Final Report: Analysis of EEG Characteristics in Identifying Congruent Modal/Lexical Stimuli

Emmanuil Simkhayev

CSC 59867 Senior Design II

Professor Michael Grossberg

December 21st, 2018

Abstract

Identifying where and when decisions in the brain are made is something that neuroscientists aim to understand. Using data from a study by Dave Britton at the City College of New York, which evaluated the effects of interference in the form of word and non-semantically significant stimuli, we were able to find signals in the brain within the spatial and time domains. Applications of machine learning were used to tackling classification problems along with the non-trivial task of finding these signals.

As data scientists, we experienced the struggles of pre-processing data to get it in an acceptable format for applying machine learning. These included removing artifacts, creating custom events for our purposes, and generating epoched data. Using the python library MNE, manipulating EEG data made this process more manageable.

The two main challenges tackled had to do with auditory versus visual, and lexical versus non-lexical binary classification. Our results were more reliable for the auditory vs visual problem compared to those of the lexicality problem. Plots of mappings to the brain made analysis of these results easy to understand and draw conclusions from.

Introduction

Identifying the effects of interference on an individual's reaction time can be traced back to the experiment performed by John Ridley Stroop in 1935[1]. For example, it showed that reading the word red in the color a blue causes interreference. This idea can be further extended by performing an analogous experiment using picture and word interference (PWI). A subject can be shown a picture of a cat with the written word dog over it to distract them. The results from PWI closely match those of the Stroop Effect [2].

The work in this project was based on research done by David Britton, a psychologist at the City College of New York. His study investigated brain activity and behavioral responses associated with semantically meaningful stimuli. Interference was in the form of the Stroop Effect, where auditory and visual modalities along with word and non-word semantically significant stimuli were both met with distractors. These distractors were bicategorical, with having the same audio-visual, and word-nonword stimulus relationships that differed in semantic content [3]. The goal was to use this paradigm defined by Britton and discover characteristics of EEG that help identify where in the brain semantically congruent stimuli is recognized. It is useful to know where and when in the brain this activity is present.

This problem was tackled using well known machine learning algorithms. More importantly, we were able to leverage results from classifiers to paint a picture of where the reaction times in the brain were most prevalent. Along with that, labeling of data into small problems such as audio versus visual stimuli allowed us to see differences in reaction times between these smaller problems.

Background

Integrated information from different modalities are crucial for information processing [4]. The convergence of information from multisensory input enhances behavioral performance (e.g. speed and accuracy) due to the increase of certain neurons activating (firing) together. The firing rate of cells of multisensory inputs far exceeds firing rate of unisensory inputs. This means that humans can realize a kind of input and storing information when we have more than one sensory organ being used (such as seeing and hearing) Semantic Congruence refers to the combination of multisensory stimuli that are presented in terms of the same meaning. The impact of multisensory input has been investigated in several studies [5,6] and significantly faster times were found in semantically congruent audio-visual pairings compared with unisensory input. Significant longer times were found for mismatched incongruent audio-visual pairings. Thus, semantic congruence has a significant impact on the integration of information across different modalities.

Two visual-audio bimodal stimulus (VABS) systems were conducted on a study [7]. The study was comparing congruent and non-congruent VABS based model. The paper aimed to see if semantically congruent stimuli can get the same performance as incongruent stimuli in Brain computer interface (BCI) systems. The results showed higher amplitude of the Event Related Potential (ERP) of semantically incongruent non-target and target stimuli. The acquired raw EEG data was preprocessed for artifact removal through Independent component analysis(ICA). Data was than filtered through Band pass of 0.5-40Hz and down sampled to 200Hz. The event related potentials of three channels (target ERPs and distractor ERPs) were calculated and averaged for each condition (congruent and incongruent).

