MACHINE LEARNING-BASED DATA ANALYSIS OF A LONGITUDINAL EEG STUDY OF THE CREATIVE PROCESS

UNIVERSITY of HOUSTON
CULLEN COLLEGE of ENGINEERING
Department of Electrical & Computer Engineering

Abigail Turcheck, Jesus Cruz-Garza, Dr. Jose Luis Contreras-Vidal
aturchec@asu.edu, jgcruz@central.uh.edu, jlcontr2@central.uh.edu
Laboratory for Noninvasive Brain-Machine Interface Systems, Dept. of Electrical and Computer Engineering, University of Houston, Houston TX, USA

NSF REU SITE AWARD #1757949

Arizona State University

INTRODUCTION

- This study has aimed to alleviate the issues of studying the human creative process by using a mobile brain-body imaging (MoBI) dry-electrode electroencephalography (EEG) headset to noninvasively record a Houston-area artist completing a long-term project [1].
- Machine learning has been increasingly used in the field of brain imaging due to its ability to cope with low signal-to-noise ratios [2].
- This application of machine learning aimed to classify the different activities of the artist within the large dataset.

TOTAL DATASET

The artist's activities were identified with class labels from video footage and include:

1: research in office
2: reading papers
3: research outside office
4: prototyping
5: scent training
6: walking
7: walking on treadmill

Artist prototyping

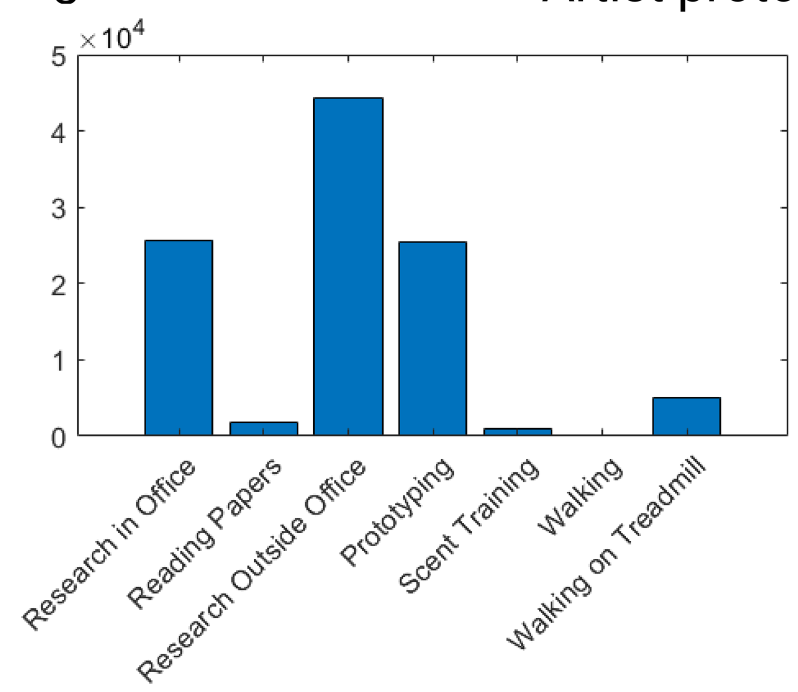

Figure 1: Histogram showing the number of samples per class vector.

"Reading papers", "scent training," and "walking" were removed from the dataset used for classification due to not having enough samples for an effective model.

