Sentimental Analysis Using BERT to Understand Impact of Covid on Peoples Mental Health

A dissertation submitted in partial fulfilment of
the requirements for the degree of
BACHELOR OF SCIENCE in Computer Science
in
The Queen's University of Belfast
by
'Tarpan Rai'

Contents

1. Abstract

In this paper we will look at how the pandemic has affected people's mental health sentiment and dietary sentiment using natural language processing. Using a pre-trained BERT model, we can find the sentiment of each free text response and attempt to find correlations with different demographics. With the help of topic modelling, we can find frequency of words to create common topic to see any similarities in responses.

2. Introduction

During the COVID-19 pandemic, issues around mental health and well-being gained widespread prominence. Although the lockdown was a short-term strategy to combat the disease, it inevitably gave rise to long-term health issues which may or may not have been aware of before. To understand this, we used a survey done by the Centre for public health at QUB. The response contains demographics of respondents and free text responses regarding their mental health and dietary health. With this dataset we can apply natural language processing methods such as sentimental analysis using BERT and topic modelling to better understand and analyze sentiment of respondents. We can go further and try to find correlation between demography and sentiment to find if different demographics were affected more or less by the pandemic.

3. Background Information and Research.

BERT (Bidirectional Encoder Representation from Transformers) is a pre-trained model and can be fine-tuned to a wide variety of task, in our case, for sentimental analysis. BERT was created by google and published their BERT blog along with a research paper by (Devlin J. et al., 2018) containing their explanation behind the BERT model. To understand the BERT model, one must learn the foundation shown in figure 4 specifically the attention layer as it was bases heavily on it. The research article 'Attention is all you need' by Vaswani A. et al., 2017 explains that "the Transformer, based solely on attention mechanisms" was "superior in quality while being more parallelizable and requiring significantly less time to train" compared to other model.

The BERT model is a 2-step framework: Pre-Training and Fine Tuning. It is initialized with a pretrain parameter and then fine-tuned with labelled data. The BERT report has 2 model sizes as shown in table 1.

BERT MODEL	Layers (L)	Hidden Size (H)	No. of Self Attention Layer (A)	Total Perimeter
$BERT_{Base}$	12	768	12	110M
$BERT_{Large}$	24	1024	16	340M

Table 1 BERT MODEL

The BERT model is pre-trained hence the vocabulary is fixed, it uses a word piece embedding with 30,000 token vocabularies and each token has a feature vector of 768 features. For words not in the vocabulary, BERT breaks the unknow word into multiple subworlds. With each sub-word starting with a '##' to indicate a sub-word except for the first sub-word. Each sequence of

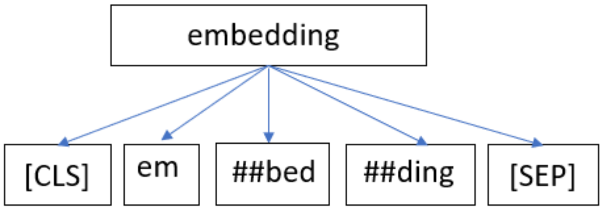


Figure 1 Embedding.

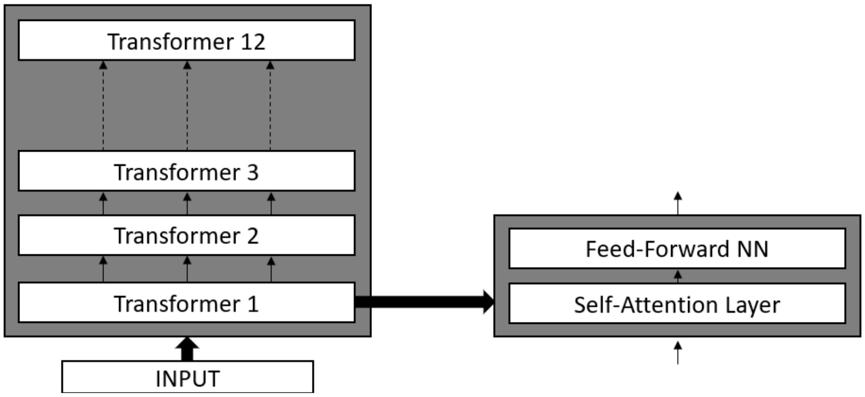


Figure 2 Transformer Layers

token starts with a special token ‘[CLS]’ and sentences are separated with ‘[SEP]’ token. For example, the word ‘embedding is not in the vocabulary:

The BERT Architecture has 12 encoders and inside each encoder there is a self-attention layer and a feed-forward neural network layer. The self-attention layer allows it to look at other position in the input sequence for clues that can help lead to a better encoding for a word. It is calculated using 3 vectors from each of the encoder’s input: Query vector (Q), Key vector (K), and Value vector (V).

1) Create vectors for each word.

$$x \cdot W^Q = Q \qquad x \cdot W^K = K \qquad x \cdot W^V = V$$

2) Calculate score.

$$Q \times K = S$$

3) Divide the score by square root of the dimension of the key vector.

$$\sqrt{d_k}$$

4) Pass the results through a SoftMax operation.

$$Softmax = \frac{e^{x_i}}{\sum_{j=1}^{n} e^{x_i}}$$

5) Multiply each vector value by SoftMax score.
6) Sum up the weighted vector value.

A multi-headed attention approach is added in the attention layer that expands the model’s ability to focus on different positions. (8 Heads) Hence it ends up with 8 different matrices then the matrices are condensed into a single matrix.