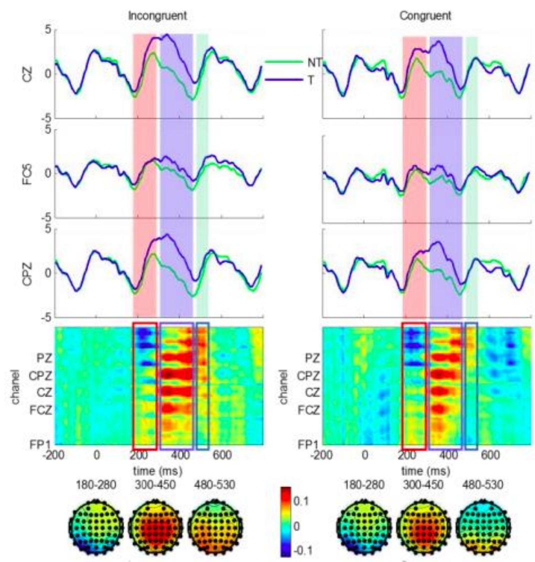

Figure 1: Readings shown between 200ms and 500ms where the highest-class discrimination was noticed. The radical change in color from neutral green shows that either ERP target amplitude was higher than ERP distractor or vice versa.

They plotted the ERP of each channel (CZ, FC5 and CPz) for each condition (congruent and incongruent) with two colors representing the target stimuli and distractor stimuli. The spatial temporal distribution of class discrimination (target and distractor) were shown for both conditions. The scalp map was shown depicting the average sign value for three specific time intervals. A Positive value represents that target ERPs have larger amplitude than distractor ERPs. Whereas a negative value represents the opposite. Support Vector Machine (SVM) was used to do binary classification of target and non-target. The illustration of

Fig 1 demonstrates that higher class discrimination values are lying at channels around FCz, Cz, CPz, and Pz with the time interval from 200ms to 500ms approximately.

From this research we can learn to visualize the patterns associated with each target ERPs and distractors ERPs so that we can find evidence to differentiate characteristics of EEG. This will help us classify the type of information contained in the EEG (e.g. visual, audio, language, non-language), and find sufficient evidence in the EEG that shows brain has understood semantic congruent stimuli was seen. We can see from figure 1 that there is no significant evidence to classify congruent and incongruent from the results because results are similar. The research explains that they might not have had enough data to get better results. We aim to make sure that our data is sufficient enough and replicate this experiment to acquire better results.

Methods

One of the most important tasks in this project was to preprocess the raw data given to us by Dave Britton. Our environment for machine learning was a powerful library called MNE-Python. This allowed us to easily load and manipulate EEG data to what was required to solve a problem. The personal contribution for this project was heavily focused in preprocessing the data for the team to use. The main tasks of preprocessing included artifact removal, generating events and their labels, and converting raw continuous EEG into epochs of 500ms intervals.

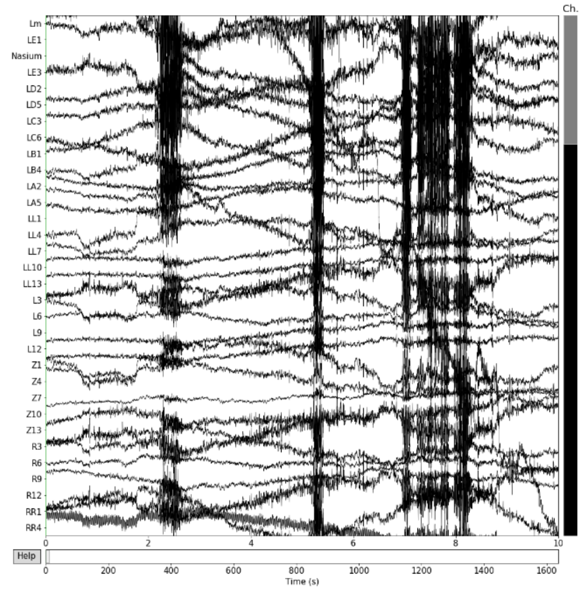

Figure 2 : Raw continuous EEG data of 30 out of the 129 total channels over 10 seconds.

Figure 1 shows EEG data over 10ms and is full of artifacts. Data useful for solving our problem requires to remove these. The method used followed three steps: notch filter, frequency filter, and Independent Component Analysis (ICA).

