Project Title
CS 598 PS - ML in Signal Processing

Ryley Higa
higa2@illinois.edu

Christian Howard
howard28@illinois.edu

Sameer Manchanda
manchan2@illinois.edu

Abstract

A set of classifiers were built that could be used to predict if an individual has Major Depression given fMRI data about them performing a set of tasks. To accomplish this task, a dense fMRI dataset was coarsened via a local voxel averaging scheme and then features were discovered using Principal Component Analysis (PCA) and Non-negative Matrix Factorization (NMF). Using the new feature representations, classifiers were constructed using the following approaches: Random Forests, Gaussian Discriminant Analysis (GDA), and a Deep Neural Network (DNN). All of the above classification methods had high testing accuracy. This high accuracy was interesting and appears to stem from the feature representation causing the reduced data to be very separable with respect to their classes.

Contents

1 Project Background

1.1 Dataset Description

The dataset used within this analysis was obtained from OpenfMRI. It contains data from 19 individuals with Major Depressive Disorder and 20 individuals with no mental illness. fMRI data was then collected for all these individuals while they did a set of tasks related to listening to different clips of audio. The resulting dataset is a time series information of Blood Oxygenation at a dense point cloud of locations within the patients brains.

1.2 What is fMRI Data?

fMRI stands for Functional Magnetic Resonance Imaging. The idea of how this works is that tissue is placed within a strong magnetic field. To help capture a picture of the tissue, radio waves are sent into the tissue. Blood that is deoxygenated will then reflect the waves in a manner that can be processed by sensors and algorithms to generate a fMRI dataset. Figure 1 shows example visualizations of some subset of the fMRI data used within this problem.
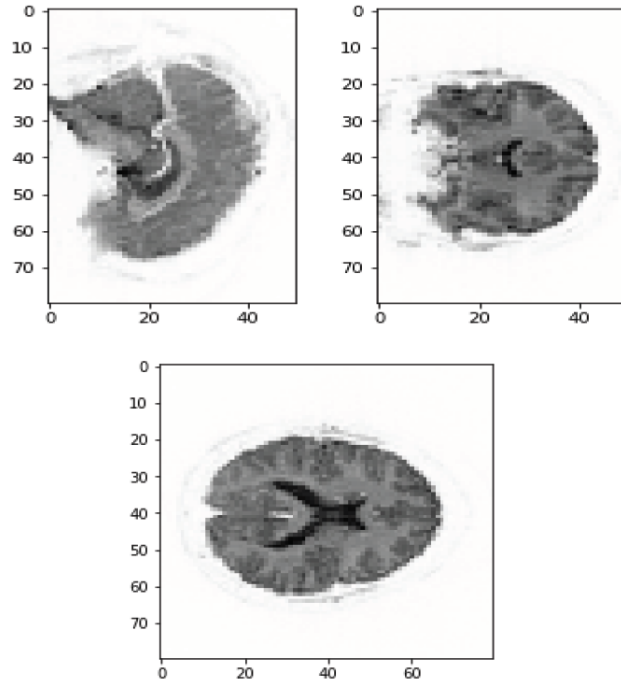

Figure 1: Image slices of the Intensities of Blood Oxygenation levels of some brain at the initial time

1.3 Problem Being Investigated

2 Problem Solution

2.1 Unsupervised Learning

2.2 Supervised Learning

2.2.1 Gaussian Discriminant Analysis

As a fairly straight forward model for classification, Gaussian discriminant functions with full covariance estimates were used. The idea is to perform a Maximum a Posteriori (MAP) estimate of which class ω^*, out of a set of classes $\{\omega_k\}$, is best representative given some input x. This can be expressed as:

$$\omega^* = \arg\max_{\omega_k} g_k(x)$$

where $g_k(\cdot)$ is the discriminant function for class ω_k. For a full Gaussian estimate, $g_k(\cdot)$ can be readily written as:

$$g_k(x) = a_k + b_k^T x + x^T C_k x$$
$$a_k = \log P(\omega_k) - \frac{1}{2}\log|\Sigma_k| - \frac{1}{2}\mu_k^T \Sigma_k^{-1}\mu_k$$
$$b_k = \Sigma_k^{-1}\mu_k$$
$$C_k = -\frac{1}{2}\Sigma_k^{-1}$$

where μ_k and Σ_k are the expected value and covariance for x given we are in class ω_k. Using this representation, a classifier can be efficiently learned and used to obtain quadratic decision boundaries.

2.2.2 Random Forests

2.2.3 Deep Neural Networks

2.3 Numerical Results

After performing the necessary coarsening and dimensionality reduction, a set of models were built for predicting if someone has Major Depressive Disorder based on their fMRI data. All models performed well but, as Figure 4 shows, their success for a given number of features was dependent on the modeling approach.

Random Forests proved to not only perform well for the provided NMF features, but was very fast to compute. GDA performed well using PCA feature as the number of PCA features increased and was also very efficient to compute. Now while the Deep Neural Network required more computational complexity
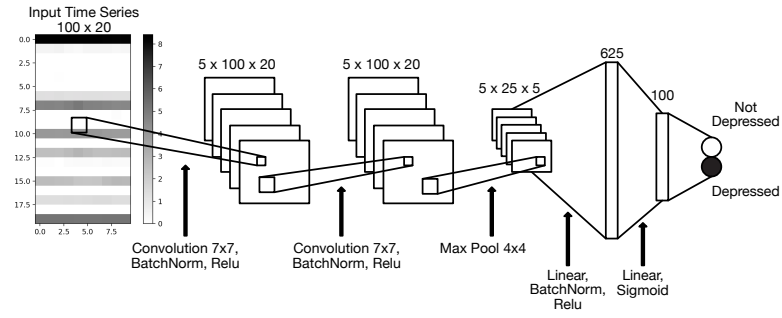
Figure 2: Deep Neural Network Configuration

to train, it showed great performance as well. As can be seen in Figure 4, GDA required the most number of features to achieve greater than 90% classification performance while the Random Forests and Deep Neural Network needed only 5 or more features to achieve high accuracy.

References

[1] Gene H. Golub, Charles F. Van Loan. *Matrix Computations 4th Edition.* The John Hopkins University Press, Baltimore, Maryland, 2013.

[2] Henrik V. Sorensen, C. Sidney Burrus. *Efficient Computation of the DFT with Only a Subset of Input or Output Points.* IEEE Transactions on Signal Processing, Volume: 41, Issue: 3, pgs. 1184 - 1200, Mar 1993.

Appendices