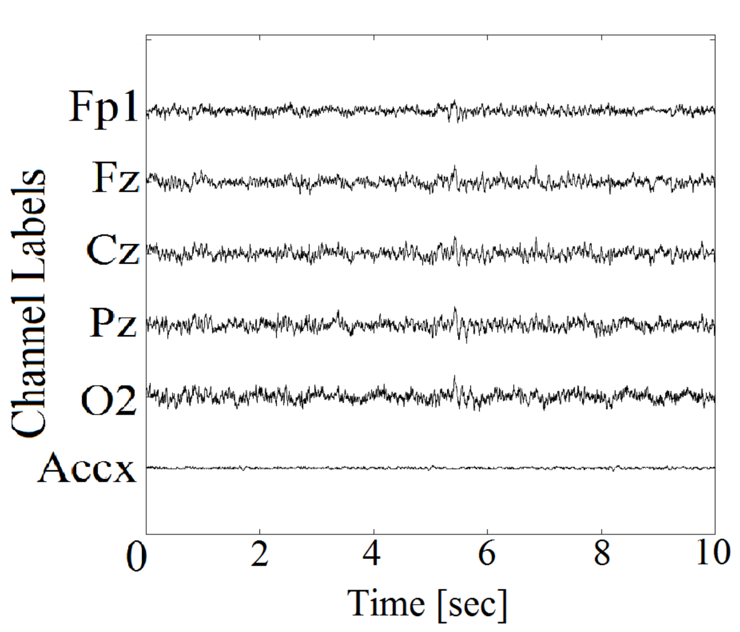

Figure 2: Example set of EEG data for a 10-second time interval.

DATA ANALYSIS PROCESS

- Precautionary measures were taken to avoid data leakage, i.e. separating the entire data files into training/validation and testing files prior to the classification process.
- To keep analysis time efficient and accurate, random reiterative resampling of the data was used to apply to the machine learning model.

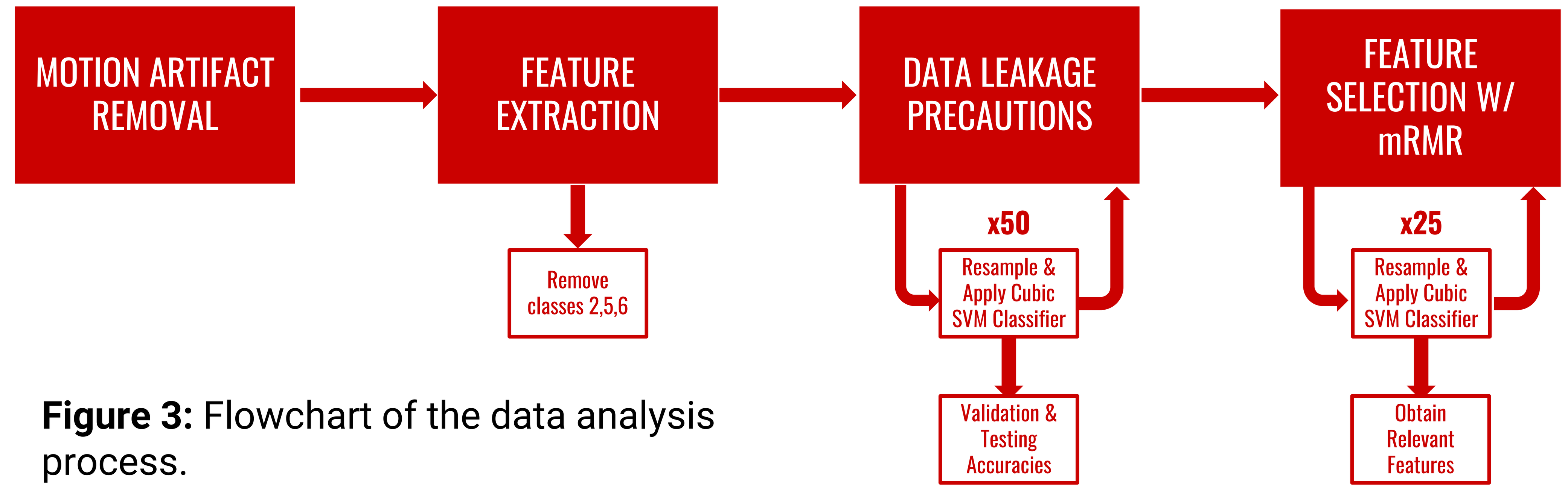

Figure 3: Flowchart of the data analysis process.

FEATURE MATRIX

The feature matrix was created by extracting different frequency bands from the EEG channels.

F7	P7
Fp1	Pz
Fp2	P4
F8	T3
F3	P3
Fz	O1
F4	O2
C3	C4
Cz	T4
P8	

Delta	1-4 Hz
Theta	4-8 Hz
Alpha	8-12 Hz
Beta	12-30 Hz
Gamma	30-40 Hz

Figure 4: Channel names and frequency bands used for feature extraction. With 19 channels and 5 frequency bands, 95 total features were extracted.

The 95-feature matrix was created with 4 sec. windows and 2 sec. overlap. This matrix was then used for resampling and creation of a machine learning model.

