Objective

For the past 60 years, the anxiety and depression medications are prescribed to patients based on “The Hamilton Depression Rating Scale” (HDRS) and “Social and Functioning Assessment Scale” (SOFAS). These scores are very subjective as they are determined by clinicians based on patient interview without incorporating scientific evidence based on Brain FMRI data. Objective of this project is to identify whether HDRS score and SOFAS scores are representative of the three antidepressants prescribed based on FMRI data of 5 brain attributes.

Data Set and Features

Dataset consists of 128 patients FMRI data obtained from Williams Pan-Lab, Precision Psychiatry and Translational Neuroscience, Stanford Medicine iSPOT-D project. There are 11 features: age, gender, education, 3 antidepressants: Sertraline, Venlafaxine, escitalopram, 5 FMRI brain scan data from brain region Amygdala, Insula and Nucleus Accumbens along with HDRS and SOFAS scores for all the patients. Our project analyzes both supervised and unsupervised methods. All methods are carried out independently both for HDRS and SOFAS scores. In supervised model, HDRS/SOFAS score is a dependent variable and models are fit using different combinations of brain data as feature variables. In unsupervised model, we studied 8 features(3 antidepressants) and 5 brain scan data to understand the association between medicine and brain attributes.

Models

We selected 20 patients randomly out of 128 as test set and use K-fold cross validation(K=10) on 108 patient data to train and validate our models. The following 6 supervised and unsupervised models are considered for the project.

<table>
<thead>
<tr>
<th>Logistic Regression</th>
<th>Linear Regression</th>
</tr>
</thead>
<tbody>
<tr>
<td>Supervised Algorithm</td>
<td>Supervised Algorithm</td>
</tr>
<tr>
<td>Bayesian Linear Regression with Laplace Prior</td>
<td>Factor Analysis</td>
</tr>
<tr>
<td>K-means Clustering</td>
<td>SVM</td>
</tr>
</tbody>
</table>

AUC for Logistic Regression

Based on Average test set misclassification error in logistic regression to predict the SOFAS logistic outcome, sensitivity and specificity of the ROC (Receiver Operating Characteristic) curve and AUC (Area under the curve) are used to understand the model performance. Based on moderately high value for AUC, we conclude that there is statistical significance between SOFAS logistic score and brain scan data.

Algorithm

Supervised Algorithm

Bayesian Linear Regression with Laplace Prior(Elastic Net) We choose a Laplace prior for the parameter \( \theta \) as Laplace distribution is symmetric around zero and it is more strongly peaked as \( \lambda \) grows. The MAP estimator is the sparse lasso solution, this is useful to pinpoint the exact Brain attribute to HDRS/SOFAS score which will establish the functional connectivity between antidepressants and the specific brain region, because some of the \( \theta \)'s goes to zero.

\[
\text{Laplace Prior} : p(\theta) = \frac{1}{2\lambda} \exp(-\frac{|\lambda|}{\lambda})
\]

\[
\text{Dataset} : S = (x_i, y_i)_{i=1}^m
\]

We search for a choice of \( \theta \) that minimizes the objective function

\[
J(\theta) = \frac{1}{2} \sum_{i=1}^m (\hat{y}_i - y_i)^2
\]

The output of Bayesian linear regression on a new test point \( x \) is the posterior predictive distribution

\[
p(y|x, S) = \int p(y|x, \theta)p(\theta|S)d\theta
\]

Results: Factor Analysis

In Factor Analysis, we transform the current set of variables into an equal number of variables such that each new variable is a combination of the current ones through some transformation. Here data gets transformed in the direction of each eigenvector and represent all the new variables or factors using the eigenvalues. An eigenvalue more than 1 means that the new factor explains more variance than the original variable. Output of our Factor Loadings shows that all 11 feature variables(3 antidepressants, age, gender, education, 5 brain scan attributes) adequately represent the factor categories for this medical data set.

Future Work

We would like to enhance our Gaussian Mixture Model with regression and sparsity as follows: instead of estimating the \( \mu_k \) for \( k = 1, K \), we would estimate only the coefficients of a sparse linear combinations of the \( \chi_s \) for all the data belonging to the same cluster using a sparsity enforcing penalty like \( l_1 \) norm of the coefficients. The main difficulty with such an approach might be to choose the right sample vector representing each cluster a priori, we would like to use Lasso as one of the potential approach to solve that problem.

References

2. MRI preprocessing, classification and pattern recognition: https://arxiv.org/abs/1804.10167

Table 1: RMSE for training and test data

<table>
<thead>
<tr>
<th>Method</th>
<th>Test RMSE</th>
<th>Training RMSE</th>
</tr>
</thead>
<tbody>
<tr>
<td>Logistic Regression</td>
<td>3.4331</td>
<td>3.8686</td>
</tr>
<tr>
<td>Ridge Regression</td>
<td>3.4244</td>
<td>3.6300</td>
</tr>
<tr>
<td>Lasso</td>
<td>3.5453</td>
<td>3.6862</td>
</tr>
<tr>
<td>Elastic Net</td>
<td>3.5195</td>
<td>3.7181</td>
</tr>
</tbody>
</table>