# Alzheimer's Disease Classification Testing Using ADNI Data

Charlee Cobb

2023-05-02

---

**Table of Contents**



---

**Abstract**

Across the globe, the number of people with Alzheimer's disease is increasing. While the cure for the disease has not been found, scores of people have come together to collect data on patients with this disease. The Alzheimer's Disease Neuroimaging Initiative (ADNI) in particular is a leading organization on neuroimaging data. Testing with ADNI data has lead to a number of discoveries of biomarkers related to Alzheimer's disease. Given the large amount of neuroimaging data, computational methods like machine learning are helpful in classifying the disease and related biomarkers. This paper reviews the use of Random Forest models to classify various cognitive states, including Alzheimer's Disease. Random Forest was found to be the best performing model against SVM models.

**Background and Motivation**

Alzheimer's Disease is the most common type of dementia in the United States, and it is estimated that 14 million people will be diagnosed with the disease by 2060 (CDC, 2020). As with many diseases, diagnosis and screening is imperative to properly treat and aid individuals with Alzheimer's Disease. Causes of Alzheimer's Disease is also associated with multiple factors, as there is yet to be a single driving cause. Luckily, data on patients diagnosed with Alzheimer's has been collected for a number of years. Imaging data of people's brains may lead to the discovery of biomarkers that correlate with the progressing states of dementia. The data may also aid in developing better diagnostic and treatment options for future patients.

Using supervised machine learning algorithms is one way to take all the collected data points and predict a diagnosis. Regular use of machine learning prediction tools could help remove biases and error from diagnosis, and get people the help they need faster (Uddin, S., 2019). The Alzheimer's Disease Neuroimaging Initiative (ADNI) is an organization that develops clinical, imaging, and genetic data that can be used to develop

diagnostic standards for Alzheimer's. ADNI has released a number of competitions, and invites teams to develop tools and test models on their data sets. The National Research Council launched a competition called "The international machine learning neuroimaging challenge for automated diagnosis of Mild Cognitive Impairment competition with ADNI data" in 2017(Castiglioni, I., 2017). The winner of this challenge found that Random Forest Classifiers were the best classifiers to predict the early diagnosis of Alzheimer's and differential diagnostic states of cognitive function (Dimitriadis, 2018). Inspired by this, I contacted ADNI for access to their database (The data prepared for the Kaggle competition in 2017 was no longer available to the public) and tested the classification capabilities of the Random Forest classifier against other supervised machine learning algorithms.

**Research Question and Hypothesis**

While single classifiers have been the preferred method of machine learning for neuroimaging machine learning tests, they have fallen out of favor to adapt to the complexity and nuances of diagnostic neuroimaging (Dimitriadis, 2018). The goal of this project is to use Random Forest modeling on the neuroimaging data and compare the results to SVM modeling. I hypothesis that the Random Forest model performs better than SVM for classifying cognitive states in a patient.

**About the data set and Methods**

This data set was developed in 2017 as part of the "QT-PAD Project Data" challenge. It is an accumulation of ADNI data from 2004 to 2017, and includes MRI data, PET data, clinical examination data, and patient demographic data. Patient identification data was obscured and represented as arbitrary identification numbers. There are 1737 patients represented across 12750 samples (rows in the data). These samples represent one singular clinical visit.

I decided to test the data in two parts. In the first run, I included the patient's clinical exam scores in the testing and training data set. I was curious to see how these additional features impacted the performance of the Random Forest and SVM models. In the second run, I excluded the clinical exam features. This was done to get the data set as close to the data set from the Dimitriadis paper as possible. The second run would test the hypothesis.

The data was split into 80% training and 20% testing. Features dropped from both rounds of testing include "VISCODE", "COLPROT", "ORIGPROT", "PTID", "RID", "SITE", "DX", "EXAMDATE", and "EXAMDATE.bl". "VISCODE", "PTID", and "RID", are patient identifiers that were arbitrary to the data. "COLPROT" and "ORIGPROT" indicate which trials the data comes from. "EXAMDATE" and "EXAMDATE.bl" are the dates of the patients' doctor visits. "DX" is the secondary diagnosis received at follow up visits if one was given. The parameters proposed for the random forest classifier are as follows:

- criterion='gini', n_estimators=25, random_state=1, n_jobs=2

Parameters proposed for SVM are as follows:

- kernel = 'rbf', gamma=0.10, C=10.0, random_state = 1

**Results**

Below are the confusion matrices and prediction scores of each classifier. Random forest was tested with and without the proposed parameters on both the data set with clinical exams and data without clinical exams. SVM was tested with and without proposed parameters on the data set with clinical exams. The best performing SVM classifier (without parameters) was tested on the data set without clinical data. The "important features graph" was produced by the best performing Random Forest classifier, in this case the

classifier with default parameters. The code that produced the following results and images will be attached in the corresponding .ipynb file ("Charlee_Cobb_ADNI_predictions.ipynb") that was used for this project.

With clinical exams in the training data set, Random Forest had 99.61% accuracy, with the most important features being CDRSB, MMSE, FAQ. SVM performed with 49.14% accuracy. In the data set without clinical exams, Random forest had an accuracy of 99.098% and SVM had an accuracy in 55.80%. The features important to the Random forest in this data set were age and hippocampus images.

**confusion matrix output for Random Forest, no parameters, with clinical exams:** rfc_model score: 0.996078431372549

```
##   index predicted_LMCI predicted_CN predicted_EMCI predicted_AD predicted_SMC
## 1  LMCI            776            0              0            0             0
## 2    CN              0          458              0            0             0
## 3  EMCI              0            1            919            0             0
## 4    AD              5            2              0           70             0
## 5   SMC              0            0              0            0             0
```

**Confusion matrix output for SVM model, no parameters, with clinical exams:** svc_model score: 0.49137254901960786

```
##   index predicted_LMCI predicted_CN predicted_EMCI predicted_AD predicted_SMC
## 1  LMCI            245            0            531            0             0
## 2    CN              0           51            407            0             0
## 3  EMCI              0            0            920            0             0
## 4    AD              0            0             76            1             0
## 5   SMC              0            0              0            0             0
```

**Confusion matrix of Random Forest with parameters proposed for this project, with clinical exams:** rfc_model score: 0.9937254901960785

```
##   index predicted_LMCI predicted_CN predicted_EMCI predicted_AD predicted_SMC
## 1  LMCI            776            0              0            0             0
## 2    CN              0          458              0            0             0
## 3  EMCI              0            3            917            0             0
## 4    AD              8            1              0           68             0
## 5   SMC              0            0              0            0             0
```

**Confusion matrix output for SVM with rbf kernel with clinical exams:** svc_model score: 0.4768627450980392

```
##   index predicted_LMCI predicted_CN predicted_EMCI predicted_AD predicted_SMC
## 1  LMCI            220            0            556            0             0
## 2    CN              0           42            416            0             0
## 3  EMCI              0            0            920            0             0
## 4    AD              0            0             77            0             0
## 5   SMC              0            0              0            0             0
```

**Dropping Clinical exams**

**confusion matrix output for Random Forest, no parameters, without clinical exams:** rfc_model score: 0.9909803921568627

```
##   index predicted_LMCI predicted_CN predicted_EMCI predicted_AD predicted_SMC
## 1  LMCI            773            0              3            0             0
## 2    CN              1          457              0            0             0
## 3  EMCI              0            0            918            0             0
## 4    AD              3            3              2           69             0
## 5   SMC              0            0              0            0             0
```

**Confusion matrix of Random Forest with parameters proposed for this project, without clinical exams:** rfc_model score: 0.9858823529411764

```
##   index predicted_LMCI predicted_CN predicted_EMCI predicted_AD predicted_SMC
## 1  LMCI            772            2              2            0             0
## 2    CN              3          455              0            0             0
## 3  EMCI              1            0            917            0             0
## 4    AD              6            7              0           64             0
## 5   SMC              0            0              0            0             0
```

**Confusion matrix output for SVM with rbf kernel without clinical exams:** svc_model score: 0.5580392156862745

```
##   index predicted_LMCI predicted_CN predicted_EMCI predicted_AD predicted_SMC
## 1  LMCI            309            0            467            0             0
## 2    CN              0          128            330            0             0
## 3  EMCI              0            0            920            0             0
## 4    AD              0            0             65           12             0
## 5   SMC              0            0              0            0             0
```

Below are the images of the correlation map of all the features in the data set, and also the most important features of Random forest with clinical exams and random forest without clinical exams respectively.

