Can A Person's MBTI Type Be Determined By A Sample Of Their Writing?

Parinita Kadamby
pkadamby.2023@gmail.com

ABSTRACT

Personality typology, in this case MBTI, can give us great insight into how a person thinks and operates and what motivates that person's decisions. Although typology is most commonly seen in pop psychology and used for fun, typing backed by more empirical evidence and less by self-reported answers, which rely solely on the self-awareness and unbiasedness of the person, could give a more accurate insight into how people fit into the theory of cognitive functions, which MBTI is based on. This project aims to lessen the reliance of self-report for personality typology through an artificial intelligence algorithm that can type people as one of the 16 MBTI types using an unedited writing sample by that person. The model used in this project is a transformer model, and the dataset used is a Kaggle dataset consisting of 8000 people's personality types and 50 written samples from each person. The results are presented through a Streamlit app that uses the model to assign a personality type according to inputted text. The accuracy measured while testing this model is approximately 57.64%, which means the model correctly classified over half the text samples it was tested on. This indicates that there is a pattern that the model is picking up on, other than complete chance, that connects a person's writing to their MBTI type. However, the writing is not 100% predictive of the person's personality type.

1. INTRODUCTION

Many places of schooling and work have students and/or employees take personality tests to determine their strengths, weaknesses, and styles of work/communication. These manual personality tests (a) take a lot of time to fill out and (b) rely on the self-awareness and unbiasedness of the test taker. The time and energy required can oftentimes act as a deterrent since at a certain point many would rather just get it over with rather than take the time necessary to provide accurate answers. Additionally, questions in these personality tests can often have multiple interpretations, resulting in even the most self-aware person to report inaccurately due to a misunderstanding. When this problem is solved, schools and workplaces can spend less time on personality assessments and get more accurate results.

This project aims to combat the problems with current personality tests through a machine learning model that is trained to identify a person's personality type using samples of text written by said person. Since the model only requires text already written by the person, little to no time would be required. Additionally, no self-report would be necessary, so this combats both issues with personality tests that are faced today.

2. LITERATURE REVIEW

There has already been some work done on typing personality via text classification. Previously done work has achieved accuracy up to 63%, mainly for prediction of Big Five traits, while MBTI is a lesser explored personality scale. In this work [1], MBTI has been further explored, and accuracies up to 88% have been observed. The model used in that work is a Random Forest model; therefore, this work differs by using a transformer model to classify the text into the 16 personality types.

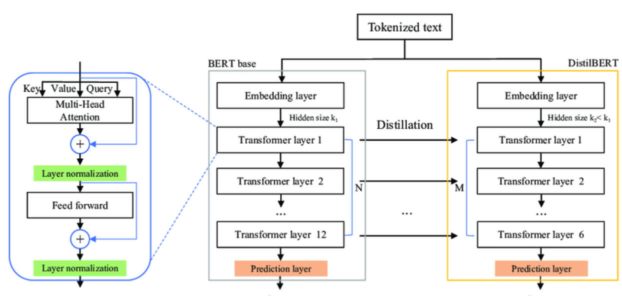
Challenges faced in these past attempts include outdated models, lack of tone recognition, and lack of weightage of words. The model used in this project is a Transformer model, which is a more cutting edge model. Additionally, the data is visualised in world clouds, which display words in different sizes according to how often they appear in posts by a certain personality type; this can help decide which words to increase and decrease the weights of, should the words be weighted.

3. METHODS

3.1 (distil)BERT

In order to build the text classifier, we used a HuggingFace tutorial for a transformer model, which we modified appropriately to fit this project.

The model we used is distilBERT, a transformer model that utilises a technique called distillation to approximate the output of BERT's large neural network using a smaller one.



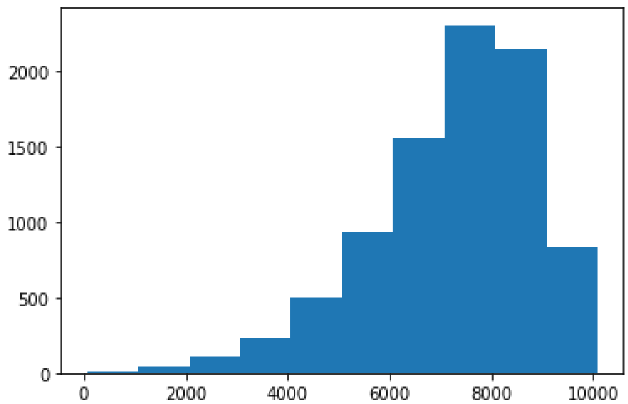
The model that distilBERT uses as its base, BERT, stands for Bidirectional Encoder Representations from Transformers. It was developed by Google with the goal of helping computers understand human language (in this case, English).

