COMP4211 Project Report
MBTI Personality Prediction from Social Media Profiles

LIU Runsheng
SID: 20944989
rliuar@connect.ust.hk

ZHANG Hongyi
SID: 20945402
hzhangdl@connect.ust.hk

Abstract

In this project, we explore the prediction of Myers-Briggs Type Indicator (MBTI) personality types based on social media profiles. Using a multimodal approach, we integrate textual, visual, categorical, and numerical data derived from user posts, profile avatars, and engagement metrics to create a comprehensive predictive model. The framework includes two main components: a Transfer Learning Module, which leverages BERT for textual data, and a Multimodal Learning Module, which combines all data modalities to capture relationships between user behaviors and MBTI labels. Experimental results demonstrate high accuracy in predicting the Feeling/Thinking (F/T), Intuition/Sensing (N/S), and Judging/Perceiving (J/P) dimensions and relatively lower performance in classifying Extroversion/Introversion (E/I). These results provide insight into the relationship between personality traits and online behaviors, while presenting avenues for further exploration to improve predictive accuracy and understand underlying correlates. This study bridges psychology and machine learning, showing the potential of integrating different types of data for personalized applications in digital environments.

1 Introduction

Social media platforms have become an integral part of modern life, offering unique opportunities to observe and analyze human behavior. Platforms like Instagram, Twitter (now named as X), and Facebook provide a digital reflection of individuals' personalities through their posts, profile descriptions, and interactions with others. This treasure trove of data has inspired interdisciplinary research that blends psychology, computational linguistics, and machine learning. A particularly fascinating application is predicting personality traits, such as those defined by the Myers-Briggs Type Indicator (MBTI), from social media activity.

The MBTI is a well-known psychological framework that categorizes individuals into 16 distinct personality types based on four binary dimensions: Extroversion/Introversion (E/I), Intuition/Sensing (N/S), Feeling/Thinking (F/T), and Judging/Perceiving (J/P). These dimensions capture fundamental aspects of personality, such as social interaction preferences, decision-making styles, and ways of processing information. For example, an extroverted individual may exhibit traits of openness and social engagement, an intuitive thinker may exhibit tendencies toward creative and abstract reasoning, and so on. Understanding these characteristics can help individuals gain awareness and support applications in areas such as recruiting, personalized marketing, and mental health assessments. For social media companies, this type of application can help conduct user personality research for a better user experience, as well as improve the recommendation mechanism. Furthermore, by analyzing a person's profile to gain an initial understanding of their personality, individuals can approach personal interactions with greater ease and confidence, making social relationships more fluid and enjoyable.

Recent advances in machine learning have enabled significant progress in personality prediction, with researchers using various types of social media data. For example, linguistic analysis of text messages has revealed that word choice, emotional tone, and content themes can indicate personality traits. Similarly, visual data, such as profile avatars, can provide information about an individual's preferences and emotional state. In addition, numerical data, including follower counts, friend connections, and engagement metrics, provide valuable contextual insights into social interactions and reach.

However, existing studies often focus on a single modality, such as text, which limits their ability to fully capture the complexity of human personality. This project addresses this gap by using a multimodal approach that integrates textual, visual, and numerical data to predict MBTI personality types. By combining different data sources, we aim to develop a richer and more nuanced understanding of the relationship between online behaviors and personality traits.

The core of our framework includes two modules: the Transfer Learning Module and the Multimodal Learning Module. The Transfer Learning Module leverages the Bidirectional Encoder Representation by Transformers (BERT) model [DCLT19], a modern natural language processing framework, to extract meaningful features from text data. The Multimodal Learning Module combines these features with information from images and numerical data, allowing the model to identify correlations between modalities and their relationships with MBTI dimensions. This holistic approach provides a comprehensive analysis of user profiles.

Through experiments, we demonstrate the effectiveness of this framework, achieving high prediction accuracy in certain dimensions (F/T, N/S, and J/P). The E/I task achieves relatively lower performance, pointing to the need for further exploration into the nuances of these traits and their representation in online behavior. These findings not only highlight the complexity of human personality but also open new avenues for leveraging multimodal data to improve predictive models.

2 Related Work

The task of predicting MBTI personality types from social media data has received increasing attention in recent years, with researchers exploring different methodologies and data sources. This section reviews relevant studies, highlights their strengths and limitations, and highlights the novelty of our approach.

Recent studies have mainly focused on using text data from social media to predict MBTI personality types. For example, Naik et al. (2022)[NDD+22] applied long-term memory (LSTM) networks to analyze user-generated posts, and succeeded in identifying personality traits based on linguistic patterns. This study demonstrated the effectiveness of deep learning models in processing sequential text data, revealing that language use provides significant clues about individuals' personalities. Similarly, a Stanford University project [Uni22] used linguistic features, such as word frequency and sentiment analysis, to create predictive models for MBTI dimensions. This work established the importance of text data in personality prediction, highlighting its ability to capture user preferences, thought processes, and emotional tones.