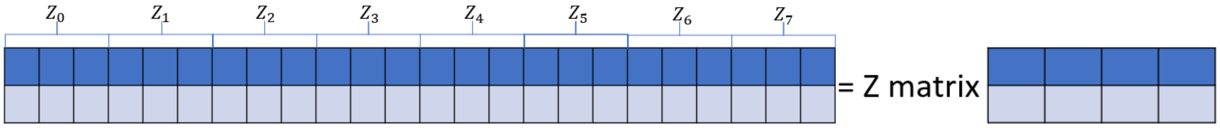


Figure 3 Attention layer matrix

3.1 Topic Modelling

Topic modelling is an unsupervised learning method in NLP that analyses the corpus and extracts the main topic presented in the data by detecting patterns like word clusters and frequency of words present. By detecting word frequency and word distance the information can be deduced into various set of topics. There are many different topic modelling methods like Latent Semantic Analysis (LSA), Latent Dirichlet Allocation (LDA) and others. In our model we will be using LDA for topic modelling and visualising it using pyLDAvis library in python. LDA is a generative probabilistic model of a corpus. It takes a corpus and represents document as random mixture over latent topics and each topic is characterized by a distribution over words as shown in figure 4.

3.1.1 How LDA works.

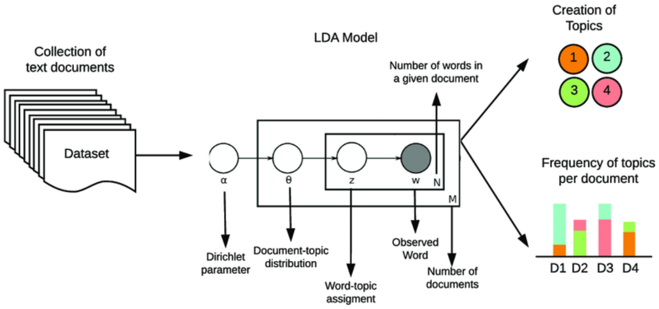


Figure 4 LDA Model

Firstly, LDA applies 2 assumptions to the corpus:

1. Documents are a mixture of topics.
2. Topics are a mixture of tokens.

Secondly, text data is then preprocessed and tokenized into 2 matrixes:

1. Document Topic Matrix which contains possible topics
2. Word Topic Matrix which contains words that those topics can contain.

Using the matrixes, it goes through iteration of assignments of topic to words. Correcting and adjusting variables across iterations until the most optimal representation of the matrixes are developed. In python the results can be visualized using the pyLDAvis

4. Dataset

Training data has been provided by Royal Victoria Hospital. The data set includes unstructured, and free-text responses by participants where they describe aspect of brain health and diet, as well as other aspects of wellbeing impacted by the COVID-19 pandemic. It has 19 question total divided into 4 different main categories: Demographic variable, General covid questions, covid and brain health, Covid and diet. The dataset has 6818 total difference respondents but only 1124 and 1217 responded to the 2 free text question respectively.

Example of question:

- Demographic Variables Question Examples:
  - Age: (16-39, 40-49, 50-65, 66-74, 75+)
  - Country (Northern Ireland, Scotland, England)
  - Gender (Male, Female, Non-binary, I'd prefer not to answer this question)
- General COVID Question Examples:
  - Do you feel impacted by COVID?
  - Have you had to check for COVID?
  - Have you received treatment for COVID-19?
- COVID and Brain Health Question Examples:
  - Has the ongoing pandemic impacted your brain health?
  - Do you feel your brain health has improved or deteriorated?
  - We are interested to know why your brain health might have deteriorated during COVID-19. Please provide a reason for this using the free text box below.
- Covid and Diet Question Examples:
  - Has COVID-19 impacted your dietary habitats?
  - Do you feel that your diet has improved or deteriorated?
  - We are interested to know why your dietary habits might have deteriorated during COVID-19. Please provide a reason for this using the free text box below.

Example of free-text responses:

- Question: We are interested to know why your brain health might have deteriorated during COVID-19.
  - "Being at home, less able to converse with friends, anxiety re being out shopping etc."
  - "Lack of exercise and more sitting"
  - "I'm now very isolated and getting a bit forgetful, nothing serious"
  - "I've been unable to see family and friends, and this is really difficult and makes me really sad"
- Question: We are interested to know why your dietary habits might have deteriorated during the COVID-19 pandemic.
  - "Eating a lot more rubbish, i.e., chocolate and biscuits!"
  - "Out of routine with diet and exercise"
  - "I'm at home more, and so treat myself to less healthy foods"
  - "I'm drinking more alcohol because I am not free to do the things I want to do.

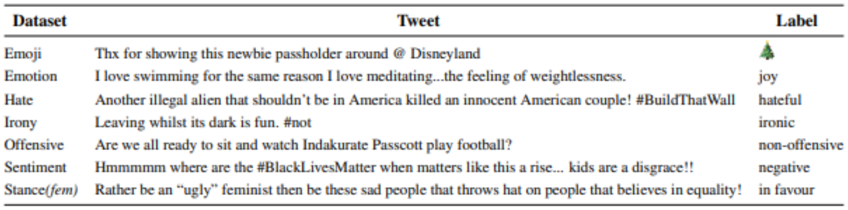
The main data we will be using are the demographic variables and free-text responses.

