Mind Reading with Artificial Intelligence

Anuhya Bhagavatula
University of Washington
Seattle, WA 98105
anuhyabs@uw.edu

Shrusti Ghela
University of Washington
Seattle, WA 98105
sghela@uw.edu

Abstract

Human communication works by turning thought into motion. However, sometimes it would be really handy if we could communicate directly from our brain. Studies on brain imaging show that different spatial patterns of neural activation are associated with thinking about different semantic categories of words and so we have attempted to "Mind Read" by applying machine learning algorithms on data from functional magnetic resonance imaging (fMRI) reports to predict the category of words that a person may be thinking of.

1 Introduction

1.1 Motivation

The brain has been a mystery for generations. The authors of the project have always been curious to understand how the human brain works, how it perceives images and other senses and how it functions so wonderfully. Now finally, we are in an age where we can understand the underlying functioning of the neural structures within the brain and make an attempt to piece together why humans respond to situations the way they do. Both of us are from a Computer Science Background and the science of the brain has awed us. This project gives us an opportunity to combine our area of study (Artificial Intelligence) to our area of interest (Neuroscience) and this is our main motivation behind taking up this project.

1.2 fMRI

Currently, scientists have two ways to figure out what is going on inside a human brain. The first is functional Magnetic Resonance Imaging (fMRI) and the other is by using electrodes. fMRI measures the flow of blood to different regions of the brain. The blood flow is correlated with neural activity, so an fMRI helps to understand what part of the brain are activated in a certain task. It records brain activities with high spatial resolution while a subject is performing a specific cognitive task. The fMRI has some shortcomings such as the machinery is expensive and it takes time to take an fMRI which means that the temporal resolution isn't great, typically a few seconds.

1.3 Machine Learning

Machine Learning is the ability of getting computers to learn and solve problems without explicitly being programmed. Machine learning algorithms do so by using historical data as input to predict outcomes for new output values. Classification and Regression models are often used to analyze fMRI data and predict In this project, we want to be able to predict the category of a word that a subject is reading based on the activation patterns in their brain.

2 Experimental Paradigm

The experiment that originated the dataset was designed to predict human brain activity associated with the meanings of nouns. Nine right handed subjects (aged between 18 and 32) participated in the experiment. Each subject was shown one word per trial and were asked to think about a set of properties associated with that word. Each word was presented for 3 seconds followed by a 7 seconds rest period. 5 words belonging to 12 categories were presented to each subject over 6 experimental epochs. An epoch is a setting where all the 60 words were presented, without repetition. Each epoch had all the 60 words but in a different order.

3 Dataset Description

3.1 Data Acquisition and Processing

The functional images were acquired from a Siemens Allegra 3.0T Scanner at the Brain Imaging Research Center of Carnegie Mellon uiversity and the University of Pittsburgh. The acquisition matrix was 64 x 64 with 3.125 x 3.125 x 5-mm voxels. The dataset processing and analysis included correction for time slicing, motion, linear trend, and temporary smoothing with a high pass filter. The data was then normalized and resampled to 3 x 3 x 6-mm voxels. The mean of the four images acquired during the 4s window, offset 4s from the stimulus onset provided the main input measure.

3.2 Dataset Description

The pre-processed and analysed data was obtained from the following site:

```
http://www.cs.cmu.edu/afs/cs/project/theo-73/www/science2008/data.html
```

The data consists of .mat files containing the dataset for 9 subjects. Each file has three variables: info, data and meta.

The 'meta' variable contains general information about the dataset such as the name of the fMRI study, identifier for the human subject, the number of trials in the dataset, the number of voxels in each image and the maximum coordinates in the brain image.

The 'info' variable describes information about each presentation trial. It is a 1 x 360 struct array and it contains information on the category of the word, the numeric index of each category, the word, numeric index of the word , and the number of times the word is presented.

The 'data' variable describes the actual image intensity data values. It is a 360 x 1 cell array, where each cell represents one trial in the experiment. Each element in this call array is a 1 x V array where V is the number of voxels.