Removing artifacts from power - line noise created by electricity is a commonly addressed through the notch filter. Based on geographical location, it removes frequencies of multiples of 50Hz or 60Hz [8]. The frequency filter was used to filter out frequencies too low that were less than 5Hz and those greater than 100Hz. ICA finds groups of components that contribute to artifacts and removes them. These three steps generated cleaner EEG data that was acceptable to use for solving our problems.

The raw data we worked with was for 32 subjects, with 1,280 trails for each. Two files were generated during experimentation, one with continuous EEG data for 128 channels at a sampling rate of 512Hz, and the other a consisting of time-stamped triggers (events). Each trial can be broken up into three 500ms intervals, where a subject is firs shown a distractor, a target, and the same distractor once again. After pre-processing for artifacts, epochs of 500ms (257 samples) were created to indicate sections of trials. Because the events were time-stamped, it was simple to figure out which events belonged to which trial. The data compressed took up 11GB of storage, but when loaded into python, this more than doubles in size.

The task of labeling the data was

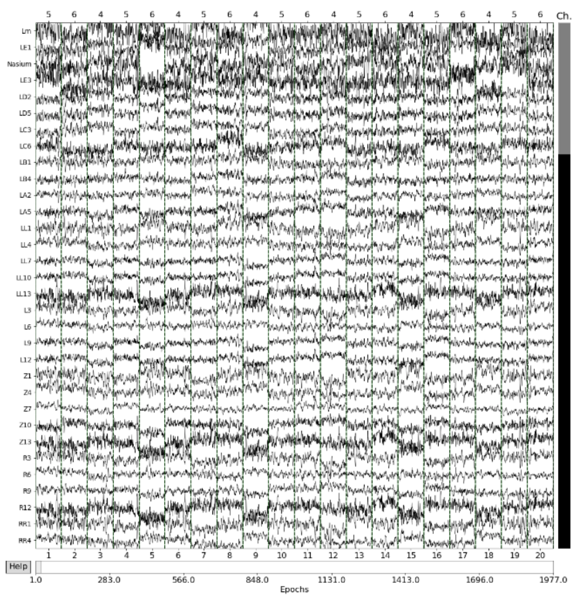

Figure3: Preprocessed data in 500ms epochs. Most artifacts are removed as compared to graph in figure 1.

a challenge due to the complexity of the experiment. Based on figure 2, there are nine events that correspond to a trial. Each event is time-stamped. Those that are labeled as *stim-code* indicate the beginning of a 500ms epoch of a trial. The other events are known describers. For example, *Accuracy* tells us if the

Event ID	Stim Combination Name	Meaning
112	A-AL-L	Auditory - Spoken word - Word
212	A-AL-N	Auditory - Spoken word - Non Word
312	A-AN-L	Auditory - Sound - Word
412	A-AN-N	Auditory - Sound - Non Word
512	A-VL-L	Auditory - Spelled Word - Word
612	A-VN-N	AudItory - Picture - Non Word
712	V-AL-L	Visual - Spoken Word - Word
812	V-AN-N	Visual - Sound - Non Word
912	V-VL-L	Visual - Spelled Word - Word
1012	V-VL-N	Visual - Spelled Word - Non Word
1112	V-VN-L	Visual - Picture - Word
1212	V-VN-N	Visual - Picture - Non Word

Figure 4: Stim code combinations that were assigned to epochs during classification.

button click was correct or not. Individually, the stim codes contain a useful amount of information. The stim code events (labeled 1 – 24) associated with the *Flanker stim-code 1, Target stim-code 1,* and *Flanker stim-code 2* would be combined to create one of twelve unique events that indicate the type of stimulus shown to a subject. These labels could then be used for the following classification problems:

1. Identify modality of the given stimuli (Auditory versus Visual)
2. Identify the Lexicality of the given stimuli (Language versus Non-Language)

Solving classification problems required setting up a pipeline that was appropriate for EEG data. To begin, the shape of our data is 3D, with the first being number of epochs, then channels, and finally the time in number of samples (512 samples per second). To be used in a binary classifier such as logistic