5. Tools Used

- Transformer model: We will be using a pre trained BERT model from the hugging face library (Cardiffnlp/twitter-roberta-base-sentiment)
- Gensim: Is an open-source python library for representing documents as semantic vectors efficiently. We will be using it for its word2vec and LDA model.
- Nltk: Contains a database of corpora used for NLP tasks.
- Matplotlib/pyLDAvis: Plotting libraries to visualize results.

6. Model Used

As stated previously, we will be using a pre-trained BERT model from the hugging face library. The model was trained using a large corpus containing 60 million English tweets [Barbieri, Neves, Camacho-Collados and Espinosa-Anke, 2021] to handle 6 different tasks: emotional recognition, emoji prediction, irony detection, hate speech detection, offensive language identification, sentimental analysis, and stance detection as shown in figure 5. Each task was trained using a range of a few hundred to 45,000 tweets.

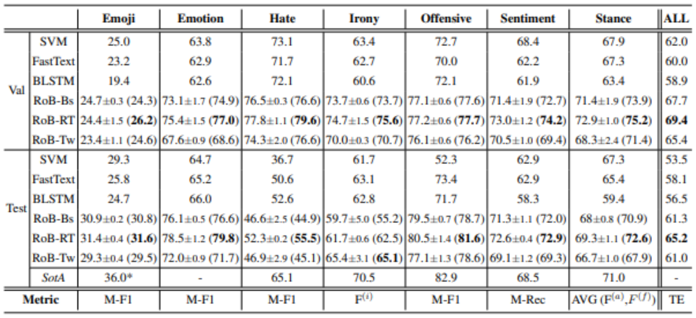


Dataset	Tweet	Label
Emoji	Thx for showing this newbie passholder around @ Disneyland	🎄
Emotion	I love swimming for the same reason I love meditating...the feeling of weightlessness.	joy
Hate	Another illegal alien that shouldn't be in America killed an innocent American couple! #BuildThatWall	hateful
Irony	Leaving whilst its dark is fun. #not	ironic
Offensive	Are we all ready to sit and watch Indakurate Passcott play football?	non-offensive
Sentiment	Hmmmmm where are the #BlackLivesMatter when matters like this a rise... kids are a disgrace!!	negative
Stance*(fem)*	Rather be an "ugly" feminist then be these sad people that throws hat on people that believes in equality!	in favour

Figure 5 Roberta models

With this model, we will be focusing on sentimental analysis specifically. Sentimental analysis which analysis if the statement is 'negative', 'neutral' or 'positive'.

The pre-trained model was trained using three variants of the RoBERTa language model due to its outstanding performance in the GLUE benchmark, which is used to eveluate performace of models in NLP tasks. As shown in figure 6 the RoBERTa model performed well on all the tasks with a score of 77 in emotion and 74.2 on sentiment, averaging 69.4 on all tasks.



		Emoji	Emotion	Hate	Irony	Offensive	Sentiment	Stance	ALL
Val	SVM	25.0	63.8	73.1	63.4	72.7	68.4	67.9	62.0
	FastText	23.2	62.9	71.7	62.7	70.0	62.2	67.3	60.0
	BLSTM	19.4	62.6	72.1	60.6	72.1	61.9	63.4	58.9
	RoB-Bs	24.7±0.3 (24.3)	73.1±1.7 (74.9)	76.5±0.3 (76.6)	73.7±0.6 (73.7)	77.1±0.6 (77.6)	71.4±1.9 (72.7)	71.4±1.9 (73.9)	67.7
	RoB-RT	24.4±1.5 (**26.2**)	75.4±1.5 (**77.0**)	77.8±1.1 (**79.6**)	74.7±1.5 (**75.6**)	77.2±0.6 (**77.7**)	73.0±1.2 (**74.2**)	72.9±1.0 (**75.2**)	**69.4**
	RoB-Tw	23.4±1.1 (24.6)	67.6±0.9 (68.6)	74.3±2.0 (76.6)	70.0±0.3 (70.7)	76.1±0.6 (76.2)	70.5±1.0 (69.4)	68.3±2.4 (71.4)	65.4
Test	SVM	29.3	64.7	36.7	61.7	52.3	62.9	67.3	53.5
	FastText	25.8	65.2	50.6	63.1	73.4	62.9	65.4	58.1
	BLSTM	24.7	66.0	52.6	62.8	71.7	58.3	59.4	56.5
	RoB-Bs	30.9±0.2 (30.8)	76.1±0.5 (76.6)	46.6±2.5 (44.9)	59.7±5.0 (55.2)	79.5±0.7 (78.7)	71.3±1.1 (72.0)	68±0.8 (70.9)	61.3
	RoB-RT	31.4±0.4 (**31.6**)	78.5±1.2 (**79.8**)	52.3±0.2 (**55.5**)	61.7±0.6 (62.5)	80.5±1.4 (**81.6**)	72.6±0.4 (**72.9**)	69.3±1.1 (**72.6**)	**65.2**
	RoB-Tw	29.3±0.4 (29.5)	72.0±0.9 (71.7)	46.9±2.9 (45.1)	65.4±3.1 (**65.1**)	77.1±1.3 (78.6)	69.1±1.2 (69.3)	66.7±1.0 (67.9)	61.0
	SotA	36.0*	-	65.1	70.5	82.9	68.5	71.0	-
Metric		M-F1	M-F1	M-F1	$F^{(i)}$	M-F1	M-Rec	AVG ($F^{(a)}$,$F^{(f)}$)	TE

Figure 6 GLUE Results of Models

7. Pre-Processing

Before using the data, some changes were made. First the file format was changed from excel format (.xlsx) to comma separated value format (.csv) due to its efficiency and flexibility with the Pandas library. Then the dataset was separated into two different files, one for brain health responses (‘Brain sentiment.csv’) and the other for dietary health responses (‘Dietary sentiment.csv’) to remove the empty cell due to it causing errors in the transformer models. Going through the data I noticed an emoji in one of the responses. Initially after seeing the emoji I was planning to use the demoji library in python to convert them to readable text but there were only a handful of them, so I manually changed them. The free text responses were also separated and merged into a Json file (‘TextForTopic.json’) for topic modelling later.

