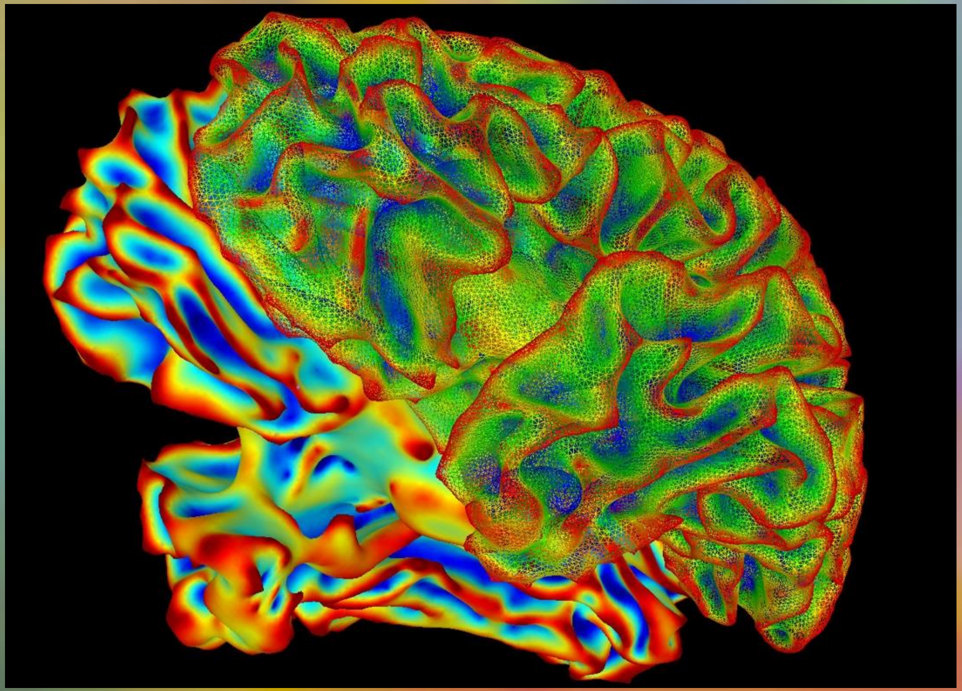
Understanding Mental Disorder Diagnosis for US Service Members



Authors: Yagmur Yavuz Ozdemir and Eyoel Berhane

PROJECT BACKGROUND

- Doctors diagnose patients for a litany of conditions for a living, and often physical ones are clearcut for them when the symptoms fall in well-organized categories.

- Diagnosing mental disorders is often a difficult process even with a mix of classifying certain symptoms and behaviors, for the affects of certain conditions linger longer and often unseen methods.

- MRIs and other tools are common methods for collecting information on the brain, such as brain waves and neural activity, in order to understand such conditions better in a physical manner. Often traumatic conditions on the brain (or affecting it) are profound yet not easy to observe directly.