After reiterative sampling and application of the cubic SVM machine learning model, average validation and testing accuracy was found.

Average Validation Accuracy: .7479
Average Testing Accuracy: .5083
St. Dev. Valid. Acc. = .0135
St. Dev. Test. Acc. = .0358

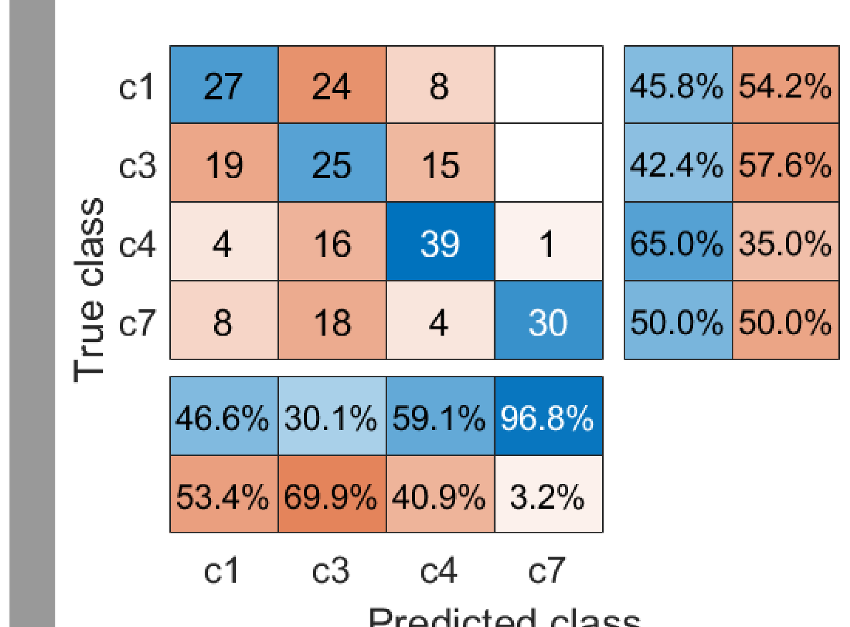

Figure 5: Average confusion matrix of testing data after 50 iterations.

MODEL ACCURACY

Relevant features were selected with a maximum relevance-minimum redundancy (mRMR) algorithm. This chooses the most relevant features to the class vector and subsequently chooses other features with the least redundancy.

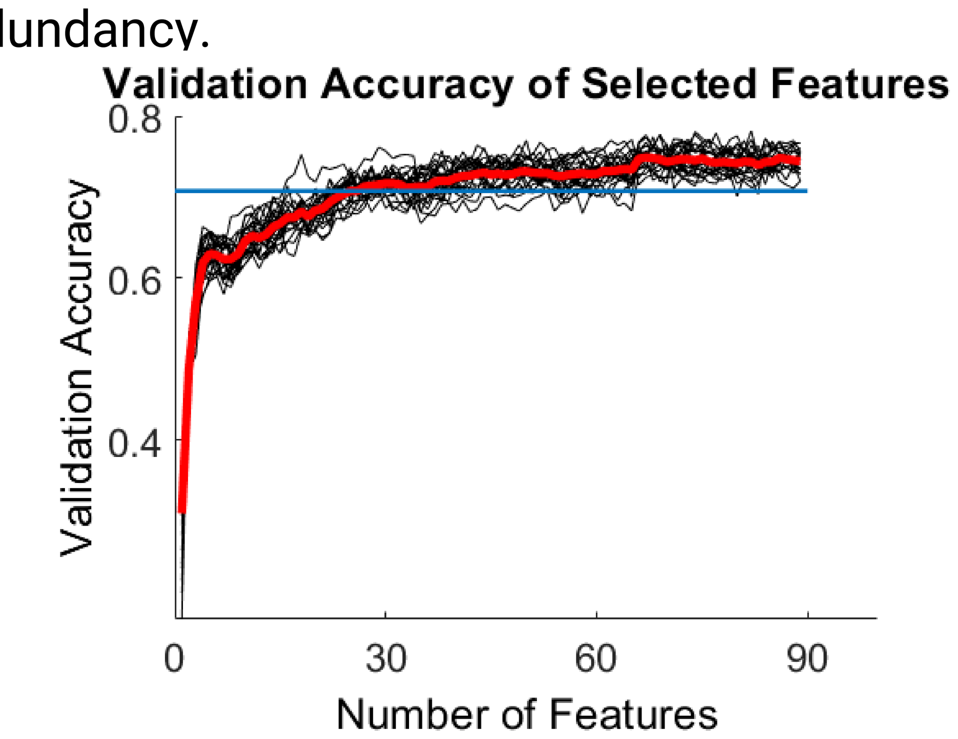

Figure 6: Validation accuracy of model over 25 resampling iterations of feature selection.