8. Sentimental Analysis

Using the pre trained Bert model tokenizer, we can find the sentiment of all the free text responses. The tokenizer is applied to two different set of responses: Health sentiment responses and dietary health sentiment as shown in table 2.

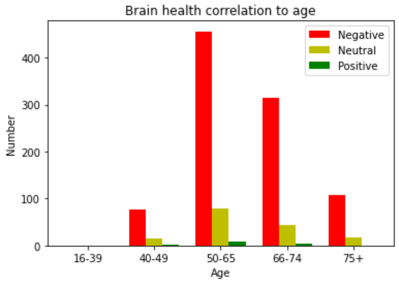
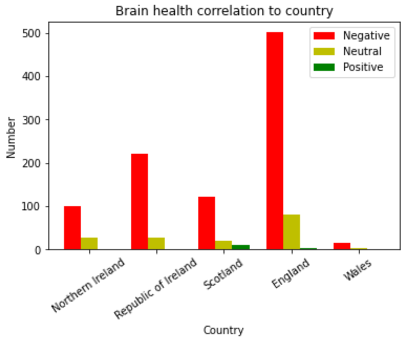
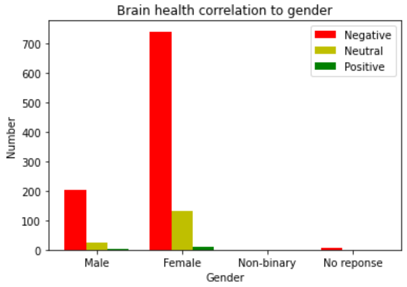
Sentiment	Health Sentiment	Dietary Sentiment
Negative	955	394
Neutral	156	814
Positive	13	4

Table 2 Number of sentiments

From the results we can see that most of the responses had negative or neutral sentiment to the pandemic, only a handful of responses were positive. For health sentiment, an overwhelming number of responses were negative but for dietary sentiment most of the responses were neutral.

8.1 Correlation to demographic differences

After finding sentiment for each response, it was then sorted to the different demographic's variables e.g., age, country, gender to find any correlation. The results are then plotted into a graph using the matplotlib library in python.

	Variables	Negative	Neutral	Positive
	16-39	0	0	0
	40-49	77	16	0
	50-65	456	78	9
	66-74	314	44	3
	75+	108	18	0
	Northern Ireland	97	25	0
	Republic of Ireland	220	28	3
	Scotland	121	20	2
	England	501	80	7
	Wales	16	3	0
	Male	205	24	2
	Female	742	131	9
	Non-Binary	0	0	0
	No Responses	8	1	1
	White	931	152	11
	Chinese	1	0	0
	Irish Traveler	0	0	0

