Finding Regions of Interest in Cognitive States during Different Activities and Training a Neural Network to Classify those Activities using Cognitive States

Final Report

Yashasvi Asthana

April 30, 2020

Contents

1. Introduction

When a person takes an action (physical or psychological), they go through multiple sets of cognitive states. This is also called the “process of thinking”. In this project, the main goal is to find those hidden cognitive states and then use them to classify what the person is thinking. The StarPlus fMRI dataset consists of fMRI scans of subjects going through a strict experiment. The experiment consists of multiple trials of understanding the sentences and relating them with the displayed images.

2. Visualizing fMRI data

Each fMRI scan is in the shape 64 x 64 x 8. This means that there are 8 vertical slices of 2 dimensional brain scans making the scan 3 dimensional. All the voxels are mapped to a particular coordinate in this space. These voxels have different activations based on the activity of the person.

*Video of this data can be created using the submitted code.

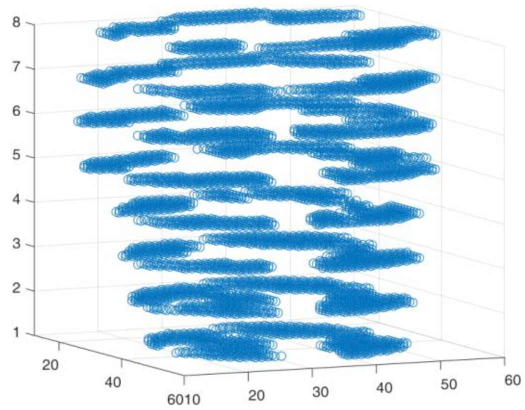

Figure 1: Voxels without activations

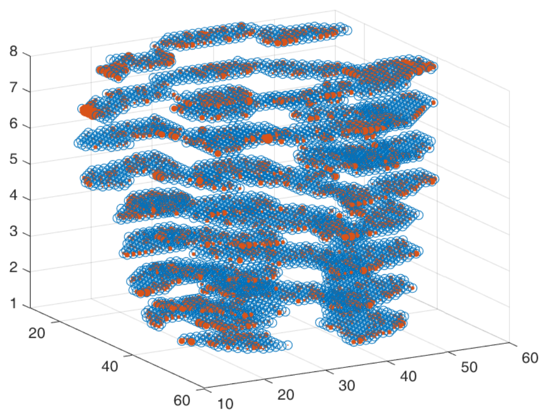

Figure 2: Voxels with activations

3. Creating Data Matrix

We label the dataset based on 3 different classes – subject staring at a fixed fixation point, subject viewing a picture and subject viewing a sentence.

Based on the subject 1, there were 132 frames of the subject staring at the fixation point and 308 frames each of the subject viewing a picture and a sentence.

4. Finding Number of Hidden Cognitive states

We will use K-means clustering algorithm with Silhouette analysis to find the most separable cognitive states in each class separately.

We will ignore the high scores at K=2 as we want to find hidden states and not just give the 2 'not thinking' and 'thinking' states. So we will look for sudden spikes in the graphs below.

4.1. Staring at Fixation Point

Here, we want to find the number of hidden cognitive states that exists when the subject was staring at the fixation point.

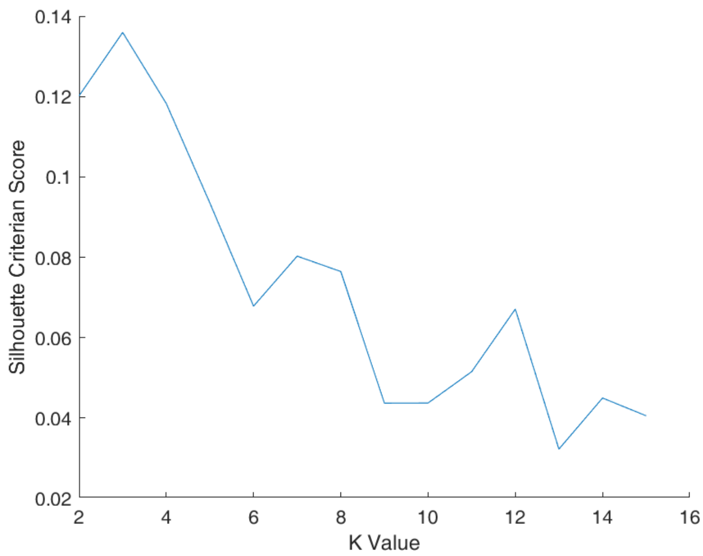

Figure 3: K Value vs. Silhouette Criteria

Based on the Silhouette analysis, we can say that there can be around 3 hidden cognitive states while the subject is just staring at the fixation point.