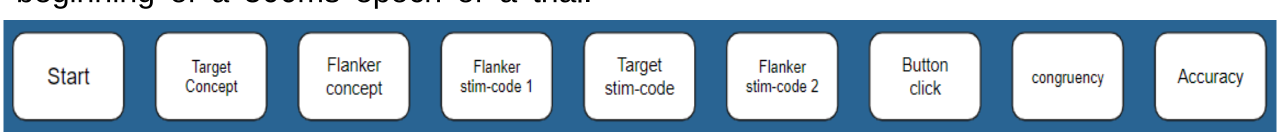

Figure 5: The nine events that make up a trial. For the purposes of the problem we attempted to solve, the three stim-codes were used.

regression, the data required to be in the form of *n samples x n features.* A vectorizer was used to convert our data into the correct dimensions. The pipeline continues with scaling the data with a *Standard Scaler,* which standardizes features by removing the mean and scaling to unit variance [9]. The final step is to fit the data with a classifier. Using This project mainly relied on Logistic Regression and Random Forest classifiers.

To classify audio vs visual stimuli, the we had to match event ids of the stim code combinations in figure 3 to audio and visual labels. The combinations that started with an ***A*** were labeled as audio, and those that start with ***V*** were labeled as visual. This is known as a binary classification, thus relabeling of epochs to indicate a 0 for audio and 1 for visual was done. Due to computational limits, we could not classify more than 5GB of data. Files were randomly selected and fit with a Linear Model Logistic Regression classifier. MNE-Python has libraries that extended those of Scikit-Learn to make this process simple. The shape of the data was ~16,000 epochs, 125 channels (We used 125 channels because the file containing electrode locations necessary to plot our results was incomplete), and 257 samples. The most useful part of the classifier were the coefficients returned. They allowed us to visually understand when and where in the brain the classifier was making its decisions to classify. Maps of a head with electrode positions were used to do this. Since, this is a binary classifier, two colors along with a neutral white were used to represent these results.

Vectorizing the data was not the only method we had available. A sliding estimator is a useful tool that classifies data over the time dimension. For our purposes, we were able to do this for each sample (~2ms window). Analysis of this was available through an ROC AUC curve for each point in time in milliseconds. This can tell us when classification scores were highest, and thus indicating when the brain provides the most information to distinguish for the labels of audio and visual.

This sliding estimator also returned coefficients that could be used to visualize activity in the brain over time. Using an Event Related Potential (ERP), an average of the all epochs, we can generalize a response in the brain.

The important part of this pipeline is that it can be used for more than just the modality problem, because the problems we are solving depend on the labeling of the data. As mentioned earlier, the other problem we attempted to solve was word versus non-word classification.

Results and Evaluation

Experimentation results were based on Logistic Regression and Random Forests. These two classifiers were ideal for us because they offered coefficients that allowed us to analyze results more profoundly. Although classification scores were not impressive in the mid 70’s and 80’s, we were able to identify signals in the brain spatially and over time. The table below in figure 5 shows average scores across multiple 5GB epoch files, between the two classifiers used. As expected, the random forest did perform better, but was computationally more expensive and took longer to complete.

Classifier	Average Score Audio vs. Visual	Average Score Language vs. Non-Language
Logistic Regression	63%	55%
Random Forests	69%	60%

Figure 5: Average classification scores using logistic regression and random forests classifiers. Five separate files of ~16000 samples were classified with 5-fold cross validation, and their scores were averaged into these numbers.

A single number to describe the performance of a classifier is not ideal for our purposes, so further analysis of results was done. For both algorithms, their coefficients were leveraged to understand when and where in the brain signals are. One of the most useful methods of visualizing these results was using a head map and displaying spatial data over small intervals of time such as 50ms. This is shown in figure 6.