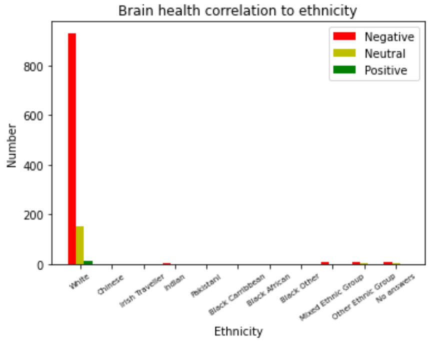
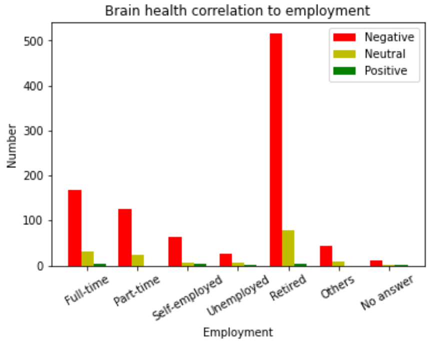
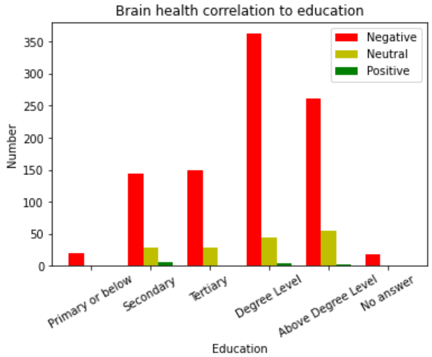
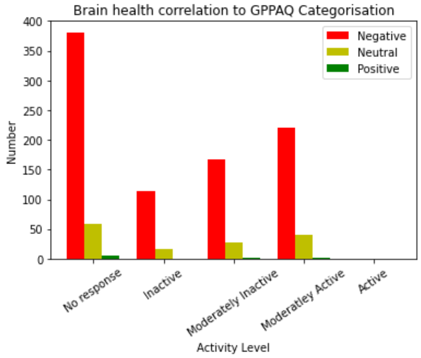
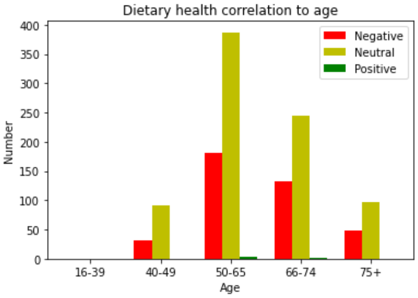
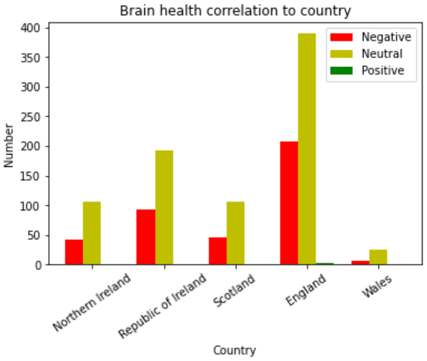
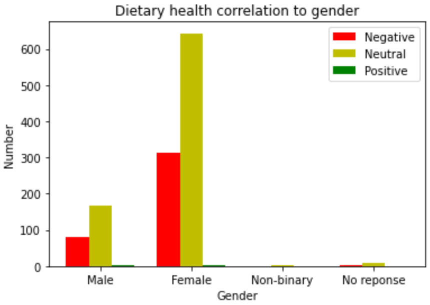
	Indian	2	0	0
	Pakistani	0	0	0
	Black Caribbean	0	0	0
	Black African	0	0	0
	Black Other	0	0	0
	Mixed Ethnic Group	6	0	0
	Other Ethnicity Group	6	2	0
	No Answer	9	2	1
	Full-Time	169	31	2
	Part-Time	126	24	0
	Self-Employment	63	7	3
	Unemployment	27	6	1
	Retired	515	78	5
	Others	44	8	0
	No Answer	11	2	1
	Primary or below	20	0	0
	Secondary	143	29	6
	Tertiary	150	28	0
	Degree	363	44	3
	Above Degree	261	54	2
	No Answer	18	1	1
	No Response	381	59	5
	Inactive	114	16	0
	Moderately Inactive	167	27	2
	Moderately Active	221	41	2
	Active	72	13	3

Table 3 Brain Health sentiment

9. Dietary Sentiment

Similar, sentimental analysis of people diet was also done. Most of the responses were neutral and only 4 of the responses were negative as shown in table 4.

9.1 Correlation to Demographic Variables

	Variables	Negative	Neutral	Positive
	16-39	0	0	0
	40-49	31	91	0
	50-65	181	387	3
	66-74	133	244	1
	75+	49	96	0
	Northern Ireland	41	106	0
	Republic of Ireland	93	193	1
	Scotland	46	105	0
	England	208	390	3
	Wales	6	24	0
	Male	80	167	1
	Female	312	643	3
	Non-Binary	0	1	0
	No Responses	2	8	0
	White	384	799	4
	Chinese	1	1	0
	Irish Traveler	0	0	0
	Indian	0	2	0
	Pakistani	0	0	0
	Black Caribbean	0	0	0
	Black African	0	0	0

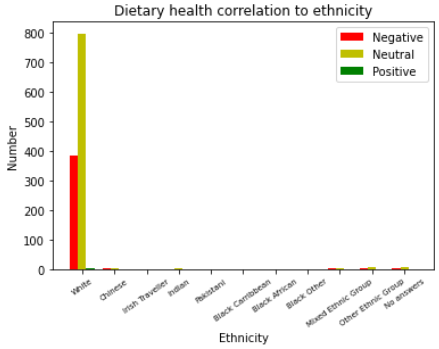
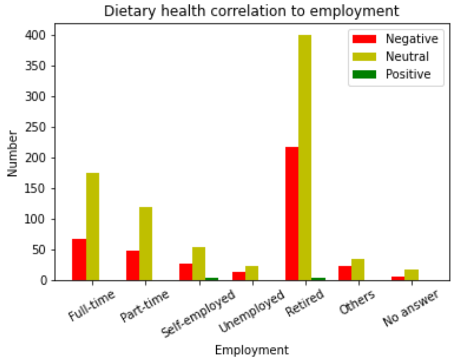
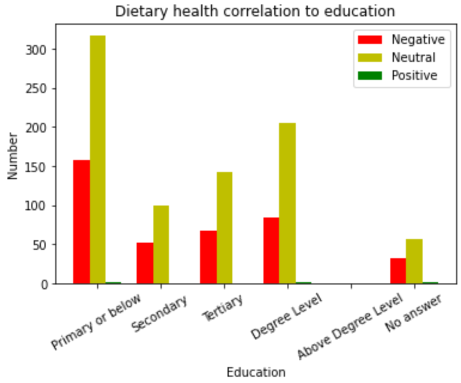
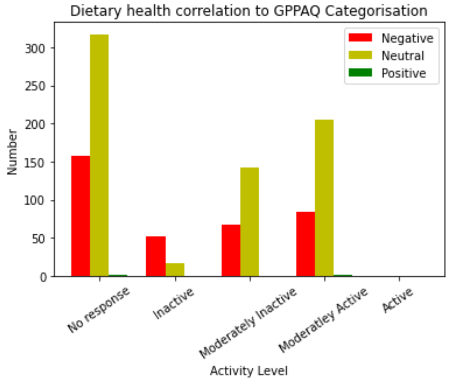
	Black Other	0	0	0
	Mixed Ethnic Group	2	3	0
	Other Ethnicity Group	3	6	0
	No Answer	4	8	0
	Full-Time	66	175	0
	Part-Time	48	119	0
	Self-Employment	26	53	2
	Unemployment	12	23	0
	Retired	216	398	2
	Others	22	34	0
	No Answer	4	16	0
	Primary or below	158	317	2
	Secondary	59	155	4
	Tertiary	68	125	0
	Degree	147	295	0
	Above Degree	106	223	0
	No Answer	4	12	0
	No Response	158	317	2
	Inactive	52	99	0
	Moderately Inactive	67	142	0
	Moderately Active	84	204	1
	Active	33	56	1