4.2. Viewing a Picture

Similarly, here we want to find the number of hidden cognitive states that exists when the subject was viewing a picture.

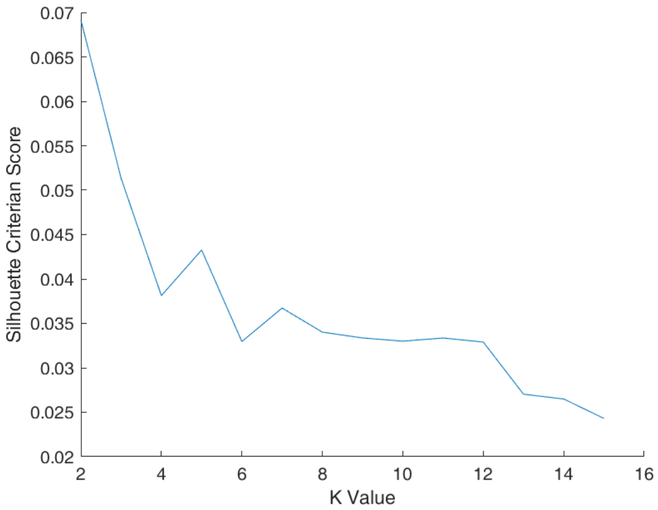

Figure 4: K Value vs. Silhouette Criteria

Based on the Silhouette analysis, we can say that there can be around 5 hidden cognitive states (due to the sudden spike) while the subject is processing a picture.

4.3. Viewing a Sentence

Now, we want to find the number of hidden cognitive states that exists when the subject was viewing a sentence.

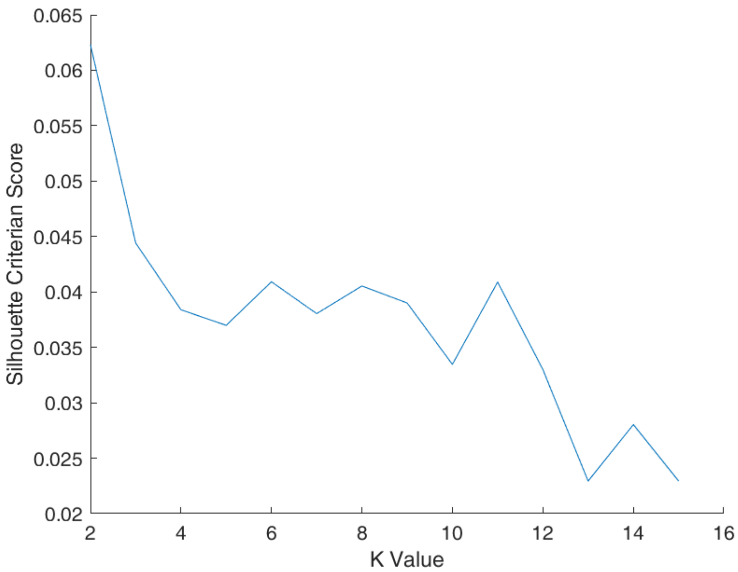

Figure 5: K Value vs. Silhouette Criteria

Based on the Silhouette analysis, we can say that there can be around 6 or 11 hidden cognitive states (due to the sudden spikes 2 times) while the subject is processing a sentence.

5. Classification

As we have divided the dataset into 3 separate classes – subject staring at the fixation point, subject viewing a picture and subject viewing a sentence, we can try to train a model to classify them based on the activations in the voxels.

5.1. With all the Features

This is the baseline. We train 2 different models on all the features:

Support Vector Machine

SVM tends to perform better than other linear learners with low training samples.

