fMRI Data Analysis

Instructions

1. You should implement the assignment in Python.
2. Ensure that the submitted assignment is your original work. Please do not copy any part from any source, including your friends, seniors, and/or the internet. If any such attempt is caught, serious actions, including an F grade in the course, are possible.
3. A single .zip file needs to be uploaded to the Courses Portal.
4. Your grade will depend on the correctness of answers and output. Due consideration will also be given to the clarity and details of your answers and the legibility and structure of your code.
5. Please start early to meet the deadline. Late submissions won't be evaluated.
6. The grading rubric for the brain encoding and decoding part of the assignment can be found at the end of the document.

Theory Questions

Please answer the following questions after reading the paper *Auditory Language Comprehension: An Event-Related fMRI Study on the Processing of Syntactic and Lexical Information*: `https://www.sciencedirect.com/science/article/abs/pii/S0093934X00923137`

1. How does the use of four stimulus conditions (normal speech, syntactic speech, real-word lists, pseudoword lists) allow the authors to dissociate syntactic and semantic processing in auditory language comprehension? (5 marks)
2. What preprocessing and statistical analysis steps were applied to the fMRI data before group-level inference, and what role does Fisher's Z transformation play in this pipeline? (5marks)

Problem Statement

The brain encoding problem aims to automatically generate fMRI brain representations given a stimulus. The brain decoding problem is the inverse problem of reconstructing the stimuli given the fMRI brain representation. In this assignment, your task would be to construct an encoder as well as a decoder for textual stimuli. Details about the dataset, methodology, and the tasks are provided in the sections below.

1 Dataset

The dataset to be used for the assignment can be downloaded from here: Download the dataset.

The dataset consists of 627 sentences and the corresponding fMRI data, recorded when the sentences were presented to a subject one by one. fMRI is provided for four different brain ROIs listed below:

- **Language:** Related to language processing, understanding, word meaning, and sentence comprehension
- **Vision:** Related to the processing of visual objects, object recognition
- **Task Positive:** Related to attention, salience information
- **Default Mode Network (DMN):** Linked to the functionality of semantic processing

The dataset consists of two files:

- `stimuli.txt`: List of 627 sentences presented as stimuli.
- `text_fmri_data.pkl`: The fMRI responses corresponding to 8 different ROIs:
  - `['language_left', 'language_right', 'vision_left', 'vision_right', 'task_positive_left', 'task_positive_right', 'dmn_left', 'dmn_right']`

2 Tasks

2.1 Brain Decoder

The first task is to build 8 different decoders, one for each ROI corresponding to each hemisphere. The decoder could be simply a regression-based model like ridge/lasso regression or any model of your choice. You should perform k-fold cross-validation and report the average of the evaluation metrics across folds.

2.1.1 Evaluation Metrics

Given a subject and a brain region, let N be the number of samples. Let $\{Y_i\}_1^N$ and $\{\hat{Y}_i\}_1^N$ denote the actual and predicted sentence vectors for the i-th sample. Thus, $Y \in \mathbb{R}^{N \times D}$ and $\hat{Y} \in \mathbb{R}^{N \times D}$, where D is the dimension of each sentence vector. **2V2 Accuracy** is computed as follows:

$$\text{2V2Acc} = \frac{1}{N(N-1)/2} \sum_{i=1}^{N} \sum_{j=i+1}^{N} I\left[\left(\cos_D(Y_i, \hat{Y}_i) + \cos_D(Y_j, \hat{Y}_j)\right) < \left(\cos_D(Y_i, \hat{Y}_j) + \cos_D(Y_j, \hat{Y}_i)\right)\right] \tag{1}$$

where $\cos_D$ is the cosine distance function, and $I[\cdot]$ is an indicator function such that $I[c] = 1$ if c is true, else it is 0. The higher the 2V2 accuracy, the better. **Pearson Correlation (PC)** is computed as:

$$PC = \frac{1}{N} \sum_{i=1}^{N} \text{corr}(Y_i, \hat{Y}_i) \tag{2}$$

where corr is the Pearson correlation function.

2.2 Brain Encoder

The second task is to build 8 different encoders, one for each brain area. Similar to a decoder, the encoder could also be simply a regression-based model like ridge/lasso regression. You should perform k-fold cross-validation and report the average of the evaluation metrics across folds.

2.2.1 Evaluation Metric

Given a subject and a brain region, let N be the number of samples. Let $\{Y_i\}_1^N$ and $\{\hat{Y}_i\}_1^N$ denote the actual and predicted voxel value vectors for the i-th sample. Thus, $Y \in \mathbb{R}^{N \times V}$ and $\hat{Y} \in \mathbb{R}^{N \times V}$, where V is the number of voxels in that region. **2V2 Accuracy** is computed as follows:

$$\text{2V2Acc} = \frac{1}{N(N-1)/2} \sum_{i=1}^{N} \sum_{j=i+1}^{N} I\left[\left(\cos_D(Y_i, \hat{Y}_i) + \cos_D(Y_j, \hat{Y}_j)\right) < \left(\cos_D(Y_i, \hat{Y}_j) + \cos_D(Y_j, \hat{Y}_i)\right)\right] \tag{3}$$

where $\cos_D$ is the cosine distance function, and $I[\cdot]$ is an indicator function such that $I[c] = 1$ if c is true, else it is 0. The higher the 2V2 accuracy, the better. **Pearson Correlation (PC)** is computed as:

$$PC = \frac{1}{N} \sum_{i=1}^{N} \text{corr}(Y_i, \hat{Y}_i) \tag{4}$$

3 Sentence Representations

You have to extract the sentence embeddings from:

1. Transformer-based language models like BERT/RoBERTa. There can be two possible sentence representations here:
   - CLS token embedding
   - The pooled embeddings of all the words (word embeddings of the sentence)
2. GloVe embeddings: Extract GloVe embedding for each word in a sentence and compute the average of all these word embeddings to get the sentence embedding.

For every sentence, you should finally have three D-sized vectors:

1. BERT/RoBERTa CLS representation
2. BERT/RoBERTa Pooled representation
3. GloVe representation

Apart from all the analysis to be done in the previous sections, you need to analyze and answer the following questions in the report:

1. Order of Accuracy/Pearson Correlation for the above three representations for Encoder and Decoder, for each ROI.
2. Reason for the above given order.

4 Grading Rubric

- Brain Encoding using BERT CLS representations (15 marks)
- Brain Encoding using BERT pooled representations (15 marks)
- Brain Encoding using Glove representations (15 marks)
- Brain Decoding using BERT CLS representations (15 marks)
- Brain Decoding using BERT pooled representations (15 marks)
- Brain Decoding using Glove representations (15 marks)
- Analysis and Plots (10 marks)

5 Deliverables

You should report the results on the bar plot, comparing the evaluation metrics for different brain ROIs. One bar plot for 2V2 Accuracy and one for Pearson correlation comparing the metrics for the brain areas. This needs to be done for both tasks. You are expected to provide some insights about the results.

You need to submit a tar file named `<roll_no.>.tar` consisting of two files:

- `code.ipynb` - consisting of the code
- `report.pdf` - consisting of the bar plots and insights from the results