Table 4 Dietary sentiment

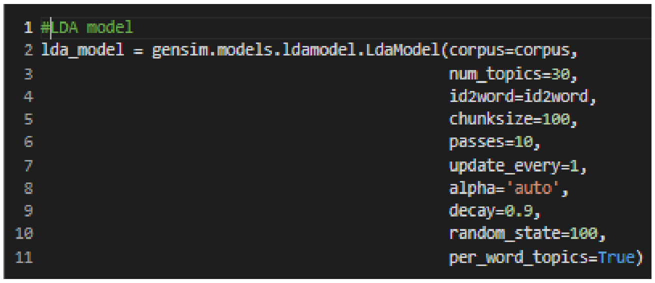
10. Pearson Correlation Coefficient

To find any correlation between brain sentiment and dietary sentiment, Pearson correlation coefficient was used. Pearson correlation measures linear correlation between datasets, it is measured from 0 to 1. The closer its is to 1, the stronger the correlation and vice versa. For our dataset we got a correlation of 0.48 with a p-value of 6.85, which tells us that there is medium correlation between brain and dietary sentiment.

11. Latent Dirichlet Allocation (LDA)

To perform LDA on our dataset some preprocessing must be done before it can be put through a LDA model.

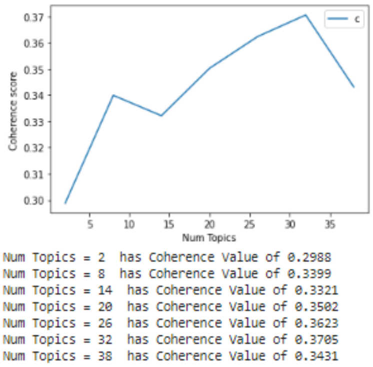
1. Punctuation and unnecessary characters are removed from the text.
2. The texts are then tokenized.
3. Bigram and Trigrams are created. Bigrams are a two-word sequence in a text commonly together. Trigrams are similar but with three words.
4. Stop words are then removed and the words are then Lemmatized.
5. Using the lemmatized words, a dictionary and corpus is created.
6. With the help of Gensim we use the dictionary and corpus to create a LDA model



```
#LDA model
lda_model = gensim.models.ldamodel.LdaModel(corpus=corpus,
                                            num_topics=30,
                                            id2word=id2word,
                                            chunksize=100,
                                            passes=10,
                                            update_every=1,
                                            alpha='auto',
                                            decay=0.9,
                                            random_state=100,
                                            per_word_topics=True)
```

Figure 7 LDA Model Parameters

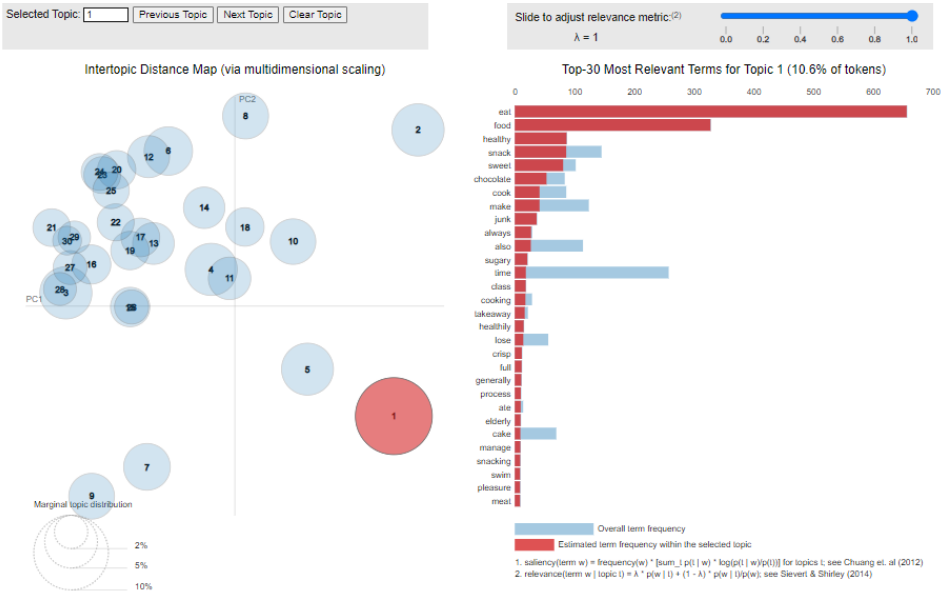
30 topics was chosen as it gave the best coherence value (0.3705) compared to other number of topics as shown in figure.



```
Num Topics = 2  has Coherence Value of 0.2988
Num Topics = 8  has Coherence Value of 0.3399
Num Topics = 14  has Coherence Value of 0.3321
Num Topics = 20  has Coherence Value of 0.3502
Num Topics = 26  has Coherence Value of 0.3623
Num Topics = 32  has Coherence Value of 0.3705
Num Topics = 38  has Coherence Value of 0.3431
```

