Cognitive Science and AI

Assignment 3: Textual Brain Encoder and Decoder

Kshitijaa Jaglan
2019115005

The assignment uses the following few key values and models:
Library for Sentence Transformers: `sentence_transformers` (or SBERT)
Model used: `all-distilroberta-v1`
Dataset: Two subjects and their data as provided
Cosine distance: Using `scipy`
2v2 metric: Implemented from scratch
Pearson Coefficient: Uses `numpy.corrcoef`

To implement the decoder and encoder, two classes - `Decoder` and `Encoder` have been created, each with methods to initialize the required values, `train` and generate a `report`. For training, the splits are made using **`KFold` with cross validation** using the following structure:

```
KFold(n_splits=n, shuffle=False, random_state=None)
```

In `KFolds`, the entire assignment was run for two values of `n` -- `3` and `5` to test what the difference would be. The findings for the same have been shown later.

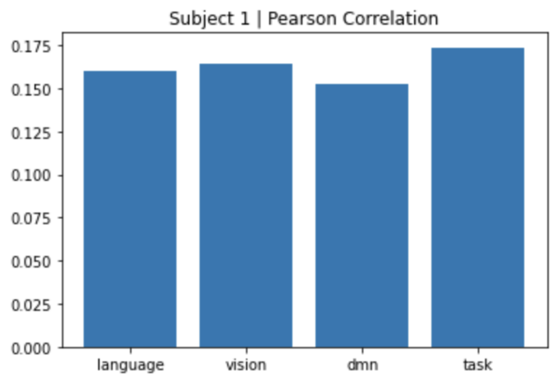
This is followed by using **`Ridge` regression** (from `sklearn`) for training on the sentence embeddings for each set of splits received from `KFold`.
For each set of test and predicted values, the corresponding **2v2 accuracy metric** and **Pearson Correlation Coefficients** are calculated.

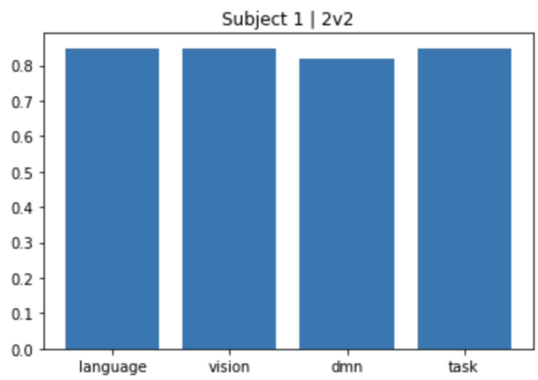
TRENDS

> Subject 1

>> n = 3

>>> Decoding

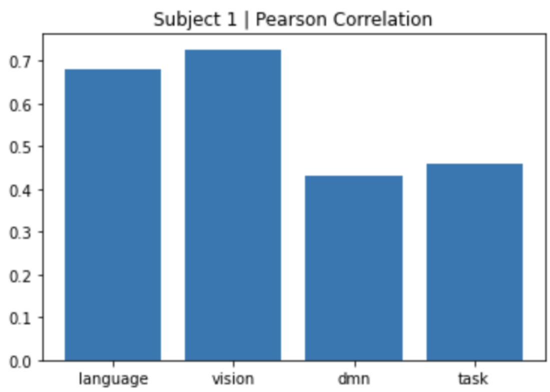


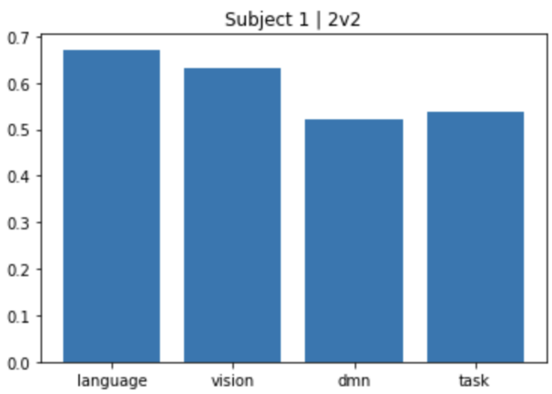



```
# report:
{'language': {'pc': 0.1602492030074432, '2v2': 0.8487302171512697},
 'vision': {'pc': 0.16448072405388228, '2v2': 0.8485615261931051},
 'dmn': {'pc': 0.15262956751939646, '2v2': 0.8191479573058521},
 'task': {'pc': 0.17361931761559793, '2v2': 0.8469512943197154}}
```

All performed almost the same in terms of 2v2, but task was slightly better if we look at Pearson Correlation (PC).

