

Autism Spectrum Disorder classification using Machine Learning techniques on fMRI

Mihir Mathur

*Computer Science Department
University of California, Los Angeles
mihirmathur@ucla.edu*

Tyler Lindberg

*Computer Science Department
University of California, Los Angeles
tlindberg@ucla.edu*

I. INTRODUCTION

Autism Spectrum Disorder (ASD) is a neurological disorder that adversely affects a person's capability to communicate and do social interactions. The current clinical standards for diagnosis of ASD include developmental screening and comprehensive diagnostic evaluation. However, these clinical methods cannot effectively diagnose ASD till about 2 years of age [1], due to which many autistic children are unable to get the requisite support and resources early on.

In recent years, the research into the effectiveness of machine learning for ASD classification has been growing. In this research project, we investigate this novel approach of classification of ASD by applying five supervised machine learning algorithms to fMRI data of ASD subjects and control subjects. The data used for this research is a subset of the Autism Brain Imaging Data Exchange (ABIDE data-set), which has fMRI data from about 500 ASD and control subjects each[2]. In this report, we outline our methodology for preprocessing data and for generating different sets of features from the fMRI time-series data. We then describe our methodology for running various machine learning models. Then we discuss and compare the accuracy of the different Machine Learning models used on the fMRI with one another and with the accuracy obtained from using the phenotypic data for the models. Finally, we summarize which feature-sets worked the best for our models and the top features for classification of ASD we found.

II. METHODOLOGY

A. Pre-processing Data and Generating Features

Two subsets of ABIDE were used for this research:

- 1) Phenotypic Data including AGE, SEX, HANDEDNESS, FIQ, VIQ, PIQ of ASD and control subjects.
- 2) fMRI scans of ASD and control subjects.

For preprocessing the phenotypic data, alphabetical features such as "handedness" were converted to numerical features and data then normalized.

Preprocessing the fMRI images took multiple steps. First, the correct format for the fMRI data needed to be selected among all possible preprocessing steps. In this study, the Data Processing Assistant for Resting-State fMRI (DPARSF) pipeline was chosen based on the methods of Bi et .al [3][4].

Next, the format of the data was chosen as fMRI represented in the Automated Anatomical Labeling (AAL) brain atlas. AAL divides the brain into 116 macroscopic brain structure and creates a vector time-series of the brain activity for the duration of the fMRI [5]. However, due to the variety of configurations used, each time-series has a different length which makes it difficult to compare time-series directly. Additionally, the time-series does not illuminate how different regions of the brain interact with one another, making it difficult to use as a feature set. To account for this issue, the Pearson correlation coefficient was computed for every pair of the 116 AAL brain structures, resulting in $(116 \times 115/2) = 6670$ features which are independent of time. Then from this data, all samples that resulted in a floating point error during the computation were thrown out of the correlation data set.

The pairwise Pearson correlation coefficients can also be used to generate measures that stem from graph theory, better characterizing the brain networks integration and segregation, with approximately 7,000 additional generated features. In order to generate these graph measures, one first needs to convert the correlation coefficients into an adjacency matrix of the absolute value of each correlation. Then all values below a fixed constant are thresholded to zero. In this case a value of 0.25 was chosen based on the work of Bi et. al [4]. This results in an adjacency matrix which can be used to compute various graph measures. For the computation of these graph measures a Python library, bctpy, was used, which mirrors the functionality of the Brain Connectivity Toolbox written in Matlab [6][7].

The motivations behind using graph measures as a feature set come from an article that explores classification of Alzheimer's using machine learning [8]. Among the graph measures there are three types of measurement: functional segregation, functional integration, and local nodal measures. In functional segregation the weighted clustering coefficient and local efficiency were used. For functional integration the weighted distances and characteristic path length were used. And with local nodal measures only the node degree was chosen. These decisions form a subset of the decisions made by Khazae et. al with the inclusion of the distance matrix values as an additional metric from the methods of Bi et .al [8][4].