Figure 8 Coherence value/topic

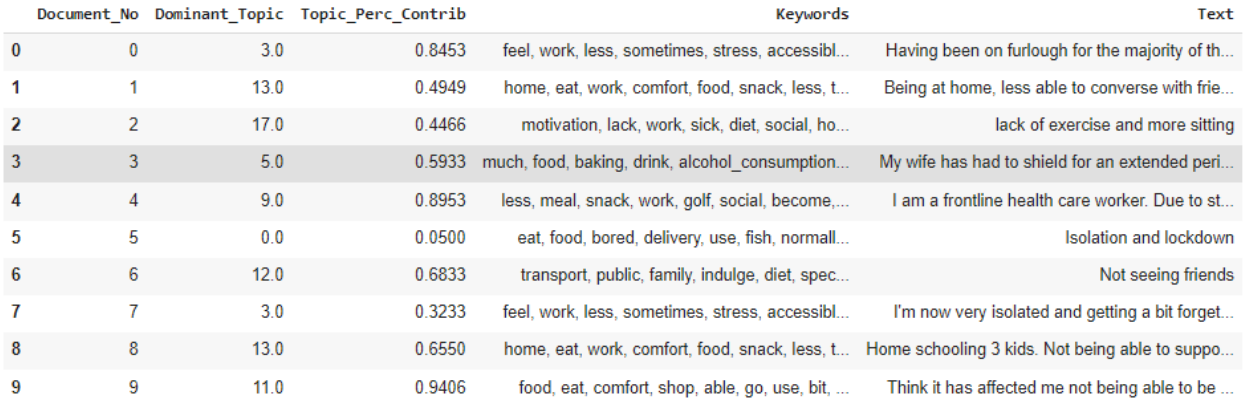
Using the pyLDAvis library in python we can plot it out as shown in figure 8.



1. saliency(term w) = frequency(w) * [sum_t p(t | w) * log(p(t | w)/p(t))] for topics t; see Chuang et. al (2012)
2. relevance(term w | topic t) = λ * p(w | t) + (1 - λ) * p(w | t)/p(w); see Sievert & Shirley (2014)

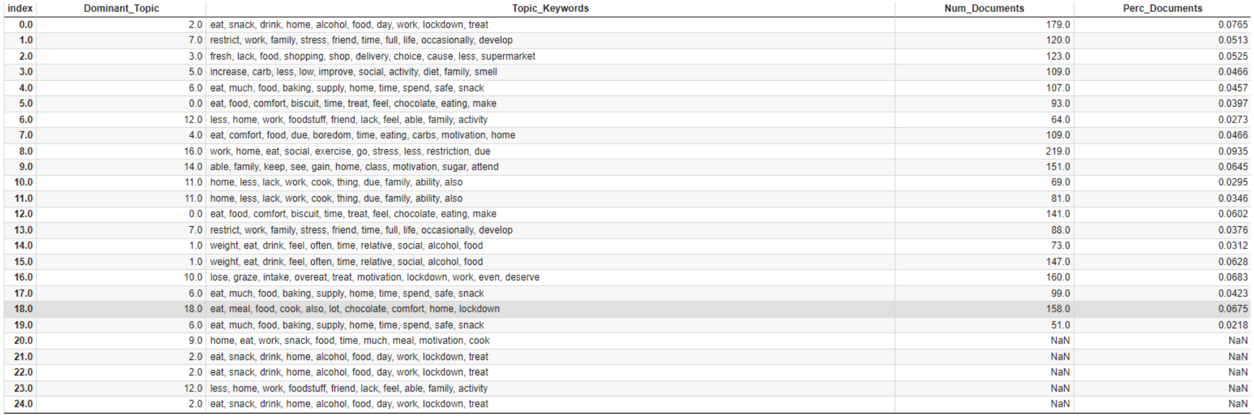
Figure 9 pyLDAvis representation

Each circle on the left grid represents a topic. The larger circle represents a higher number of texts in the corpus about the topic. Circles that are closers together in a cluster are more similar and circles that are further away from each other represent how different they are. On the left graph the blue bar represents the number of times the word appeared in the corpus and the read bar gives an estimation of the times a given term was generated by a given topic. In the bottom right grid, topic 1-5 seems to be about drinking and eating which on the opposite end seems to be negative sentiments.