**Interpretation of Results**

In the above uses of the default classifier settings for Random Forest and SVM, Random Forest out performed the SVM model in all classifications. Both the Random Forest and the SVM models performed better with the default parameters, so custom parameters were not as useful for this data set. The SVM model performed poorly, often making all of its classifications to be EMCI. When looking at the heat map, there are a number of areas where features are highly correlated with other features. This will be a useful map to have for feature reduction tests outside of separating features based on clinical exams, demographic data, and imaging data. As far as feature importance, The Rey's Auditory Verbal Learning Test (RAVLT) test scores were a relatively important feature.This is interesting because the RAVLT test wasn't developed just to diagnose Alzheimer's disease. Clinical exams, demographics, and MRI data were more important features for Random Forest classification than PET data (tau / pau data). Baseline values were consistently more important than follow up visits for diagnosis. This may be because patient's cognitive states didn't change in follow ups, or patient's only had one doctor visit recorded.

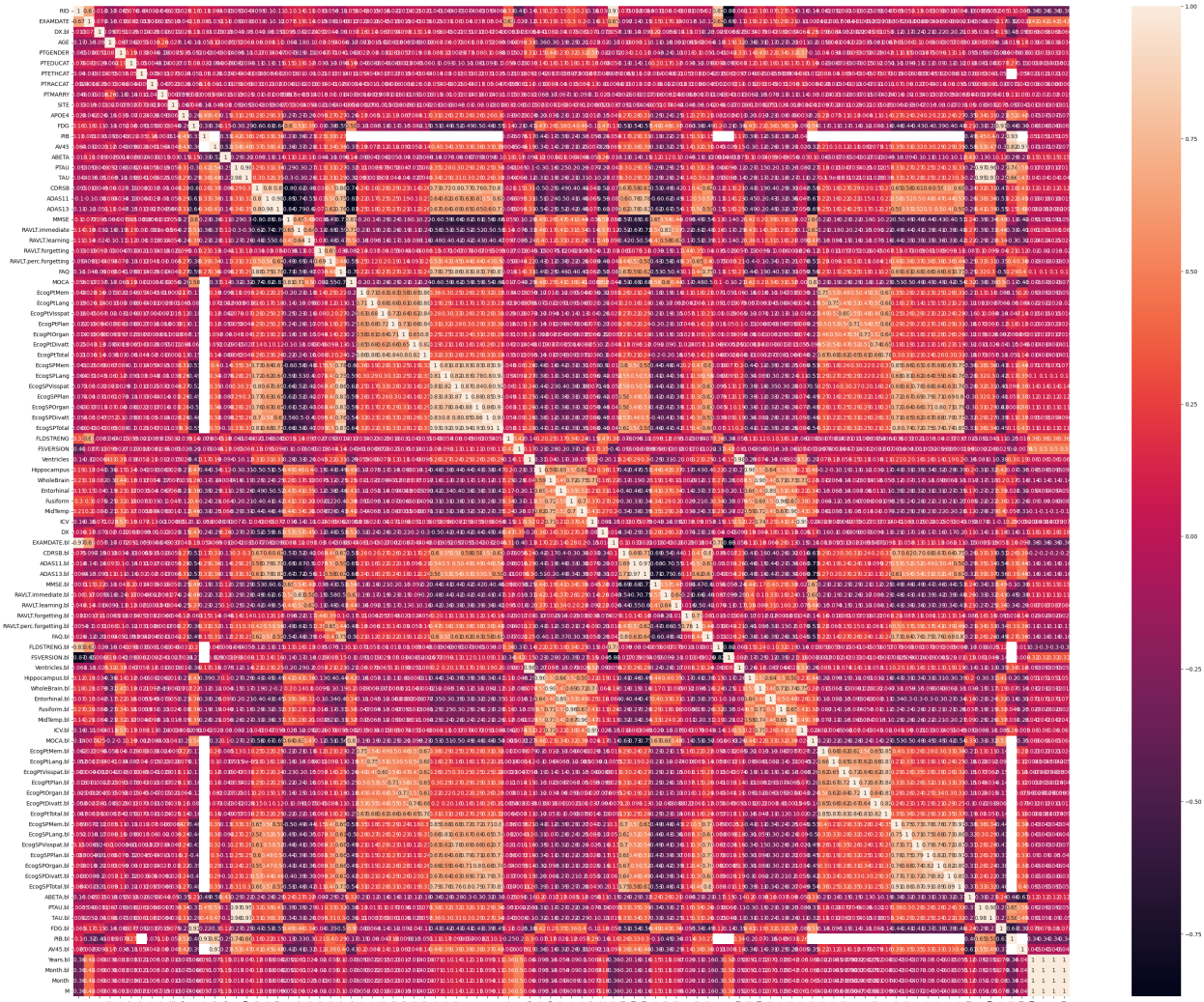

Figure 1: Correlation Map of all features

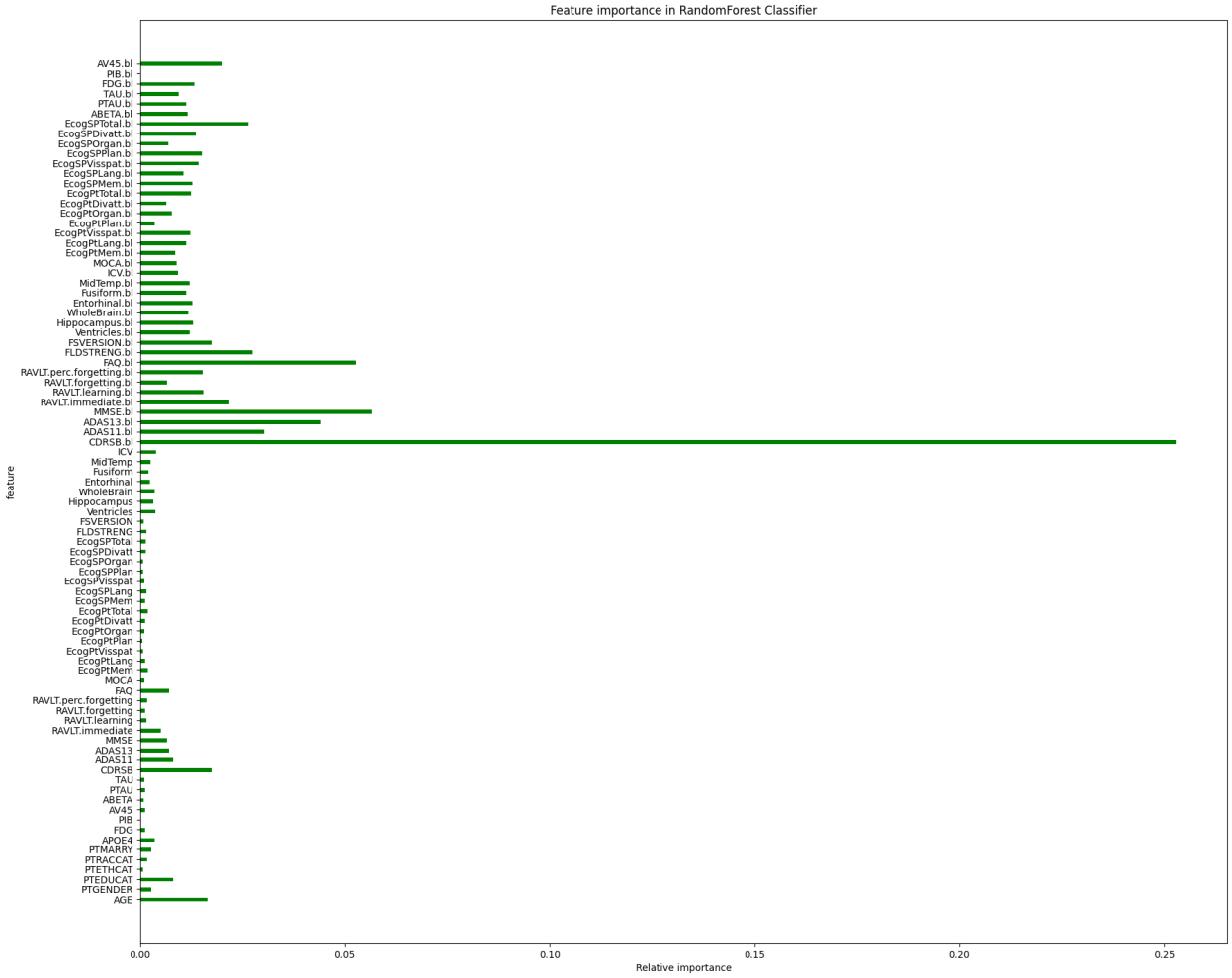


Figure 2: Important features of Random Forest with clinical exam data