This model used transfer learning, meaning that it drew from a pretrained model (BERT) and tweaked it instead of building a model from scratch. We tailored BERT to be able to classify text samples into one of sixteen personality types.

3.2 DATASET

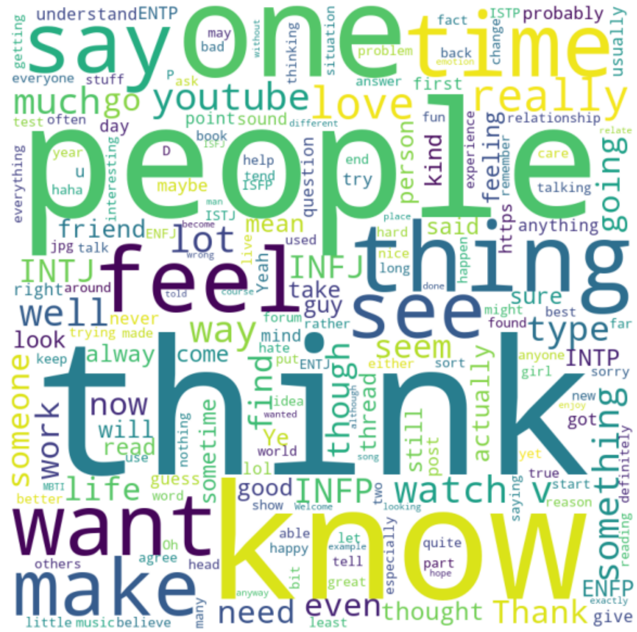
The dataset I used for the project was from Kaggle. It consists of 8600 people, providing each person's MBTI type and their last 50 posts from PersonalityCafe, a forum in which people talk about the different types of personality typology. To train the model on this data, I assigned a number to each of the sixteen personality types.

In a histogram, we plotted the lengths of the posts in the dataset. The x axis indicates the number of characters in a post, and the y axis indicates the number of posts with that many characters in it.

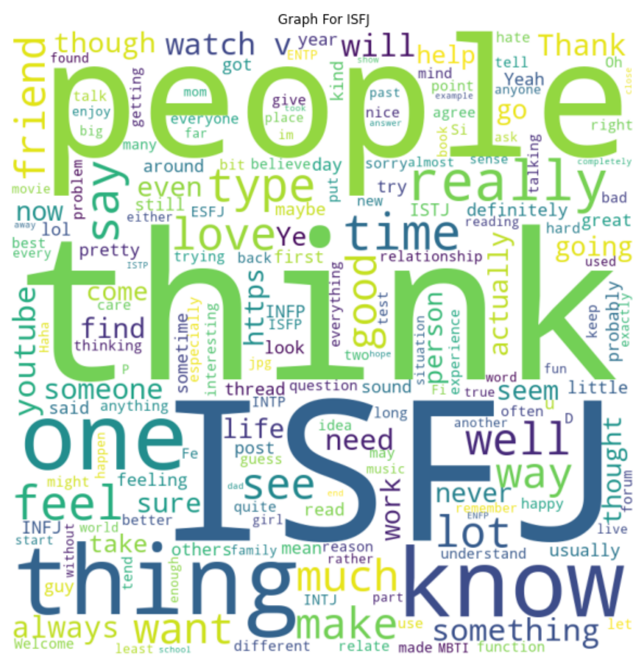


Another data visualisation method we employed is the word cloud. We generated a word cloud for each personality type to find out which words were used most commonly in their posts and if there were any differences in word choice that indicated a particular personality type.