>>> Encoding



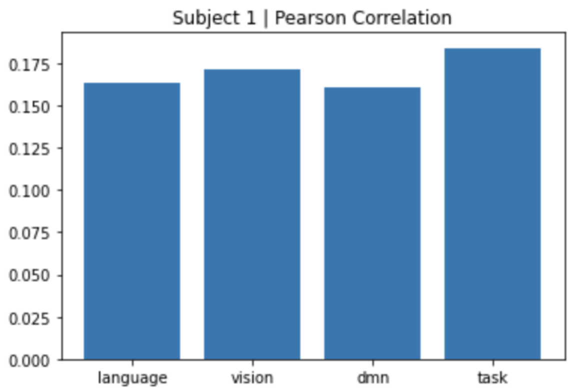


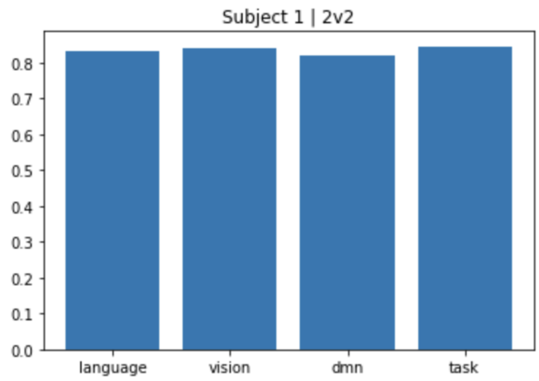

```
# report:
{'language': {'pc': 0.6790922566013622, '2v2': 0.6716813887866518},
 'vision': {'pc': 0.7272369629801676, '2v2': 0.6317323027849344},
 'dmn': {'pc': 0.4306317305671888, '2v2': 0.5201509017298491},
 'task': {'pc': 0.45912743818025814, '2v2': 0.5363452337136547}}
```

Overall, vision and language perform better, both with PC and 2v2. However, this difference is more visible with PC than 2v2. Moreover, Vision performs better (however slightly) than Language if we look at PC and opposite in case of 2v2.

>> n = 5

>>> Decoding

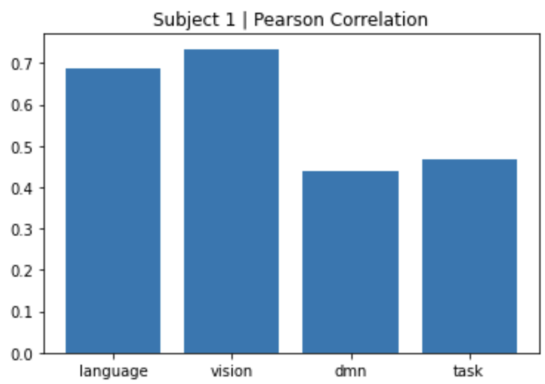


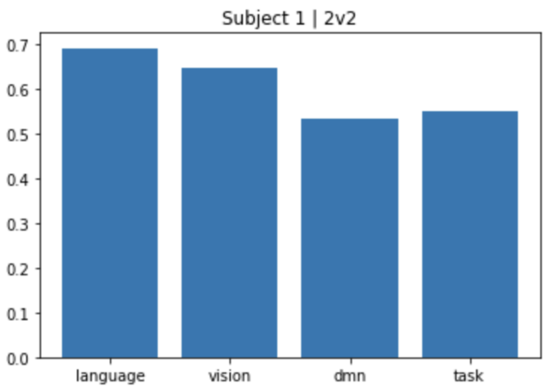



```
# report:
{'language': {'pc': 0.16338641040373386, '2v2': 0.8335975422427035},
 'vision': {'pc': 0.17165816398182077, '2v2': 0.8424942140296979},
 'dmn': {'pc': 0.16075523896944072, '2v2': 0.8221374295954942},
 'task': {'pc': 0.18409911508117865, '2v2': 0.8461927291346646}}
```

All performed almost the same in terms of 2v2, but task was slightly better if we look at Pearson Correlation (PC).

>> Encoding





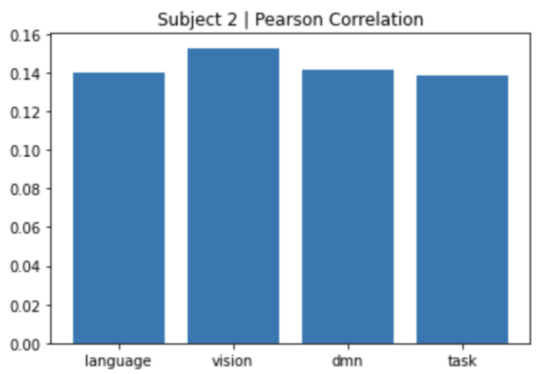

```
# report:
{'language': {'pc': 0.6859656528424612, '2v2': 0.6916174091141833},
 'vision': {'pc': 0.7341015903770238, '2v2': 0.6488749615975422},
 'dmn': {'pc': 0.43984202399196226, '2v2': 0.5348452636968766},
 'task': {'pc': 0.4671064544125694, '2v2': 0.5514982078853047}}
```