4 Challenges and Limitations

One big challenge in the dataset was that the category of the words were not independent. For example, two of the categories were building and buildpart. Even if the words within these categories are independent there would exist some amount of intersection in the set of properties thought about each word. Another challenge was that we had 12 categories of words and a relatively small data size so essentially just 60 training points for each category. The limitation in the size of the fMRI data is due to the limited amount of data that can be gathered from a single subject in a single session. The fMRI dataset has a large number of voxels (or features in ML terms) but a small number of time points due to limitations of samples per experiment session. Multi-subject fMRI datasets would be one way to fix this limitation. However, the number of voxels of each patient are not aligned. The variabilities come from multiple resaons - difference in brain size, or different prior experience between subjects. It is necessary to spatially align responses across subjects which requires some amount of topographical factor analysis - which is currently beyond the scope of our project.

5 Classification Algorithms

5.1 Naive Bayes Algorithm

Naive Bayes is a classification technique based on the assumption of covariates independence, also known as Bayes' theorem. It is assumed that the presence of one feature in a class has no bearing on the presence of any other feature. Building a Bayesian model is simple and especially useful when dealing with large data sets. Naive Bayes is recognized to outperform complicated classification methods in addition to its simplicity. The Bayes theorem allows you to calculate posterior probability $P(c|x)$ using $P(c)$, $P(x)$, and $P(x|c)$. The following is the formula for Posterior Probability:

$$P(c|x) = \frac{P(x|c)P(c)}{P(x)}$$

$P(c|x)$ is the posterior probability of class (target) given predictor (attribute). $P(c)$ is the prior probability of class. $P(x|c)$ is the likelihood which is the probability of predictor given class. $P(x)$ is the prior probability of predictor.

5.2 K-Nearest Neighbor Algorithm

The K nearest neighbors technique is a simple classification and regression algorithm. It merely saves all available examples and uses a majority vote of its K neighbors to classify new cases. The class's assigned case is the most common among its K closest neighbors, as determined by a distance function (Euclidean, Manhattan, Minkowski, and Hamming).

For continuous variables, the three previous distance functions are employed, whereas for categorical variables, the Hamming distance function is utilized. If $K = 1$, the case is simply allocated to the nearest neighbor's class. When performing kNN modeling, selecting K can be difficult at times. We performed cross-validation to select K.

5.3 Decision Tree

Decision Tree is a supervised learning technique that can be used to solve both classification and regression problems, however it is most commonly employed to solve classification issues. It is a tree-structured classifier in which internal nodes contain dataset attributes, branches represent decision rules, and each leaf node represents the result.A Decision tree has two nodes: the Decision Node and the Leaf Node. Leaf nodes are the output of those decisions and do not contain any more branches, whereas Decision nodes are used to make any decision and have several branches. The decisions or tests are made based on the characteristics of the given dataset. It's a graphical depiction for obtaining all feasible solutions to a problem/decision depending on certain parameters. It's termed a decision tree because, like a tree, it starts with the root node and grows into a tree-like structure with additional branches. We utilize the CART algorithm, which stands for Classification and Regression Tree algorithm, to form a tree. A decision tree simply asks a question and divides the tree into subtrees based on the answer (Yes/No).

The procedure for determining the class of a given dataset in a decision tree starts at the root node of the tree. This algorithm compares the values of the root attribute with the values of the record (actual dataset) attribute and, based on the comparison, follows the branch and jumps to the next node. The algorithm compares the attribute value with the other sub-nodes and moves on to the next node. It repeats the process until it reaches the tree's leaf node.

5.4 Random Forests

As the name implies, a random forest is made up of a huge number of individual decision trees that work together as an ensemble. Each tree in the random forest produces a class prediction, and the class with the most votes becomes the prediction of our model. The wisdom of crowds is the basic principle of random forest, and it's a simple yet effective strategy. The reason why the random forest model works so well is A large number of reasonably uncorrelated models (trees) acting as a committee will outperform any of the constituent models individually. While some trees may be incorrect, many others will be correct, allowing the trees to move in the correct direction as a group.