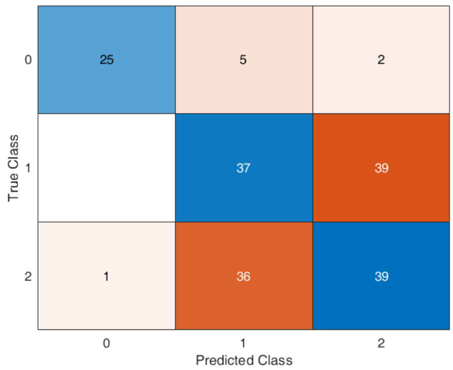

Figure 6: SVM Baseline

The train accuracy was 100% and the test accuracy was **55.38%**.

Shallow Neural Network

The Neural Network consists of 1 hidden layer with 20 neurons. We had to keep the neural network very shallow due to the lack of training data.

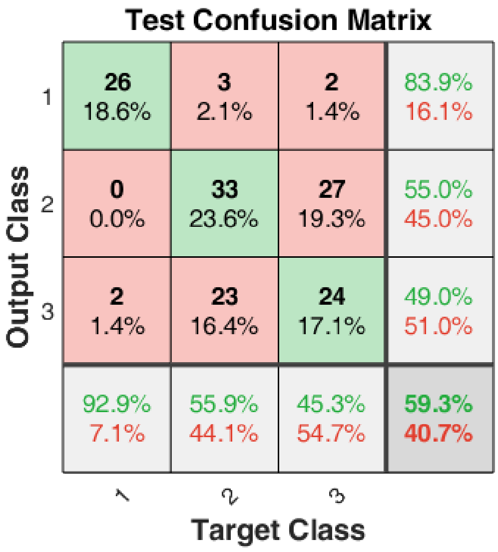

Figure 7: Neural Network Confusion Chart Baseline

The training accuracy was 94.3% and the test accuracy was **59.3%.**

5.2. With suggested Regions of Interest

These Regions of Interest were suggested by the journal published by CMU while analysing this dataset - CALC, LIPL, LT, LTRIA, LOPER, LIPS, and LDLPFC.

Using these ROI, we find all the voxels that are part of these regions and use them as features.

Support Vector Machine

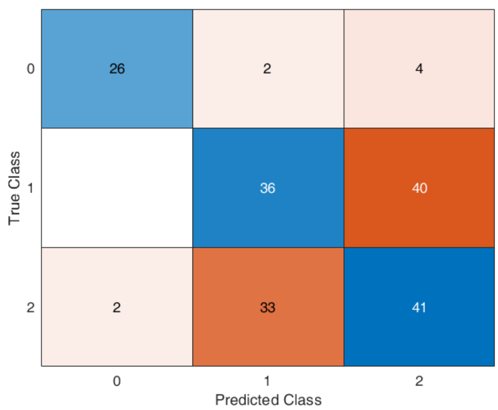

Figure 8: SVM Confusion Chart with ROI

The train accuracy was 100% and the test accuracy was **60.86%**.

Shallow Neural Network

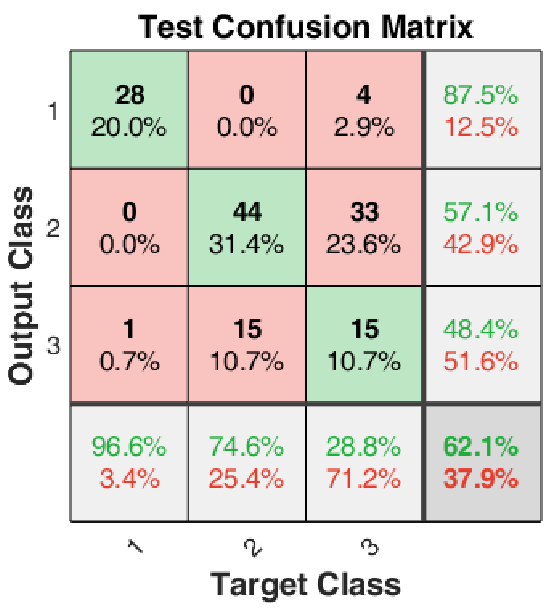

Figure 9: Neural Network Confusion Chart Baseline

The training accuracy was 86.3% and the test accuracy was **62.1%.**

5.3. With Feature Selection using Variance Ratio

Variance Ratio score was used for feature selection. The Variance Ratio score for a feature F is given by:

$$VR(F) = \frac{Var(S_F)}{\frac{1}{C}\sum_{k\varepsilon c} Var_k(S_F)}$$

, where *C* is the total number of classes, *c* is the set of classes and $Var_k(S_F)$ is the variance of the subset of values from feature *F* which belongs to class *k*.