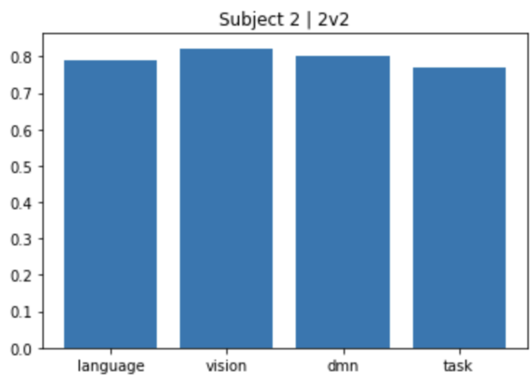
Overall, vision and language perform better, both with PC and 2v2. However, this difference is more visible with PC than 2v2. Moreover, Vision performs better (however slightly) than Language if we look at PC and opposite in case of 2v2.

> Subject 2

>> n = 3

>>> Decoding

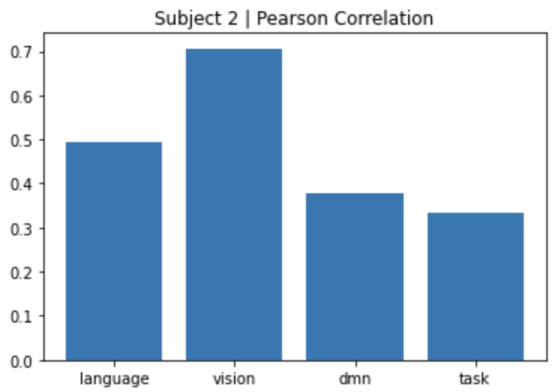


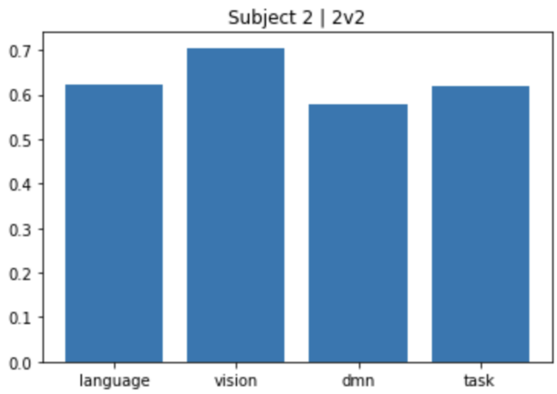



```
# report:
{'language': {'pc': 0.14012307508204397, '2v2': 0.7914519690835481},
 'vision': {'pc': 0.15272578079192725, '2v2': 0.8225524475524475},
 'dmn': {'pc': 0.1412803279137329, '2v2': 0.8006839651576495},
 'task': {'pc': 0.13883735960011612, '2v2': 0.7685253343148081}}
```

Vision performs the best, followed by language and then dmn at a very close third, followed by task at near fourth. This trend is consistent in both PC and 2v2 metrics.

>>> Encoding



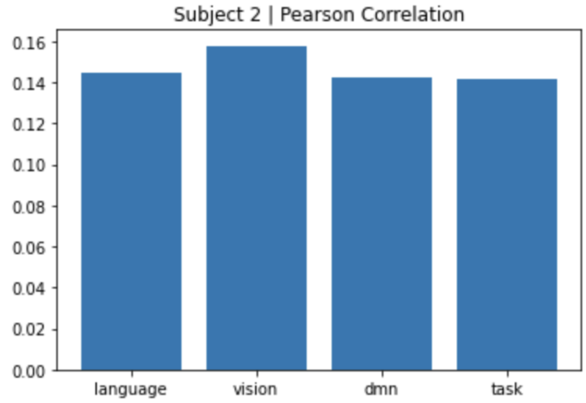


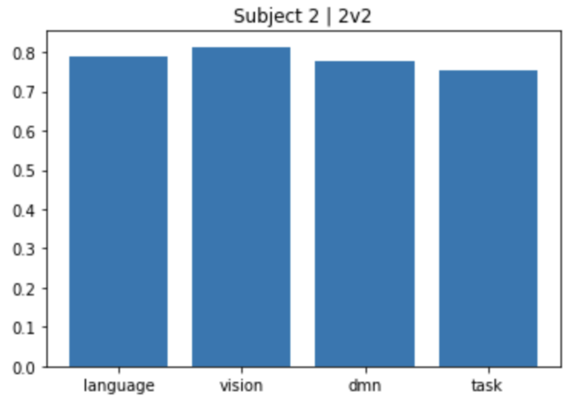

```
# report:
{'language': {'pc': 0.49272033777750114, '2v2': 0.6236504723346828},
 'vision': {'pc': 0.7061633131475054, '2v2': 0.7051435406698564},
 'dmn': {'pc': 0.37748706168141194, '2v2': 0.5796681388786652},
 'task': {'pc': 0.33201641507590024, '2v2': 0.6175009201324991}}
```