However, a major limitation of the text approach is its reliance on a single data modality, which limits the scope of personality information captured. While language reveals valuable features, it often leaves out contextual clues that can be provided by visual or numerical data. This limitation also extends to the datasets primarily used in these studies, such as those available on Kaggle or GitHub. Although rich in textual features, these datasets generally lack diversity in data types, offering limited scope for exploration of the multimodal nature of human personality. Additional resources on GitHub and Kaggle have provided datasets and models that rely solely on posts to predict MBTI types [Ved22, kag22].

Novelty of Application. Unlike previous studies that only focused on textual data from posts, our application extends the use of MBTI personality prediction to a richer and more practical social media context by introducing diverse data types. By incorporating different types of data, such as text, profile avatars, and other numerical information, our approach aims to obtain a more complete picture of online users' personalities, yielding personality predictions that are more suitable for real-world applications. such as personalized interactions on social media and improved user experiences on digital platforms.

3 Methodology

This section includes the description of all the methods and process of dataset obtaining, problem formulation, data exploration, feature engineering, and model construction.

3.1 Dataset

The dataset used is an online public dataset called "Twitter MBTI Personality Types Dataset" [Rai22]. It contains detailed user information and MBTI personality labels of 8386 users. The data is organized as follows:

- User Information: `user_info.csv` comprises 24 columns, including identifiers like ids and screen names, textual information such as descriptions and user locations, categorical features like if an account is verified or not, along with numerical metrics such as count of followers and count of friends.
- MBTI Labels: These are stored in a separate file (`mbti_labels.csv`) and consist of 2 columns: user ids and their corresponding MBTI labels.
- Tweets: `user_tweets.csv` provides a maximum of 200 most recent tweets (including retweets) posted by a user, on or before March 31, 2020.

Notice this dataset does not contain avatar images. To access the images, X API[1] was utilized to download the avatars based on the users' screen names provided in `user_info.csv`. These images are square (i.e., width equals height), but vary in size. During the process, a small portion of the accounts were discovered to be canceled. These data points are removed from the dataset. At last, 6011 data points are obtained, with each having its user information, MBTI label, and avatar image.

3.2 Problem Formulation

In this project, the objective is to predict MBTI personality types using a diverse set of input data characterized by multimodality, including textual, image, numerical and categorical features. The target is the MBTI label, which traditionally consists of 16 personality types. However, instead of treating this as a simple multi-class classification task, we use a multi-label classification strategy. To conclude, the machine learning task involved is **supervised classification task**.

The MBTI classification is inherently structured into four independent dichotomies: Extroversion/Introversion (E/I), Intuition/Sensing (N/S), Feeling/Thinking (F/T), and Judging/Perceiving (J/P). Each dichotomy is represented as a binary classification task. By training four separate classifiers for these dichotomies, we aim to enhance model performance. This approach is informed by insights from previous work [Var20], which suggests that handling these as independent binary tasks can yield better results.

The rationale for this strategy lies in the observed low dependency between the four MBTI classes. This observation is supported by correlation analyses (see Fig. 3), which indicate minimal interdependence among the dichotomies. In addition, the original 16-class MBTI personalities have a very uneven distribution in population (Fig. 1), while

[1] https://developer.x.com/en/products/x-api

as breaking into 4 binary class, the unevenness has been largely reduced (Fig. 2). By isolating each dichotomy into its own classification task, we can tailor the modeling process to the unique characteristics of each category, potentially improving predictive accuracy and model robustness.

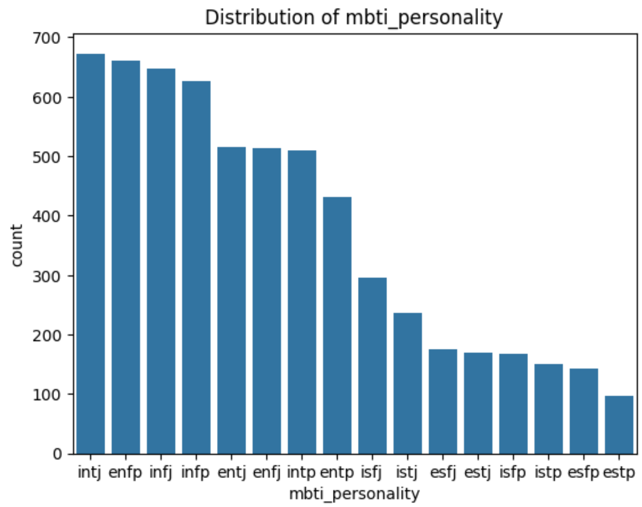

Figure 1: Distribution of the 16 MBTI personality types.