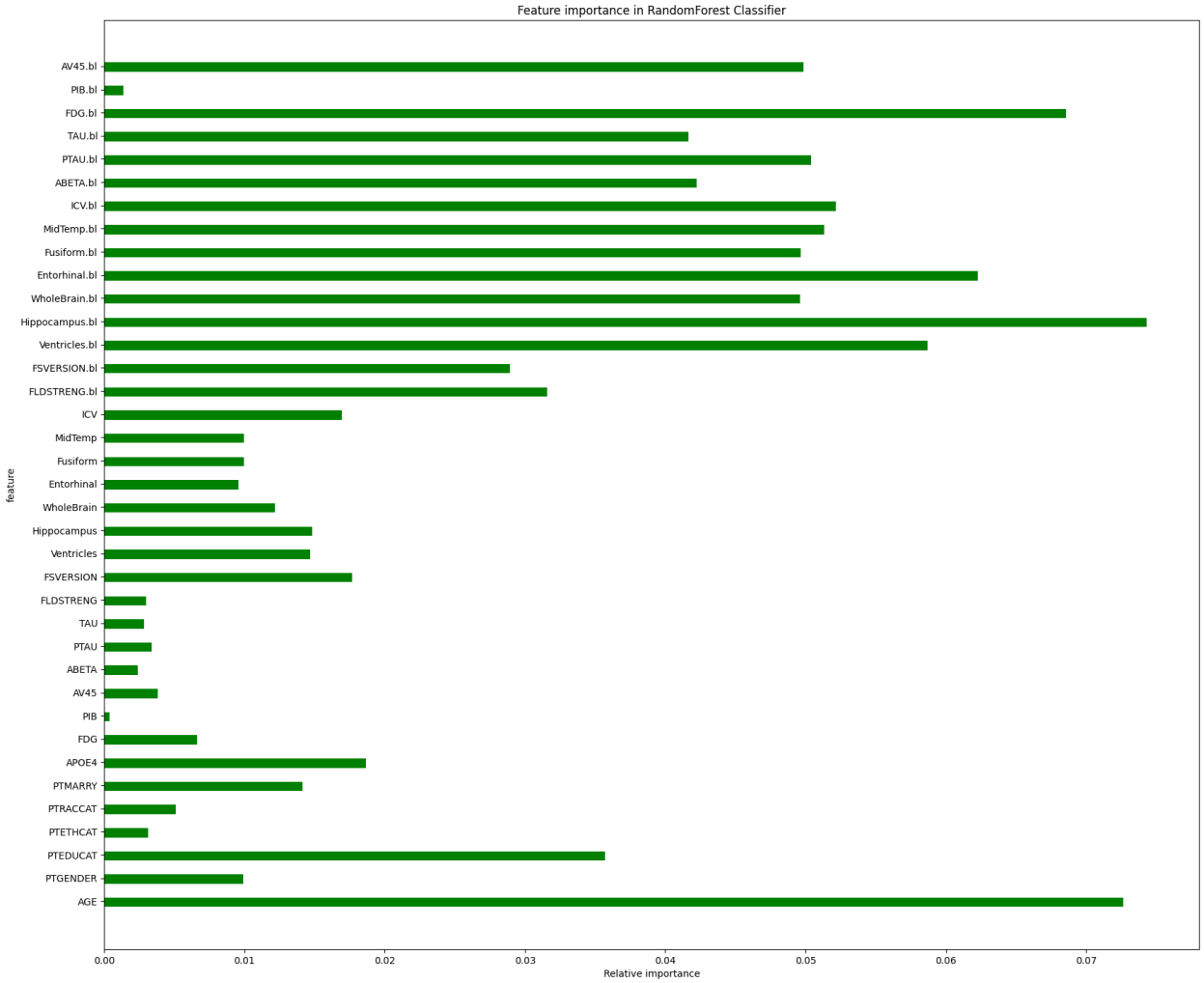


Figure 3: Important features of Random Forest without clinical exam data

## Discussion of Limitations and Alternative Interpretations

Random forest worked very well with and without hyper parameter testing. In the future, it would be interesting to see if further feature reduction would increase the precision of classification. The SVM model probably failed to perform well because the data was too noisy and unbalanced. There were only 3826 healthy controls (CN) in the entire data set compared to the remaining 8923 impairment states. Scaling the data may have helped improve the SVM models classification accuracy. I did note that SVM worked better with less features. If I had reduced features down to only neuroimaging, SVM may have worked as well as Random Forest. However, the integration of multiple features for diagnostic purposes seems to be more reliable and more realistic for use in a clinical setting. Therefore, I would still retain that Random Forest works better at classifying cognitive states in patients.

I would also like to do more research on the diagnostic values and clinical exam scores to discover any bias that is introduced from doctors. Alzheimer's disease cannot be definitively diagnosed without a biopsy of the patient's brain. These diagnostic values, even without clinical exams, are still made based