	Document_No	Dominant_Topic	Topic_Perc_Contrib	Keywords	Text
0	0	3.0	0.8453	feel, work, less, sometimes, stress, accessibl...	Having been on furlough for the majority of th...
1	1	13.0	0.4949	home, eat, work, comfort, food, snack, less, t...	Being at home, less able to converse with frie...
2	2	17.0	0.4466	motivation, lack, work, sick, diet, social, ho...	lack of exercise and more sitting
3	3	5.0	0.5933	much, food, baking, drink, alcohol_consumption...	My wife has had to shield for an extended peri...
4	4	9.0	0.8953	less, meal, snack, work, golf, social, become,...	I am a frontline health care worker. Due to st...
5	5	0.0	0.0500	eat, food, bored, delivery, use, fish, normall...	Isolation and lockdown
6	6	12.0	0.6833	transport, public, family, indulge, diet, spec...	Not seeing friends
7	7	3.0	0.3233	feel, work, less, sometimes, stress, accessibl...	I'm now very isolated and getting a bit forget...
8	8	13.0	0.6550	home, eat, work, comfort, food, snack, less, t...	Home schooling 3 kids. Not being able to suppo...
9	9	11.0	0.9406	food, eat, comfort, shop, able, go, use, bit, ...	Think it has affected me not being able to be ...

Figure 10 Dominant Topics in each text



index	Dominant_Topic	Topic_Keywords	Num_Documents	Perc_Documents
0.0	2.0	eat, snack, drink, home, alcohol, food, day, work, lockdown, treat	179.0	0.0765
1.0	7.0	restrict, work, family, stress, friend, time, full, life, occasionally, develop	120.0	0.0513
2.0	3.0	fresh, lack, food, shopping, shop, delivery, choice, cause, less, supermarket	123.0	0.0525
3.0	5.0	increase, carb, less, low, improve, social, activity, diet, family, smell	109.0	0.0466
4.0	6.0	eat, much, food, baking, supply, home, time, spend, safe, snack	107.0	0.0457
5.0	0.0	eat, food, comfort, biscuit, time, treat, feel, chocolate, eating, make	93.0	0.0397
6.0	12.0	less, home, work, foodstuff, friend, lack, feel, able, family, activity	64.0	0.0273
7.0	4.0	eat, comfort, food, due, boredom, time, eating, carbs, motivation, home	109.0	0.0466
8.0	16.0	work, home, eat, social, exercise, go, stress, less, restriction, due	219.0	0.0935
9.0	14.0	able, family, keep, see, gain, home, class, motivation, sugar, attend	151.0	0.0645
10.0	11.0	home, less, lack, work, cook, thing, due, family, ability, also	69.0	0.0295
11.0	11.0	home, less, lack, work, cook, thing, due, family, ability, also	81.0	0.0346
12.0	0.0	eat, food, comfort, biscuit, time, treat, feel, chocolate, eating, make	141.0	0.0602
13.0	7.0	restrict, work, family, stress, friend, time, full, life, occasionally, develop	88.0	0.0376
14.0	1.0	weight, eat, drink, feel, often, time, relative, social, alcohol, food	73.0	0.0312
15.0	1.0	weight, eat, drink, feel, often, time, relative, social, alcohol, food	147.0	0.0628
16.0	10.0	lose, graze, intake, overeat, treat, motivation, lockdown, work, even, deserve	160.0	0.0683
17.0	6.0	eat, much, food, baking, supply, home, time, spend, safe, snack	99.0	0.0423
18.0	18.0	eat, meal, food, cook, also, lot, chocolate, comfort, home, lockdown	158.0	0.0675
19.0	6.0	eat, much, food, baking, supply, home, time, spend, safe, snack	51.0	0.0218
20.0	9.0	home, eat, work, snack, food, time, much, meal, motivation, cook	NaN	NaN
21.0	2.0	eat, snack, drink, home, alcohol, food, day, work, lockdown, treat	NaN	NaN
22.0	2.0	eat, snack, drink, home, alcohol, food, day, work, lockdown, treat	NaN	NaN
23.0	12.0	less, home, work, foodstuff, friend, lack, feel, able, family, activity	NaN	NaN
24.0	2.0	eat, snack, drink, home, alcohol, food, day, work, lockdown, treat	NaN	NaN

Figure 11 Topic Distribution across documents

From Figure 10 and 11, the dominant topics were mainly about food and restricted social interactions.

12. Discussion and Reflection on Project

Using sentimental analysis and topic modelling we found that many of the respondents had a negative impact on their brain health and dietary health due to the pandemic. For brain health over 80% of the respondents were affective negatively and only about 1% felt positive. Even across different demographics the negative responses to positive responses were proportional. For dietary health, most respondents (~70%) felt neutral but there were still large amounts of negative sentiment. Pearson correlation coefficient between brain and dietary sentiment was calculated. A coefficient of 0.48 which indicated moderate collation between the two subjects. When using topic modeling, we found that most dominant topics were about food and lack of social interactions. This can be inferred that the negative sentiment for the pandemic was mainly due to the lack of social interaction. Reflecting on my work and results so far after, I see that I should have used a more appropriate model for my sentimental analysis or even created my own model to use for comparison against the one I have used to see if I could have increased my accuracy. Optimization of the genism LDA model could also have been improved. Plotting of results, analysis and topic modelling could have also been done using purpose built tools like

Power BI and tabulae for more flexible and clean diagrams/plots. In the future I will revisit this project once I have broadened my knowledge.

References

- Barbieri, F., Neves, L., Camacho-Collados, J. and Espinosa-Anke, L., 2021. *TWEETEVAL: Unified Benchmark and Comparative Evaluation for Tweet Classification*. Available at: https://arxiv.org/pdf/2010.12421.pdf
- Batra, H., Punn, N., S. Sonbhadra, S., K., Agarwal, S., 2021 BERT-Based Sentiment Analysis: A Software Engineering Perspective
- Blei D., B., Ng A., Y. Jordan M., I. (2003) Latent Dirichlet Allocation