Looking at PC alone, vision performs better by a **great** margin! However, this margin is comparatively very less through 2v2. Moreover, in 2v2 task follows with a close second, while it performs the worst if we look at PC.

>> n = 5

>>> Decoding

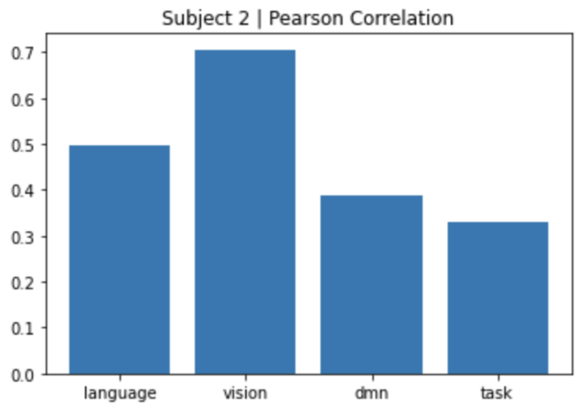


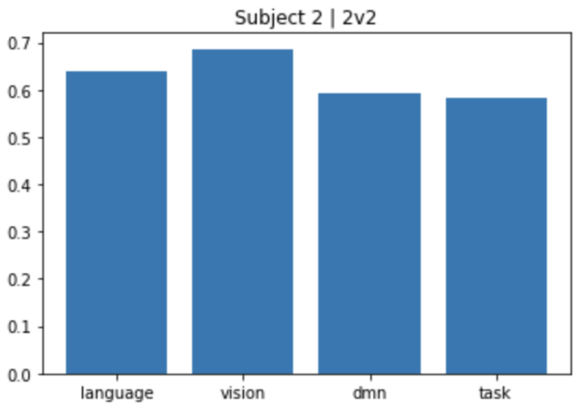



```
# report:
{'language': {'pc': 0.1445872041368747, '2v2': 0.7875244239631336},
 'vision': {'pc': 0.1577808998979287, '2v2': 0.8137335381464414},
 'dmn': {'pc': 0.1420760141148344, '2v2': 0.7758484383000511},
 'task': {'pc': 0.14160560646073123, '2v2': 0.7541374295954941}}
```

Vision performs the best, followed by language and then dmn at a very close third, followed by task at near fourth. This trend is consistent in both PC and 2v2 metrics.

>>> Encoding






```
# report:
{'language': {'pc': 0.4978534333135844, '2v2': 0.6390824372759857},
 'vision': {'pc': 0.7056094804167307, '2v2': 0.6860972862263185},
 'dmn': {'pc': 0.38734507728027057, '2v2': 0.5929011776753712},
 'task': {'pc': 0.329073675049426, '2v2': 0.5831254480286738}}
```

Looking at PC alone, vision performs better by a **great** margin! However, this margin is comparatively very less through 2v2. Moreover, in 2v2 task follows with a close second, while it performs the worst if we look at PC.

Overall Observations and inferences:

Between n = 3 and n = 5:

There was no such difference observed in the trends. Even with the values, the maximum difference was around 0.02 which wasn't significant looking at the overall values

Between PC and 2v2:

PC seems to highlight greater margins in differences as compared to 2v2 overall in all the plots. An interesting observation was that the scale of PC was significantly very less in all the decoder plots as compared to the corresponding 2v2 plots.

Decoder vs Encoder:

Overall, Encoder works better by a **wide** margin if we look at only PC, whereas it comes very close to performing at par with Decoder if we look at only 2v2. Moreover, it should also be noticed that the values of PC jump maybe even more than 4x when we shift from decoder to encoder!

General observations and inferences:

- Since the ROI of language would also involve vision to some extent (as people tend to picturise what they read), it has been observed that the values for both of these are usually very close to each other.
- The ROI of Default Mode Network (DMN) is linked to functionality of semantic processing but has been observed to be not as involved as language and vision
- The ROI of task - related to attention and salience information has also been observed to affect at a level similar to that of DMN