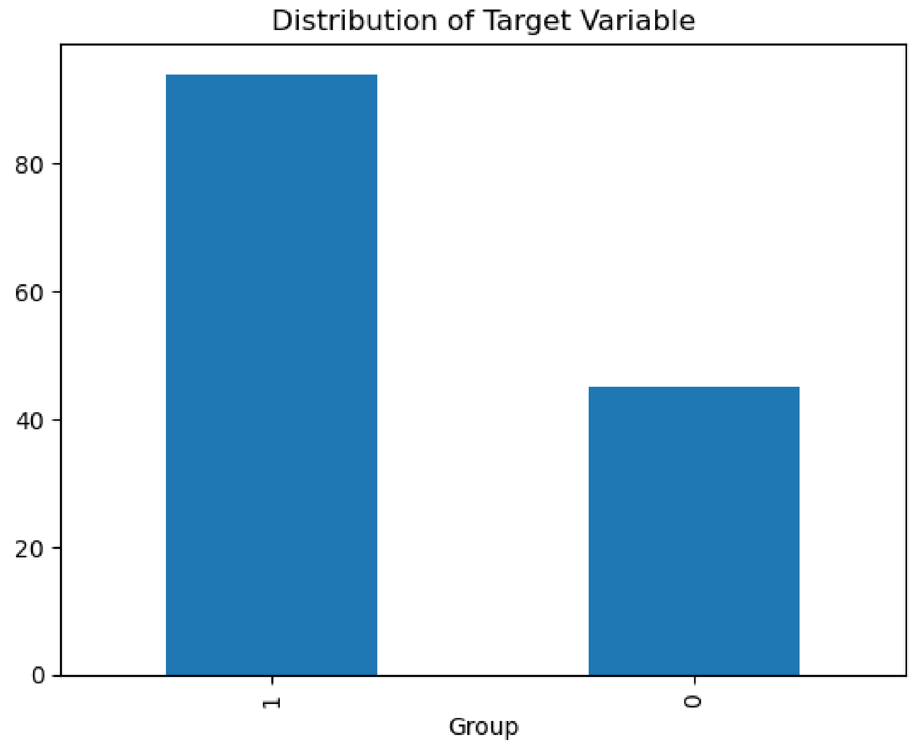
DATASET BACKGROUND

- Signals between two brain regions are obtained from brain images (scanned in Auburn University – MRI Center). **139** active-duty male US Army soldiers were recruited for this study from Fort Benning, GA, and Fort Rucker, AL, USA. All groups were matched for age, race, education, and deployment history.
- There are **677** variables that are all numeric and between -1 and 1 since they represent correlations between the signals of 2 brain regions.

PROJECT PRIORITY

- **Classification techniques:** to automatize the detection of these mental health problems using brain signals.
- Accuracy
- Sensitivity
- Specificity

- **Feature importance:** since this is a high-dimensional data, we would like to extract the important brain regions related to PTSD
- Random forest
- Lasso logistic



Dataset Organization

- We decided that since the symptoms between PTSD and PCS patients are often similar, we decided to classify the dataset as a binary variable between control (0) and the affected (1) groups.
- We split the dataset into an 80/20 training vs validation set for whichever model we would choose.

FEATURE IMPORTANCE

Random Forest :

-The minimum number of samples required to be at a leaf node = 1

- The minimum number of samples required to split an internal node = 5

-The number of trees in the ensemble= 50

-Validation Accuracy = 0.857

Lasso Logistic:

-5-fold CV

-Validation accuracy = 0.949

FEATURE IMPORTANCE

- The following table shows the brain regions of the selected variables by both random forest model and lasso logistic model.

Brain region	How many times it is selected
Frontal	26
Temporal	17
Occipital	7
Cingulum	5
Lingual	5