Table 1: Accuracy of 12 categories classification

Algorithm	Accuracy
Naive Bayes	13.3%
kNN	21.67%
Decision Trees	11.67%
Random Forest	28.33%

Table 2: Accuracy of 2 categories classification

Algorithm	Accuracy
Naive Bayes	68.33%
kNN	68.33%
Decision Trees	53.33%
Random Forest	83.33%

6 Results and Findings

The dataset was split into training and tetsing by considering the the first 5 epochs of 300 observations as training and the remaining epoch with 60 observations in the testing data. This was preferred over random sampling because the data size was small and we wanted to give each category of words equal instances of training. By random sampling of the 360, there was a possibility for higher bias for a single category. Since each epoch already had category of words in a random order, the first 5 epochs were used for training. The dataset for Patient 1 had 21674 voxels (or number of features). Dimensionality Reduction was performed on the features using Principal Component Analysis to transform the large set of features into a smaller one. Classification Algorithms were then applied on this transformed dataset. Originally all the 12 categories of the words: manmade, clothing, furniture, tool, vehicle, vegetable, building, insect, bodypart, buildpart, kitchen, animal were used to predict. The results obtained are shown in Table 1. The low performance of the algorithms can be mainly attributed to the smaller training size available per category and the non-independence in properteries of the categories. To overcome the small data size and non-independent catgories challenges that have also been described earlier, we tried to combine the 12 categories into two broader categories - manmade (vehicle, tool, building, buildpart, kitchen, furniture, manmade, clothing) and natural (animal, insect, bodypart, vegetable). This approach made the category more separable and independent and improved the overall accuracy of all the four classification algorithms. The following results are shown in Table 2.

We have considered and reported only one patient results due to space constraints. The complete analysis on the patients and the code are presented at the following link.

https://github.com/anuhyabs/Mind-Reading-with-AI

7 Conclusion

We attempted to "mind read" by predicting the category of word that a subject was thinking based on the voxels from obtained fMRI reports. We overcame some challenges as mentioned in the report. Among all the four classification algorithms, random forests performed the best in both 2-category classification and 12-category classification. As a part of the future works, we plan to apply dense neural networks and convolutional neural networks to the fMRI dataset in the future. We also plan to work on multi-subject fMRI datasets by handling the variabilities using topographical factor analysis.

References

[1] Mitchell, T. M., Shinkareva, S. V., Carlson, A., Chang, K. M., Malave, V. L., Mason, R. A., & Just, M. A. (2008). Predicting human brain activity associated with the meanings of nouns. science, 320(5880), 1191-1195.

[2] Pereira, F., Mitchell, T., & Botvinick, M. (2009). Machine learning classifiers and fMRI: a tutorial overview. NeuroImage, 45(1 Suppl), S199–S209. https://doi.org/10.1016/j.neuroimage.2008.11.007.

[3] Abraham, A., Pedregosa, F., Eickenberg, M., Gervais, P., Mueller, A., Kossaifi, J., ... & Varoquaux, G. (2014). Machine learning for neuroimaging with scikit-learn. Frontiers in neuroinformatics, 8, 14.

[4] Yiu, T. (2019, June 19) Understanding Random Forest. Towards Data Science. https://towardsdatascience.com/understanding-random-forest-58381e0602d2

[5] Upasana. (2020, November 25) Introduction to Classification Algorithms. edureka! https://www.edureka.co/blog/classification-algorithms/

[6] Gong, D. (2022, February 22) Top 6 Machine Learning Algorithms for Classification. https://towardsdatascience.com/top-machine-learning-algorithms-for-classification-2197870ff501

[7] Hossenfelder, S. [Sabine Hossenfelder]. (2021, June 26) Mindreading With Artificial Intelligence [Video]. Youtube. https://www.youtube.com/watch?v=rA5k2S8xPK8

[8] Zhang, H. (2020). Machine learning for multi-subject fmri analysis (Doctoral dissertation, Princeton University).

[9] Bonakdarpour, M. Using Machine Learning to Predict Human Brain Activity.

[10] http://www.cs.cmu.edu/afs/cs/project/theo-73/www/index.html