ridge and lasso), random forest, support vector machine (linear and radial kernel), and convolutional neural network (CNN).

Data Preperation In order to prepare data for all 3D methods, we selected the middle slice for each patient, creating an array of dimensions 176x208x176. Next, to improve the speed of model analysis, data was further cleaned by removing axes containing all zero values.

a) Logistic Regression
Logistic regression (LR) was perferred over linear discriminant analysis (LDA) because it does not require the independent variables to satisfy the assumptions of linearity, normal distribution, or equal variance. LR also provides a deterministic model yielding weighting factors for each contributing variable, while avoiding overfitting the data.

In order to prepare data for logistic regression, we selected the middle slice for each patient, creating an array of dimensions 176x208x176. Next, to improve the speed of model analysis, data was further cleaned by removing axes containing all zero values. Parameter tuning was performed through 10-fold cross-validation analysis via `cv.glmnet`. Both ridge and lasso methods were used for regularization. On that basis, we further incorporated demographic variables with neuroimaging data and fit ridge logistic regression.

b) Random Forest
Random forests is an ensemble learning method. For classification tasks, its output is the class selected by most trees. It doesn't overfit with more features and its efficiency is particularly notable in large data sets.

Parameter tuning was performed through 10-fold cross-validation analysis via `trainControl` which performs control on train data from `caret` package. The random forests was called from `randomForest`. On that basis, we further incorporated demographic variables with neuroimaging data and fit the model again.

c) SVM
Support vector machine (SVM) is a supervised learning method. It is effective in high dimensional spaces

and in cases where number of dimensions is greater than the number of samples. It is also memory efficient since it uses a subset of training points in the decision function, known as support vectors.

Parameter tuning was performed through 10-fold cross-validation analysis via `trainControl` which performs control on train data from `caret` package. Both linear and radial kernel were used for regularization. The SVM with linear kernel was called from `e1071`and the SVM with radial kernel was called from `kernlab`. On that basis, we further incorporated demographic variables with neuroimaging data and fit both models again.

d) Naive Bayes
Naive Bayes is a classifier based on applying Bayes' theorem with strong independence assumptions between the features. It coupled simple Bayesian network models with kernel density estimation, and can achieve higher accuracy levels.Unlike ther types of classifiers wihch use expensive iterative approximation, its maximum-likelihood training can be done by evaluating a closed-form expression, which only takes linear time.

Parameter tuning was performed through 10-fold cross-validation analysis via `trainControl` which performs control on train data from `caret` package. The Naive Bayes was called from `klaR`. On that basis, we further incorporated demographic variables with neuroimaging data and fit model again.

e) CNN

Deliverable and Code Repository

Shiny App

Shiny Application We built our Shiny application using the CNN model(?) as our classifier embedded behind the dashboard. This means that the dashboard can take patients MRI data and some demographic and clinical input to predicts of Dementia diagnosis using CNN model. The Shiny application can be viewed by running this command, shiny::runGitHub('625OASIS','y1zhong', subdir = 'shiny'), on the R console. Packages like Shiny and oro.nifti must be installed before this app can be run.

GitHub

A GitHub repository is created to store, update, display, and share the results of our project. Here is the link: https://github.com/y1zhong/625OASIS

Google Colab?

Evaluation

To evaluate the performance of each model, we compared th test sensitivity, specificity, and accuracy across models (Table 2). Overall, the Ridge logistic regression performed the best predication with the highest accuracy(0.79), sensitivity(0.75), and specificity(0.83), followed by Lasso regression with a relative high accuracy (0.68). However, RF and LDA had a higher sensitivity than Lasso regression (0.65 and 0.63 respectively). Additionally, Figure 2 presented the ROC curve for each model along with the AUC value which also indicated that Ridge regression was the best model to predict Alzheimer dementia. We further adjusted demographic information in our model and compared the result with image data only. For logistic penalized regression, since the Ridge method performed better than the Lasso, we only adjusted for demographic in Ridge regression. However, further adjusting demographic did not improve the prediction result of Ridge regression, LDA, and SVM. However, the specificity of Random Forest method increased 23%, while sensitivity dropped to 28%. performance

	Image Only					Image+CS					
	Lasso	Ridge	LDA	SVM-L	RF	Ridge	LDA	SVM-L	RF		Lasso
Sensitivity	0.6	0.75	0.63	0.49	0.65	0.56	0.75	0.63	0.49	0.28	0.56
Specificity	0.78	0.83	0.51	0.47	0.51	0.63	0.83	0.53	0.47	0.74	0.63
Accuracy	0.68	0.79	0.57	0.48	0.58	0.59	0.79	0.58	0.48	0.51	0.59

Discussion

In conclusion, we compared the prediction of multiple classification models and the Ridge regression performed the best in terms of accuracy, sensitivity, and specificity. Comparing to literature, our analyses had several important strengths. Some literature may have outstanding accuracy that is more than 95% when compared Alzheimer's dementia MRI with health control MRI from young population. However, young population were not at risk of developing AD and using their MRI as control will introduce bias and over predicting the result. Our analysis only included elder adults so our model classification would be more practice to distinguish AD and healthy aging brain in reality. Secondly, our outcome AD was balanced among cases and controls which would increase our accuracy. However, accuracy of AD prediction was different by the different stage of AD. AD was hard to diagnosis at early statement (MCI) as usually there was no significant symptoms, but the degeneration of the brain had began. AD was more likely to be diagnosis or discovered at more severe stage, in our case, which was when CDR equaled one or two. Therefore, accuracy would be differentiated by the level of comparison and differentiating MCI from healthy aging was the most difficult and past classification models showed much lower accuracy between 40% to 60%. In our analysis, we combined MCI, AD and severe AD as one cased group which made prediction harder with lower accuracy. Our cNN method was the most challenging model as it not only required large sample size, but also advanced computing environment as cluster computing or GPU computing to incorporate the high-dimension structure. Many research teams already recruited extraordinary engineers and computer science expertise to improve the algorithm, but with less focus important biomarkers of AD such as APOE and CSF and other important clinical information. Future research should consider including more biostatistics, neurology, and epidemiology field to create a multi-disciplinary research.

Contributions

References