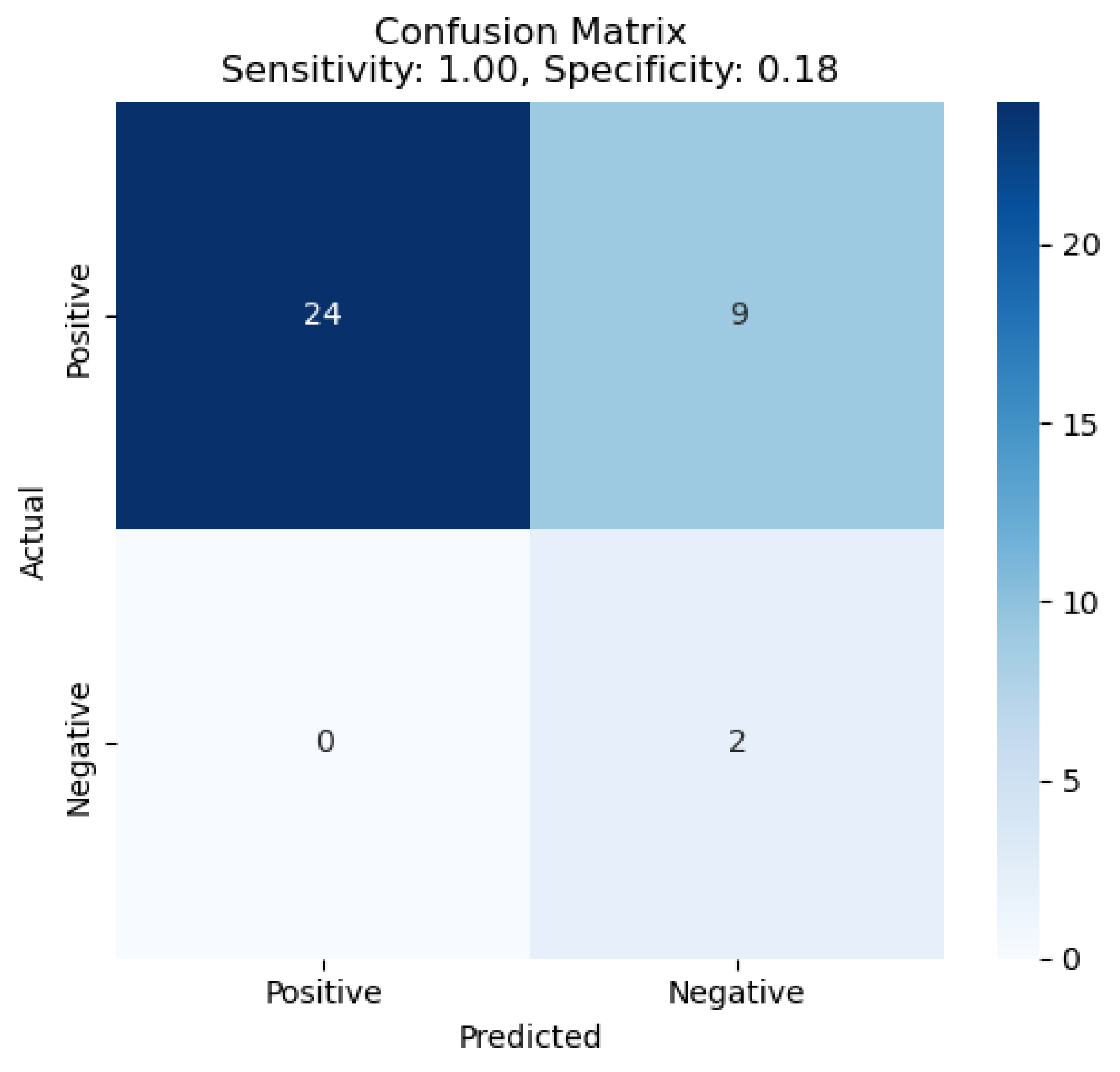
FINAL MODEL

- Since the validation accuracy is higher in Lasso model, we select it as our final model. The following shows the accuracy matrix.

```
Test Accuracy: 0.7429
Classification Report:
              precision    recall  f1-score   support

           0       1.00      0.18      0.31        11
           1       0.73      1.00      0.84        24

    accuracy                           0.74        35
   macro avg       0.86      0.59      0.57        35
weighted avg       0.81      0.74      0.67        35
```



CONCLUSION

- We used a high dimensional brain image data to automate the detection of PTSD and to extract the important brain regions.
- Since the aim is to detect, we aggregate the PCS and PTSD.
- For feature importance purposes, **random forest** and **lasso logistic** are employed.
- The **temporal** and the **frontal** lobes are found to be the most important features as well as **occipital**, **cingulum** and **lingual** which have been reported to have alterations in PTSD before (Yin et al. 2012; Zhang et al. 2016; Liu et al. 2015)
- Lasso logistic has been selected as the final model with a **74%** test accuracy.
- The model has the highest sensitivity which shows perfect detection of the positive cases.

References

- Lanka, Pradyumna, et al. "Supervised machine learning for diagnostic classification from large-scale neuroimaging datasets." *Brain imaging and behavior* 14 (2020): 2378-2416.
- Yin, Y., Jin, C., Eyler, L., Jin, H., Hu, X., Duan, L., et al. (2012). Altered regional homogeneity in post-traumatic stress disorder: a restingstate functional magnetic resonance imaging study. Neuroscience Bulletin, 28(5), 541–549.
- Zhang, Q., Wu, Q., Zhu, H., He, L., Huang, H., Zhang, J., & Zhang, W. (2016). Multimodal MRI-based classification of trauma survivors with and without post-traumatic stress disorder. Frontiers in Neuroscience, 10, 292.
- Liu, F., Xie, B., Wang, Y., Guo, W., Fouche, J.-P., Long, Z., Wang, W., Chen, H., Li, M., Duan, X., Zhang, J., Qiu, M., & Chen, H. (2015). Characterization of post-traumatic stress disorder using resting-state fMRI with a multi-level parametric classification approach. Brain Topography, 28, 221–237.