The top 2000 features with the highest Variance Ratio score were selected.

Support Vector Machine

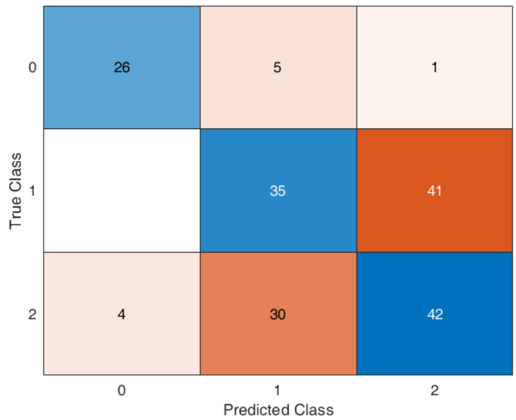

Figure 10: SVM Confusion Chart with selected features

The train accuracy was 100% and the test accuracy was **60.86%**.

Shallow Neural Network

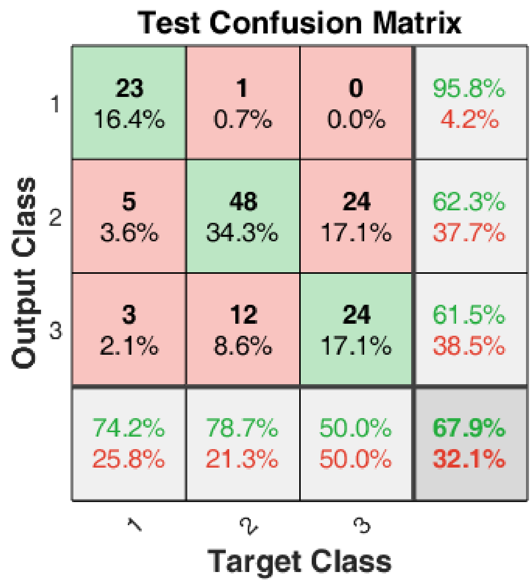

Figure 11: Neural Network Confusion Chart with selected features

The training accuracy was 83.7% and the test accuracy was **67.9%.**

6. Conclusion

In the clustering part, we saw that there are 3 hidden cognitive states when the subject was staring at the fixation point, 5 hidden cognitive states when the subject was viewing a picture and more than 6 hidden states when the subject was viewing a sentence. This shows that there are more hidden states while processing a sentence or a picture. And processing a sentence requires going through more cognitive states than processing a picture. Hence, we can say that processing a sentence is more difficult for humans.

In the classification part, we achieved the highest test accuracy of 67.9% with the Shallow Neural Network after selecting top 2000 features using the Variance Ratio scores. This is better than the 62.1% accuracy with the Neural Network after selecting features based on the suggested Regions of Interest. As SVM performs the same with both the methods of feature subspace selection, we can say that our method of feature selection is better.

References

[1] *StarPlus fMRI data* http://www.cs.cmu.edu/afs/cs.cmu.edu/project/theo-81/www/

[2] Wang X., Mitchell T.
Detecting Cognitive States Using Machine Learning
https://www.cs.cmu.edu/afs/cs.cmu.edu/project/theo-73/www/papers/ XueruiReport-10-2002.pdf

[3] John Ramish *Learning Common Features from fMRI Data of Multiple Subjects*
https://www.cs.cmu.edu/afs/cs.cmu.edu/project/theo-73/www/papers/ JohnRamishReport2004.pdf

[4] Ester, Martin and Kriegel, Hans-Peter and Sander, J¨org and Xu, Xiaowei
A Density-Based Algorithm for Discovering Clusters in Large Spatial Databases with Noise
https://www.aaai.org/Papers/KDD/1996/KDD96-037.pdf

[5] Y. Liu, K.L. Schmidt, J.F. Cohn and R.L. Weaver *Facial Asymmetry Quantification for Expression Invariant Human Identification* https://ri.cmu.edu/pub_files/pub3/ liu_yanxi_2002_2/liu_yanxi_2002_2.pdf