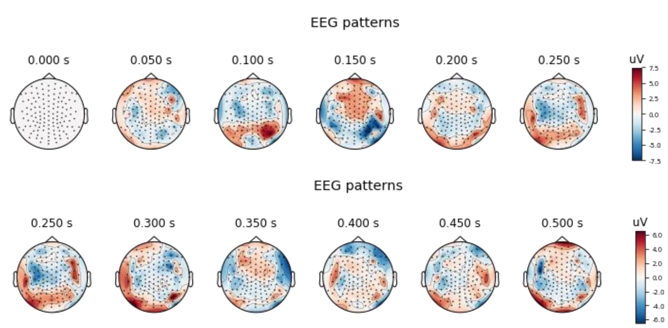

Figure 6: Coefficients of logistic regression classifier plotted on a head map. Red and blue indicate classes of audio and visual. The time periods of 100ms and 300ms show the most brain activity.

Brain activity in the brain is strongest during the 100ms to 300ms period. The parts of the brain activated are near the back of the head and sides. Both correspond to the parts of the brain responsible for processing visual and auditory responses respectively. To confirm this interpretation of figure 6, we can compare the results from the sliding estimator, to see classification scores over time.

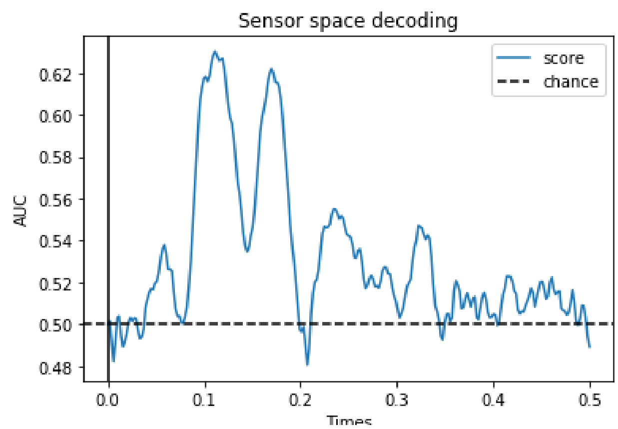

Figure 7: Sensor space decoding is MNE-Python term for classifying EEG data using a sliding estimator. Each sample (~2ms) was classified individual for all epochs and channels.

The results from the sliding estimator match those found in figure 6. Although we cannot tell where in the brain from figure 7, we can confirm when there is the most activity. Our results in the time domain matched those found in Britton's dissertation that highlights when he found when and where the most activity in the brain is starting from 150ms to 250ms.

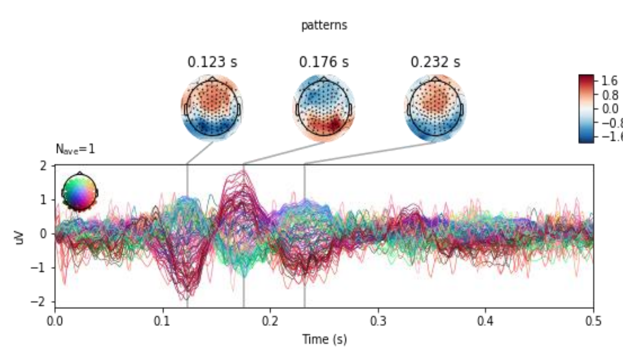

Figure 8: ERP of coefficients provided by logistic regression classifier. This shows the points in time where signals are strongest. This is also evident spatially as blue and red colors for the classes audio and visual.

The figure above shows another visual interpretation of coefficients from logistic regression classification results. There is a trail of red lines starting from 100ms, that dip and then at 123ms and then increase at 176ms. This indicates that these coefficients affect classification between the two classes.

To confirm consistency in our results, we thought to use Random Forests because we got higher classification scores with it. Coefficients for this classifier are only positive as opposed to logistic regression. They tell us how important the features (in this case the EEG channels) are for classification. Figure 9 shows the importance of particular channels during classification. Comparing it to the plot for logistic regression, we can see that there is consistency throughout both.