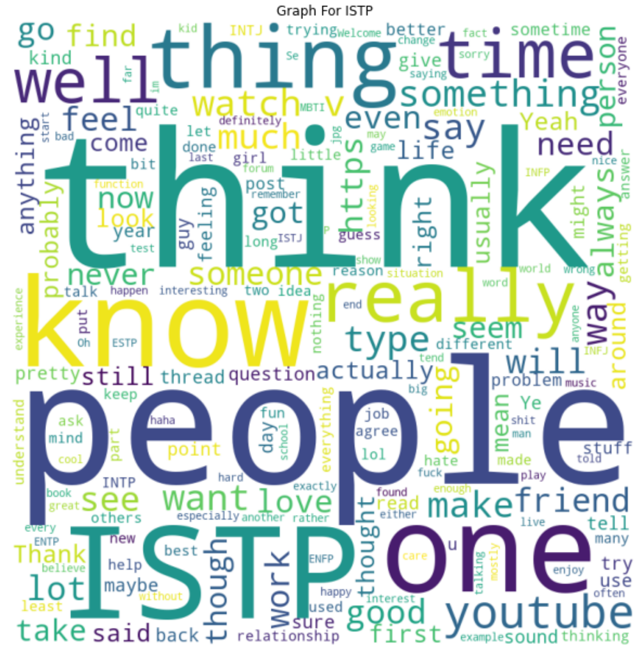
Word Cloud Generated Using All Posts By All Personality Types:



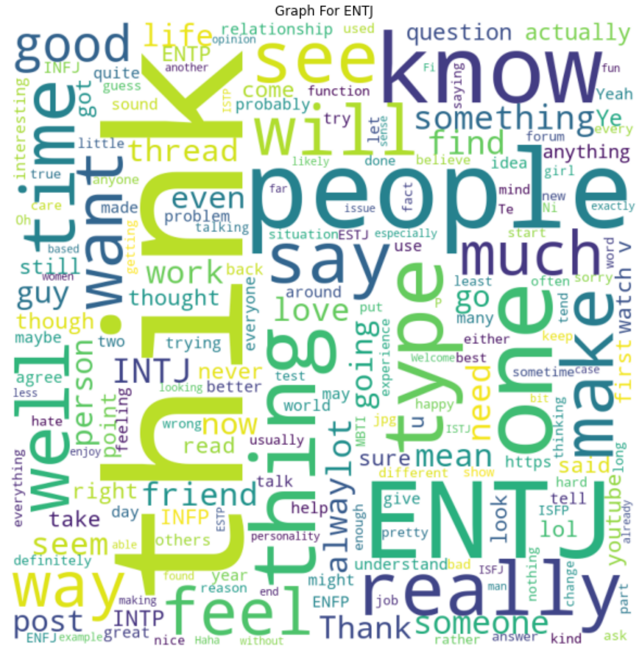
ISFJ:



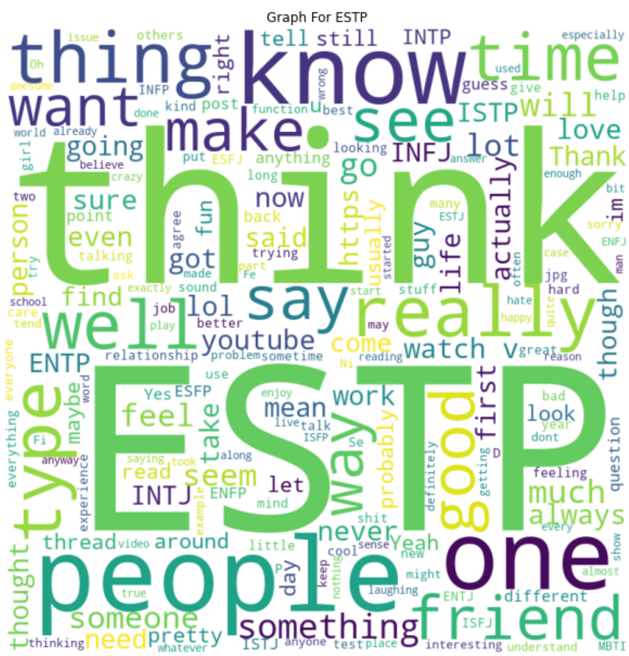
ISTP:



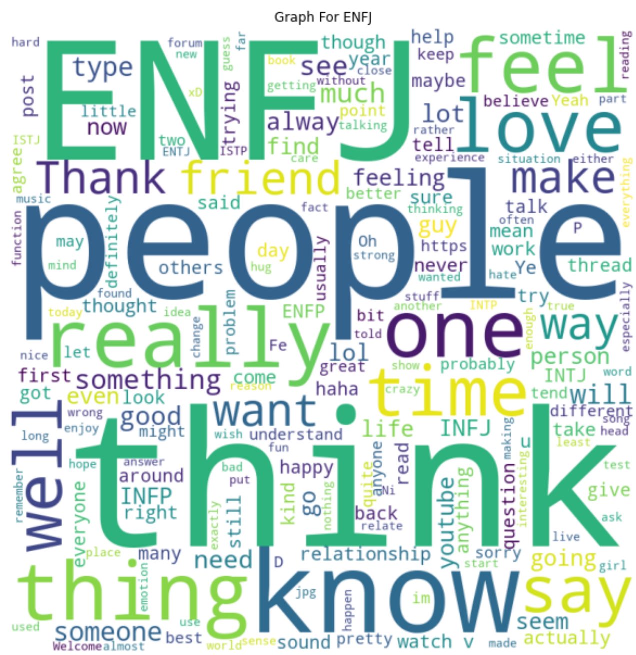
ENTJ:



ESTP:

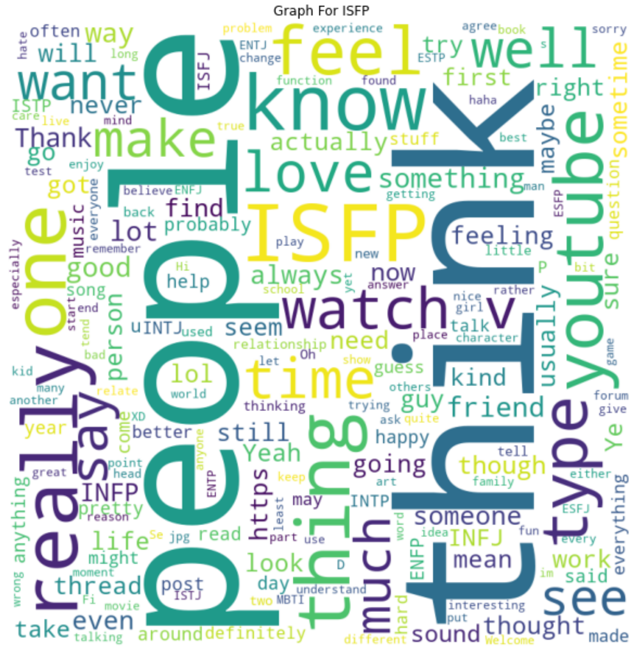


ENFJ:

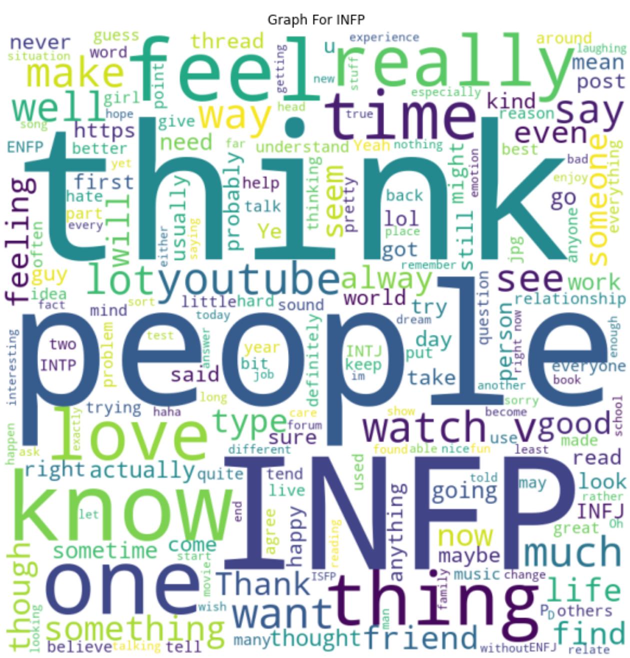


ESFP:

ISFP:



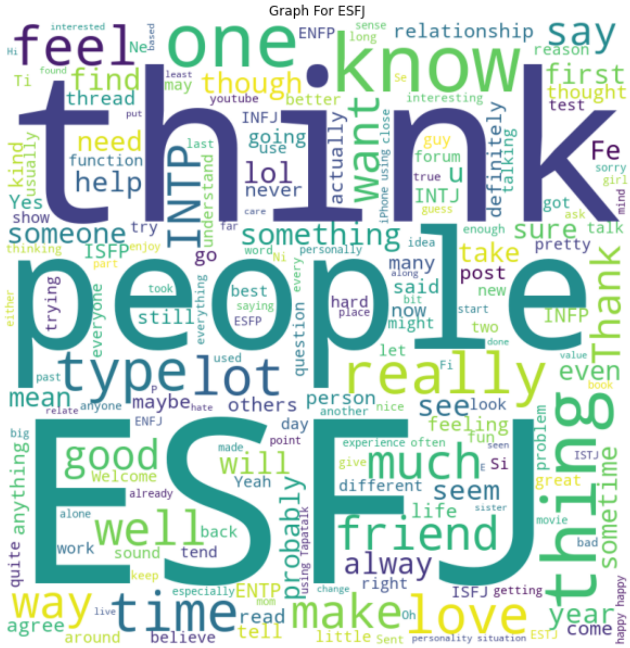
INFP:



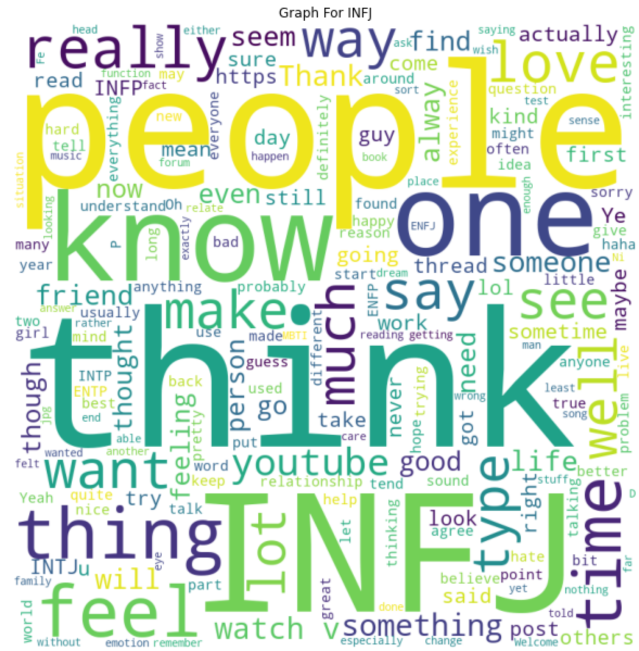
INTP:

ISFP:

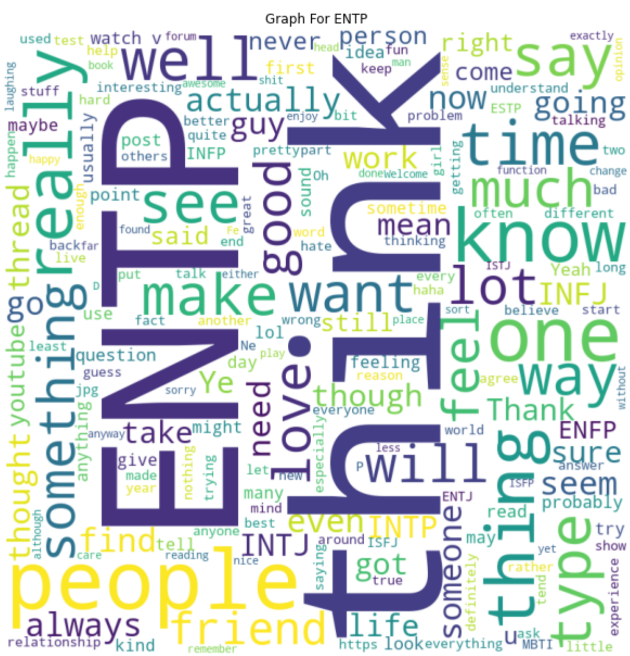
ESFJ:



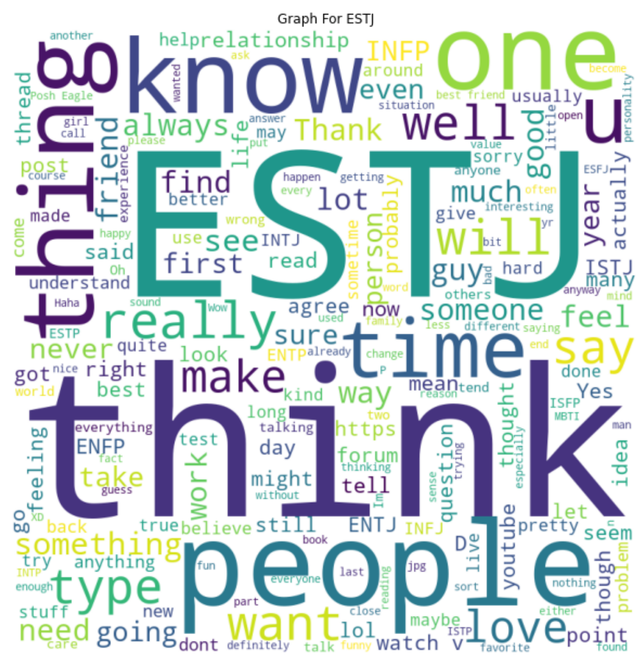
INFJ:



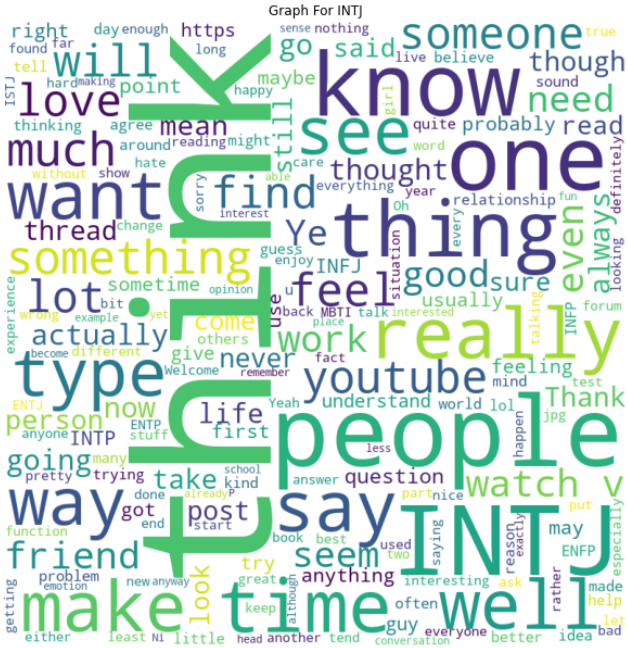
ENTP:



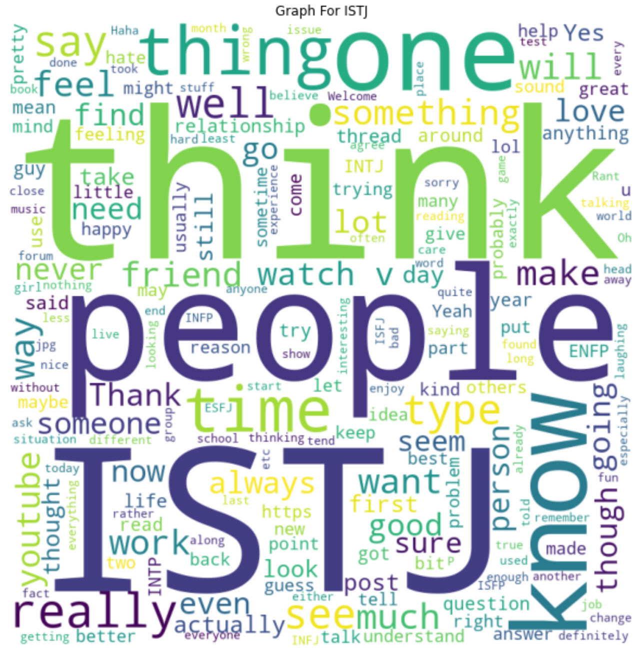
ESTJ:



INTJ:

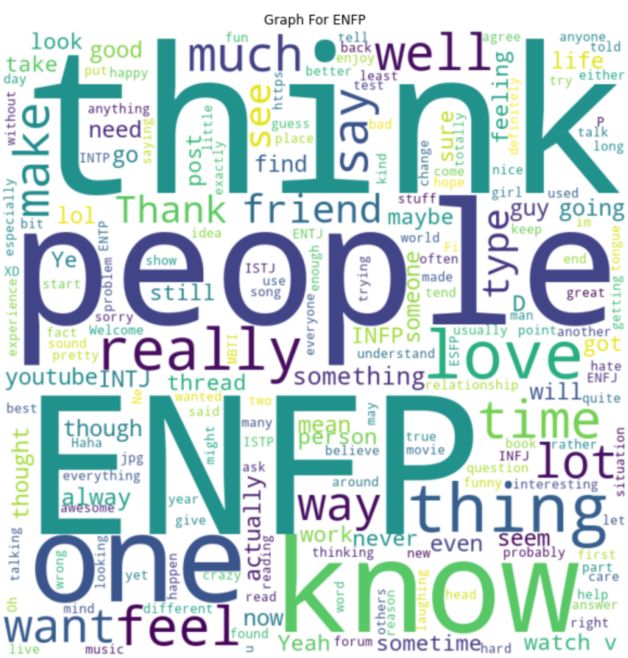


ISTJ:



Words like “people” and “think” seem to be commonly used words all across the board. “Know” and “one” also seem to be common all across the board, though to a lesser extent, indicated by the smaller size. Since PersonalityCafe is a forum on the topic of personality itself, the words “people” and “think” together could indicate the discussion of what other (people) think about a certain personality type or aspect of personality. It could also indicate discussion of what the speaker thinks about people.

ENFP:



4. RESULTS

The accuracy measured while testing this model is approximately 57.64%, which means the model correctly classified over half the text samples it was tested on. This indicates that there is a pattern that the model is picking up on, other than complete chance, that connects a person’s writing to their MBTI type. However, the writing is not 100% predictive of the person’s personality type. The precision, which determines the number of true positives versus false positives, was approximately 0.574. The recall indicates the ratio of the number of correctly classified positive samples to the number of total classified positive samples. In this project, the recall score was 0.576. Lastly, the F1, which is a combination of the precision and recall, was 0.5709.

5. LIMITATIONS OF THE STUDY

Tone-dependent factors like sarcasm or irony are difficult if not impossible for AI to detect, and these are necessary to determine the mood and underlying message of a text sample (which is our input). Additionally, we cannot guarantee that the people in the dataset have been typed correctly, which can screw what our model takes for granted is true. Furthermore, we will need much more data to train the model on to obtain more accurate results.

6. FUTURE WORK

Since MBTI is based on Carl Jung's theory of cognitive functions, a potential way to improve this model is to utilise the principles of that theory (cognitive functions). Instead of having one model choose from a selection of 16 personality types, we can train it on 8 dominant functions (Fe, Fi, Te, Ti, Ne, Ni, Se, Si) and then, depending on the output, have the text be typed by a subsequent model that is focused on the two personality types with that dominant function.

7. REFERENCES

[1] S. Bharadwaj, S. Sridhar, R. Choudhary and R. Srinath, "Persona Traits Identification based on Myers-Briggs Type Indicator(MBTI) - A Text Classification Approach," 2018 International Conference on Advances in Computing, Communications and Informatics (ICACCI), 2018, pp. 1076-1082, doi: 10.1109/ICACCI.2018.8554828.

[2] (2022). Retrieved 30 August 2022, from https://ap.themyersbriggs.com/content/Research%20and%20White%20Papers/MBTI/MBTI_Social_Media_Report.pdf

[3] uClassify - Free text classification. (2022). Retrieved 30 August 2022, from https://www.uclassify.com/browse/?searchString=myers

[4] Data Science Final Project: Myers-Briggs Prediction. (2018). Retrieved 23 September 2022, from https://medium.com/@bian0628/data-science-final-project-myers-briggs-prediction-ecfa203cef8

[5] The Distilbert model architecture and components. Available at: https://www.researchgate.net/figure/The-DistilBERT-model-architecture-and-components_fig2_358239462 (Accessed: December 17, 2022).