Additionally, there exists a nuance when computing mea-

measurements related to functional integration. When computing the graph distance between entries in the adjacency matrix, it is natural for those with lower value separations to be considered “closer” to one another. However, when using the Pearson correlation coefficient, nodes which are more closely related have a larger correlation coefficient. For this reason, the correlation adjacency matrix needs to be “inverted” to match the expectations of the functional integration measurement. To do so, all values k in the adjacency matrix are transformed according the following transformation.

$$k \rightarrow \begin{cases} \infty & \text{if } k = 0 \\ 1 - k & \text{if } k > 0 \end{cases}$$

Through this transformation, nodes with higher correlation are considered “closer” to one another, and those with no correlation are infinitely far from one another. In combination with the thresholding operation performed in the previous steps, the resultant matrix will contain many values of infinity, hence making the shortest distance between nodes a nontrivial calculation that is performed by bctpy [6].

Additionally, another, much simpler, preprocessing was done by computing the mean intensity of each AAL region and generating another feature vector with 116 elements.

B. Analyzing Data with Machine Learning Algorithms

For running machine learning algorithms, the preprocessed data with Pearson correlation coefficient, graph measures, and intensity averages used as features was divided in a 80:20 ratio for training and test. For each feature set combination the feature values were normalized to a range of -1 to 1 to account for differences in the range of measurements between the different features. We chose five supervised machine learning algorithms for training on each of the feature-set: Support Vector Machines (Linear and RBF Kernel), Adaboost (Decision Tree Base Classifier), Neural Networks (Multi-layer Perceptron Classifier), Random Forests, K-Nearest Neighbors. These algorithms were chosen since each of these has a very distinct approach of classification. We used Python’s machine learning library, sklearn, for implementing these algorithms.

For tuning the hyperparameters, a 3-fold cross validation with a grid search of hyperparameters was performed. Accuracy, area under ROC-curve, Precision and Recall were examined to compare the different models.

Of all the methods SVM’s were explored in more depth than the others. In addition to the methods used above, 5-fold cross validation was tested as another method of cross validation. Another component of SVM’s that was tuned was the kernel. When training, a linear, radial basis function (RBF), and sigmoid kernel were tested with SVM’s to compare results. Furthermore, the number of features used in training was added as an additional hyperparameter when training SVM’s. To tune this hyperparameter, the top K features were chosen in each of the training sets during cross validation, and subsequently the value of K that achieved the best validation accuracy was chosen for final testing.

In addition, feature importance was computed for each of the features after training the model. To do so feature importance was calculated by running Random Forest 100 times and examining the top 10 features on each iteration. In the case of correlation coefficients, this could be further decomposed into determining the importance of each AAL region by taking into account which regions the important connections are between.

III. RESULTS AND DISCUSSION

With three types of feature sets (Pearson correlation coefficients, mean intensities, and graph measures), all possible combinations were attempted when training the models. Of all the combinations, the Pearson correlation coefficients yielded the highest accuracies when used without any of the other features.

Classifying using the 5 methods described above gave relatively similar accuracies and AUROCs for each method, with the highest accuracy attained being 66.08% from Adaboost and the highest AUROC being 0.70 for Neural Networks. Classifications of all accuracies can be found in Figure 1. The ROC curves for the top three models can be found in Figures 2 and 3.

With the additional tests performed on SVM’s it was found that Linear and RBF kernels produced the best accuracy at approximately 65%, with the linear kernel achieving a slightly higher AUROC at 0.68 rather than 0.67. Additionally, 5-fold cross validation tended to perform better than 3-fold cross validation, potentially implying that it allowed the model the better generalize to the data. Lastly, approximately 800 features is where linear SVM’s attained the best accuracy, showing that not all correlation coefficients are necessary for an effective classification.

It is worth noting that 66% is similar to accuracies achieved in other studies using machine learning techniques on the ABIDE dataset, indicating that these results are in line with the current state of the art [9][10].

For the subset of phenotypic data we analyzed, simple Decision Trees had the best accuracy of 62%. However, due to the structure of the ABIDE dataset, not all samples had the phenotypic features used for training, so the sample size was significantly smaller for these tests.