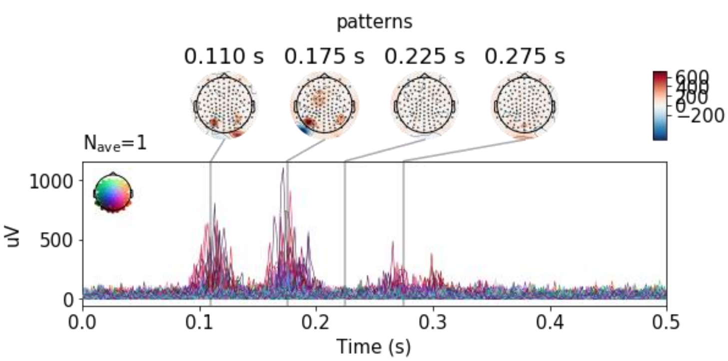

Figure 9: Plot of random forest coefficients for audio versus visual classification problem. This method is more precise in where the largest signals occur in the brain, as each map of the head has one dark spot for the most active region at that time.

Based on these results with regards to audio versus visual classification, we can tell where in the brain and when in time responses are made to stimuli. The visual cortex is the most active regions in with both classifiers. This matches what is expected during this type of stimulus processing. Unfortunately, in our observations, the auditory response was much more difficult to identify.

The process for lexical versus non-lexical classification followed the same process as the previous problem tackled. We found that finding signals in the brain for this problem was harder to confirm. Using the same techniques but changing the labels for lexicality did not prove to show the same confident results. The figures are not as exciting, and certainly need further evaluation.

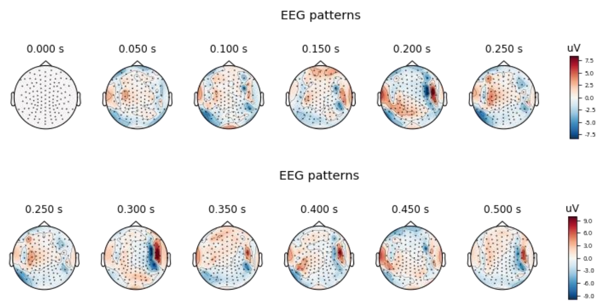

Figure 10: Coefficients of Logistic Regression for lexical versus non-lexical classification problem. According to the head maps, a portion of the brain on the right side is most responsible for classification.

The results in figure 10 show us that only a small portion of the brain was helpful in classifying lexical versus non-lexical stimuli. We know from prior research that language processing is localized in the left side of the brain [10]. This conflict with the results shown in figure 10, as the most active region is on the right side. Despite this, some lighter colored, but active regions are visible on the left side as well. The uncertainty is seen over again with other the AUC curve and ERP of coefficients. Although we can attempt to solve this lexicality problem, a difference between

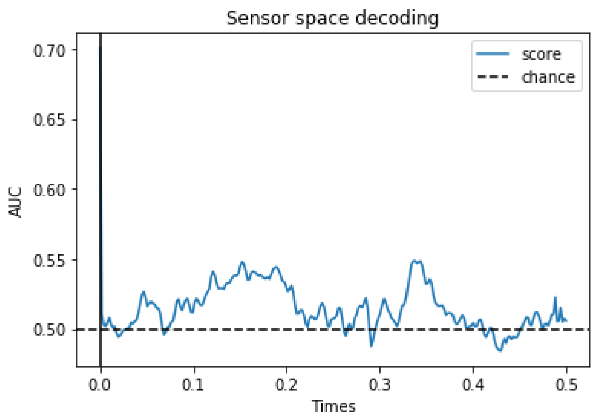

Figure 11: AUC curve of lexical versus non-lexical classification problem. Scores are near the chance line and are clearly unreliable to say when activity in the brain is most prevalent.

the two classes are not as obvious as the task of classifying modality. This score may be due to the weakness of the logistic regression classifier. We have shown that random forests have proven to be more reliable for this type of classification, and so we used that method to try to get stronger signals in the brain.

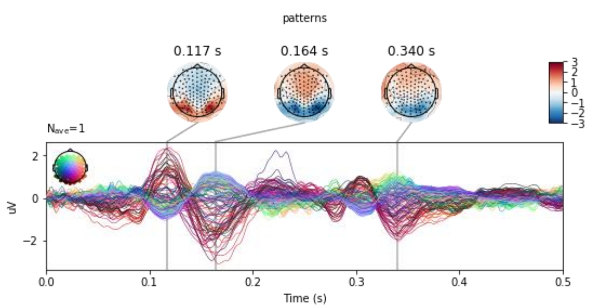

Figure 12: The three points in time that the logistic regression classifier found the most useful information for classification. Notice how the lines are more densely grouped towards the center, indicating a weaker signal.