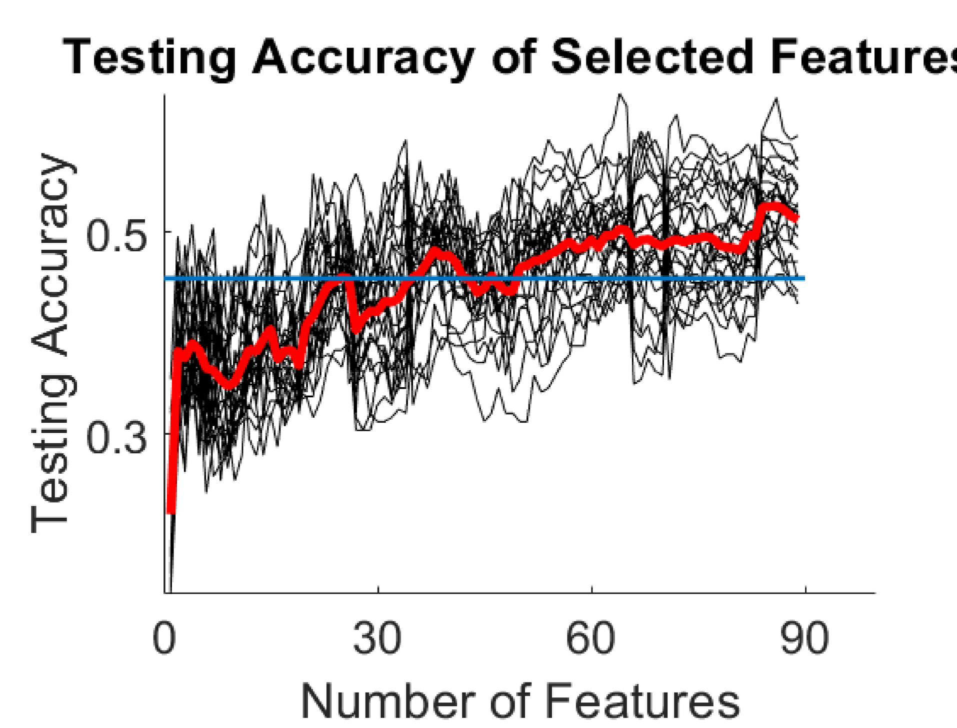

Figure 7: Testing Accuracy of model over 25 resampling iterations of feature selection. The number of relevant features was found to be **26**, with validation accuracy of **.7075** and testing accuracy of **.4545**.

SELECTED FEATURES

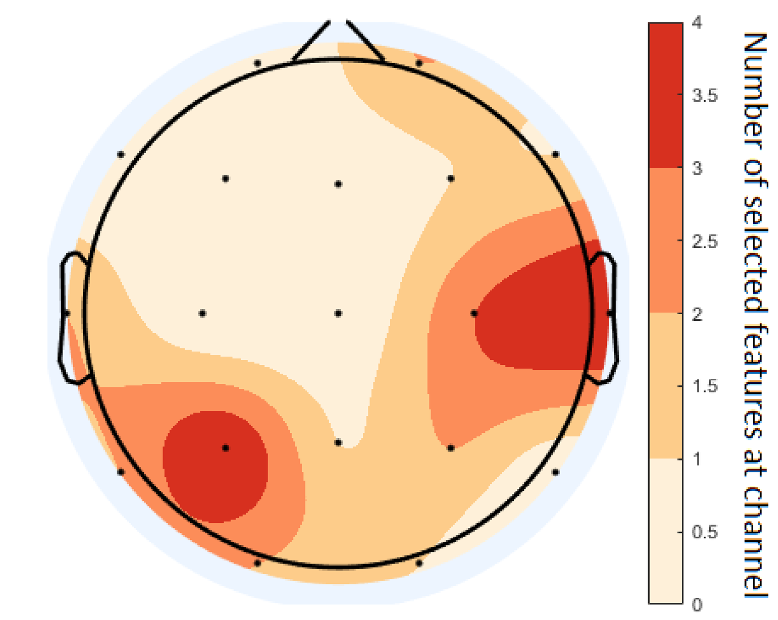

Figure 8: Scalp map of channel location power in the selected features indicates higher activity in the right temporal, right parietal, right frontal, and left parietal lobes.