on a physician's classification. It would be very interesting to test this type of data set with the inclusion of the post-mortem diagnosis of patients who were diagnosed with Alzheimer's disease while alive. Inclusion of this data could get us closer to a definitive biomarker that causes Alzheimer's disease.

## Conclusion

In all, this project was enlightening and encouraging. With the use of machine learning classifiers, it's possible to quantify biomarkers that are related to classification of a disease. Random Forest classifier is still the best classifier for this data set. While data collection is slow and ongoing, I hope that more machine learning techniques can be integrated into clinical diagnostics.

## References


Centers for Disease Control and Prevention. (2020) Site accessed April 2023. https://www.cdc.gov/aging/aginginfo/alzheimers.htm

A. Armstrong, R. (2013). Review articleWhat causes alzheimer's disease?. Folia Neuropathologica, 51(3), 169-188. https://doi.org/10.5114/fn.2013.37702

Uddin, S., Khan, A., Hossain, M. et al. Comparing different supervised machine learning algorithms for disease prediction. BMC Med Inform Decis Mak 19, 281 (2019). https://doi.org/10.1186/s12911-019-1004-8

Dimitriadis SI, Liparas D; Alzheimer's Disease Neuroimaging Initiative. How random is the random forest? Random forest algorithm on the service of structural imaging biomarkers for Alzheimer's disease: from Alzheimer's disease neuroimaging initiative (ADNI) database. Neural Regen Res. 2018 Jun;13(6):962-970. doi: 10.4103/1673-5374.233433. PMID: 29926817; PMCID: PMC6022472.