Based on figure 13 and prior research, the feature that resembles the location of language processing is at 30ms. However, given the same pattern of spikes in other regions of the brain, we cannot conclude that this is reliable. This

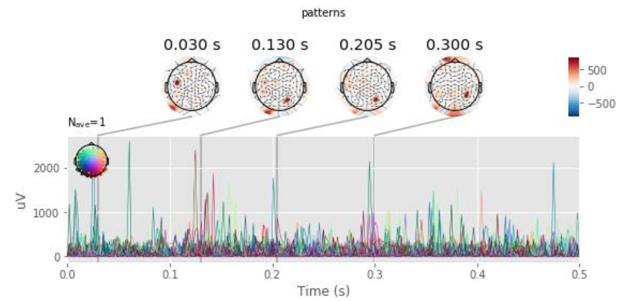

Figure 13: Plot of Random Forests coefficients for binary lexicality problem. There are a few features that standout, but we cannot conclude on a general signal based on this.

shows that the language versus non-language problem is much more difficult to find signals for.

Conclusion

The challenges faced individually with preprocessing required a profound understanding of the experiment to overcome. Providing clean and labeled data to the team was priority before we could begin doing machine learning and dive into the problems we were trying to solve. Fortunately, the data was more than acceptable to work with, and pipelines were easily created to work with it.

Despite the lackluster results for the lexicality and arguably for modality as well, we were able to begin understanding where and when in the brain decisions were made by classifiers. More specifically, we were able to tell which channels on an EEG cap contributed the most. Specific times in milliseconds were shown where signals were the strongest. Improving upon this work would certainly by trying more complex classifiers such as neural networks would give another perspective and possibly better insight into the problems we tried to solve.

With clean data, and a robust labeling system, this project can be taken further to solve more challenging problems that deal with the concepts presented during the trial, as well as if the stimuli shown was semantically congruent. Furthermore, the pipelines created here, will certainly be of use to this future work.

Citations

1. Stroop, John Ridley (1935). "Studies of interference in serial verbal reactions". Journal of Experimental Psychology. 18 (6): 643–662. doi:10.1037/h0054651. Retrieved 2008-10-08.
2. van Maanen L, van Rijn H, Borst JP (December 2009). "Stroop and picture-word interference are two sides of the same coin". Psychon Bull Rev. 16 (6): 987–99. doi:10.3758/PBR.16.6.987. PMID 19966248.
3. Britton, David. "Semantic Attention: Effects of Modality, Lexicality and Semantic Content. 2017. Graduate Center, City University of New York. New York, NY.
4. King AJ, Calvert GA. Multisensory integration: perceptual grouping by eye and ear.
5. Laurienti P, Kraft R, Maldjian J, Burdette J, Wallace M. Semantic congruence is a critical factor in multisensory behavioral performance.
6. Laurienti PJ, et al. Cross-modal sensory processing in the anterior cingulate and medial prefrontal cortices.
7. Xingwei An, Yong Cao, Jinwen Wei, Shuang Liu, IEEE Member, Xuejun Jiao, Dong Ming*, Senior IEEE Member. The Effect of Semantic Congruence for Visual-auditory Bimodal Stimuli
8. A. Gramfort, M. Luessi, E. Larson, D. Engemann, D. Strohmeier, C. Brodbeck, L. Parkkonen, M. Hämäläinen, MNE software for processing MEG and EEG data, NeuroImage, Volume 86, 1 February 2014, Pages 446-460, ISSN 1053-8119, [DOI]
9. Scikit-learn: Machine Learning in Python, Pedregosa et al., JMLR 12, pp. 2825-2830, 2011.
10. Friederici, Angela D. "The Brain Basis of Language Processing: From Structure to Function." *Physiological Reviews* 91, no. 4 (2011): 1357-392. doi:10.1152/physrev.00006.2011