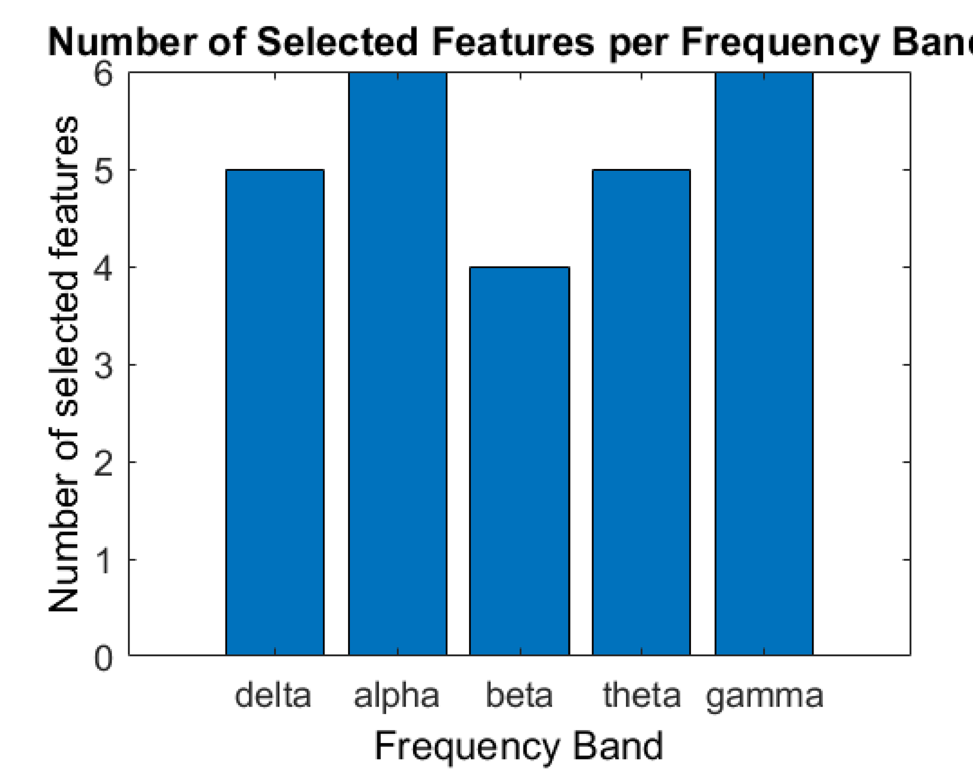

Figure 9: Frequency band power in the selected features indicates relevance of all frequency bands in classification.

CONCLUSION

Right temporal lobe function is associated with memory skills with nonverbal material. Both parietal lobes are associated with integrating sensory information, especially for spatial recognition. Right parietal lesions can lead to problems with construction of objects and drawing [4]. These functions are highly associated with the selected class vector activities of walking and prototyping. The results also show that all frequency domains are relatively equal in relevance to classification.

It was possible to classify creative tasks solely based upon EEG data collected in a real-world setting.

REFERENCES

[1] Alarcon, C., Bellman, D., Contreras-Vidal J.L., & Cruz-Garza J.G. (2019).A longitudinal mobile brain-body imaging (MoBI) study of the creative process over the span of 18 months in real-world settings. Manuscript in preparation.
[2] Lemm, S., Blankertz, B., Dickhaus, T., & Müller, K. R. (2011). Introduction to machine learning for brain imaging. *Neuroimage*, *56*(2), 387-399.
[3] Smialowski, P., Frishman, D., & Kramer, S. (2009). Pitfalls of supervised feature selection. *Bioinformatics*, *26*(3), 440-443.
[4] Warrington, E. K., & Taylor, A. M. (1973). The contribution of the right parietal lobe to object recognition. *Cortex*, *9*(2), 152-164.