Castiglioni, I., Salvatore, C., Ramírez, J., Górriz, J.M., Machine-learning neuroimaging challenge for automated diagnosis of mild cognitive impairment: Lessons learnt, Journal of Neuroscience Methods, Volume 302, 2018, Pages 10-13, ISSN 0165-0270, https://doi.org/10.1016/j.jneumeth.2017.12.019. (https://www.sciencedirect.com/science/article/pii/S0165027017304375)

Moradi E, Hallikainen I, Hänninen T, Tohka J; Alzheimer's Disease Neuroimaging Initiative. Rey's Auditory Verbal Learning Test scores can be predicted from whole brain MRI in Alzheimer's disease. Neuroimage Clin. 2016 Dec 18;13:415-427. doi: 10.1016/j.nicl.2016.12.011. PMID: 28116234; PMCID: PMC5233798.

Ivelina Gueorguieva, Laiyi Chua, Brian A. Willis, John R. Sims, Alette M. Wessels, Disease progression model using the integrated Alzheimer's Disease Rating Scale, Alzheimer's & Dementia, 10.1002/alz.12876.

Pedregosa, F. and Varoquaux, G. and Gramfort, A. and Michel, V. and Thirion, B. and Grisel, O. and Blondel, M. and Prettenhofer, P. and Weiss, R. and Dubourg, V. and Vanderplas, J. and Passos, A. and Cournapeau, D. and Brucher, M. and Perrot, M. and Duchesnay, E., Scikit-learn: Machine Learning in {P}ython, Journal of Machine Learning Research, 2011.