The most important features calculated from the Random Forest model were the Pearson correlations between the left Thalamus, right medial frontal gyrus, left gyrus Rectus and Left middle temporal gyrus and other AAL regions shown in figure 4. Of the important regions, connections with the left Thalamus consistently attained the highest feature importance of all 116 AAL regions. This result coincides with other research in the field that claims there may be a correlation between impaired thalamocortical connectivity and ASD [11].

IV. CONCLUSION

The accuracy results show that there is promise in pursuing this research further. However, due to relatively low accuracy scores, machine learning techniques will most likely not

replace current clinical diagnostic methods for individuals with ASD anytime soon.

One question for further exploration is whether increasing the number of samples would help the models achieve a higher accuracy. With only 1000 individuals in the ABIDE dataset, it is difficult to conclude that the samples are completely indicative of the wider population.

Additionally, this research could be used to better explore and discover some of the neuroanatomical causes or results of having ASD. With the connection between the left Thalamus and ASD being found in this project's models and the wider medical community, machine learning models could help illuminate other possible brain regions that are correlated with ASD or confirm proposed correlations from other studies.

Finally, the lower accuracies achieved when classifying autism in comparison to other disorders may be indicative of the wide variety of individuals who fall under the categorization of ASD. Since ASD is a "spectrum" disorder, there are a large array of different phenotypic characteristics that may classify an individual as having ASD, hence this wide range of diagnosis may make it harder to use machine learning techniques for this problem. An interesting study to assert this claim would be to include only individuals with severe cases of ASD in the dataset and see if better classification accuracies could be achieved by the resultant models.

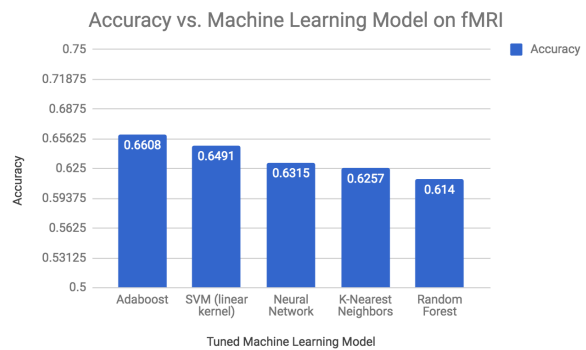


Fig. 1. Accuracies of the five machine learning models on Pearson Correlation of AAL regions

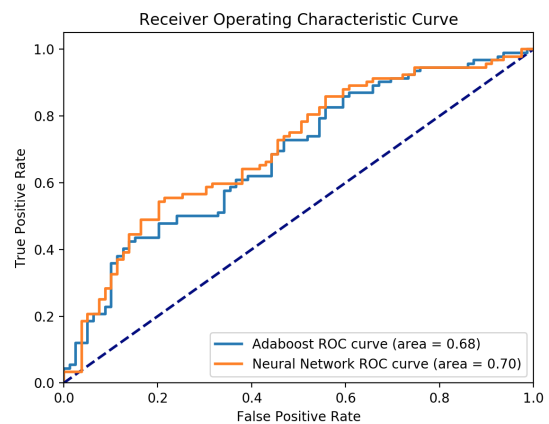


Fig. 2. ROC curves of Adaboost and Neural Net models

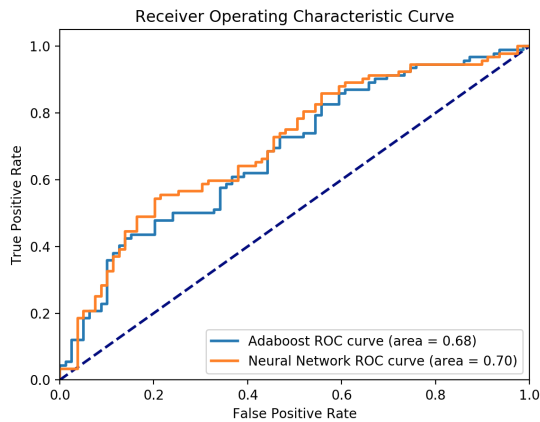


Fig. 3. ROC curves of SVM models with three kernel types

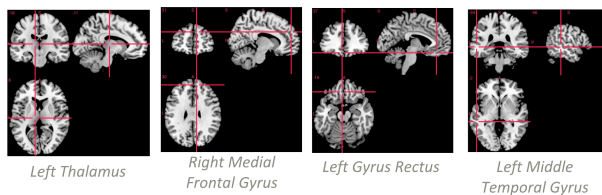


Fig. 4. The brain regions with the highest frequency of important connection features for ASD classification

REFERENCES

- [1] "Autism spectrum disorder," <https://www.cdc.gov/ncbddd/autism/screening.html>, February 2015, accessed: 2018-06-09.
- [2] C. Craddock, Y. Benhajali *et al.*, "The neuro bureau preprocessing initiative: open sharing of preprocessed neuroimaging data and derivatives," *Frontiers in Neuroinformatics*, no. 41, 2013. [Online]. Available: <http://www.frontiersin.org/neuroinformatics/10.3389/conf.fninf.2013.09.00041/full>
- [3] C. Yan and Y. Zang, "Dparaf: a matlab toolbox for "pipeline" data analysis of resting-state fmri," *Frontiers in Systems Neuroscience*, vol. 4, p. 13, 2010. [Online]. Available: <https://www.frontiersin.org/article/10.3389/fnsys.2010.00013>
- [4] X.-a. Bi, Y. Wang, Q. Shu, Q. Sun, and Q. Xu, "Classification of autism spectrum disorder using random support vector machine cluster," *Frontiers in Genetics*, vol. 9, p. 18, 2018. [Online]. Available: <https://www.frontiersin.org/article/10.3389/fgene.2018.00018>
- [5] N. Tzourio-Mazoyer, B. Landeau *et al.*, "Automated anatomical labeling of activations in spm using a macroscopic anatomical parcellation of the mni mri single-subject brain," *NeuroImage*, vol. 15, no. 1, pp. 273 – 289, 2002. [Online]. Available: <http://www.sciencedirect.com/science/article/pii/S1053811901909784>
- [6] "bctpy," <https://github.com/aestrivex/bctpy>, 2018.
- [7] M. Rubinov and O. Sporns, "Complex network measures of brain connectivity: Uses and interpretations," *NeuroImage*, vol. 52, no. 3, pp. 1059 – 1069, 2010, computational Models of the Brain. [Online]. Available: <http://www.sciencedirect.com/science/article/pii/S105381190901074X>
- [8] A. Khazaee, A. Ebrahimzadeh, and A. Babajani-Feremi, "Identifying patients with alzheimers disease using resting-state fmri and graph theory," *Clinical Neurophysiology*, vol. 126, no. 11, pp. 2132 – 2141, 2015. [Online]. Available: <http://www.sciencedirect.com/science/article/pii/S1388245715002242>
- [9] A. S. Heinsfeld, A. R. Franco, R. C. Craddock, A. Buchweitz, and F. Meneguzzi, "Identification of autism spectrum disorder using deep learning and the abide dataset," *NeuroImage: Clinical*, vol. 17, pp. 16 – 23, 2018. [Online]. Available: <http://www.sciencedirect.com/science/article/pii/S2213158217302073>
- [10] C. P. Chen, C. L. Keown *et al.*, "Diagnostic classification of intrinsic functional connectivity highlights somatosensory, default mode, and visual regions in autism," *NeuroImage: Clinical*, vol. 8, pp. 238 – 245, 2015. [Online]. Available: <http://www.sciencedirect.com/science/article/pii/S2213158215000698>
- [11] A. Nair, J. M. Treiber, D. K. Shukla, P. Shih, and R.-A. Mller, "Impaired thalamocortical connectivity in autism spectrum disorder: a study of functional and anatomical connectivity," *Brain*, vol. 136, no. 6, pp. 1942–1955, 2013. [Online]. Available: <http://dx.doi.org/10.1093/brain/awt079>