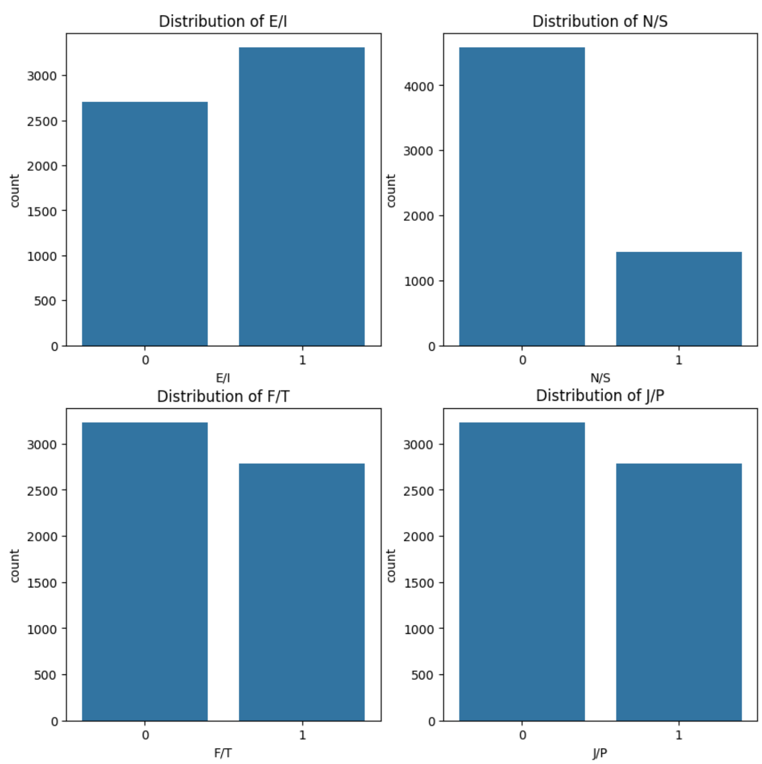

Figure 2: Distribution of the individual divided MBTI labels. The original MBTI labels were decomposed into four binary categories: Extroversion/Introversion (E/I), Intuition/Sensing (N/S), Feeling/Thinking (F/T), and Judging/Perceiving (J/P).For a label `L0/L1`, `L0/L1 = 0` means it belongs to `L0` and `L0/L1 = 1` means `L1`.

3.3 Data Exploration and Preprocessing

During this step, comprehensive exploration and preprocessing of the dataset are conducted to ensure the data are well-prepared for machine learning modeling. Our dataset comprises user information, MBTI labels, and avatar images, providing a rich source of multimodal data, including text, image, numerical, and categorical features.

3.3.1 MBTI labels

As the task was formulated as a multi-label classification problem, the original MBTI labels were decomposed into four binary categories: Extroversion/Introversion (E/I), Intuition/Sensing (N/S), Feeling/Thinking (F/T), and Judging/Perceiving (J/P). Each category was encoded with binary values. For a label `L0/L1`, `L0/L1 = 0` means it belongs to `L0` and `L0/L1 = 1` means `L1`. For example, ESFP can be represented by `E/I = 0`, `N/S = 1`, `F/T = 0`, and `J/P = 1`.

3.3.2 Tweets

Given the computational constraints associated with analyzing up to 200 tweets per user, we reduced the number of tweets to the 30 most recent ones. These tweets were concatenated with the user's description field to form a single text input for processing.

3.3.3 Profile avatars

The profile avatars are resized to the same size ($128 \times 128 \times 3$) for easier training.

3.3.4 User information selection

As `description` is also a textual feature, it is concatenated with user tweets to be precessed together. `id`, `id_str`, `name`, `screen_name`, `location` are assumed not associated with a person's personality and therefore removed. After removal of these labels, all remaining labels do not contain empty entries.

Explore the dataset by examining the correlations between features using a heat map (Fig. 3). `verified` describes if an account is verified or not. A verified account tends to have larger number of followers. Additionally, the correlation analysis also reveals very high correlations among certain features, particularly those related to media counts. Features like `total_retweet_count` and `total_favorite_count` are found to be redundant when compared to their respective average values, prompting us to retain only the average features instead of the sums to mitigate the effects of sample size variability. `number_of_tweets_scraped` is a feature indicating the total amount of the tweets of an account in the dataset and it can be dropped because we only use 30 recent tweets for our study.

In the end, `verified` is the only categorical feature selected together with 14 numerical features: `followers_count`, `friends_count`, `listed_count`, `favourites_count`, `statuses_count`, `number_of_quoted_statuses`, `number_of_retweeted_statuses`, `average_tweet_length`, `average_retweet_count`, `average_favorite_count`, `average_hashtag_count`, `average_url_count`, `average_mentions_count`, `average_media_count`.

3.3.5 Data imputation

There is no need for data imputation as all remaining labels after feature selection do not contain empty entries.

3.3.6 Further processing for numerical features

Draw the distributions of the numerical features selected as shown in the left of Fig. 4. `average_tweet_length` appears to be in normal distribution, while other features display different extend of skewness.

Logarithmic algorithm. Notice that all skewed numerical features are non negative and skewed towards low end. In order to processed these numerical features into normal distribution like to enhance model performances, the logarithmic algorithm is applied to these features. To address the log zero problem, a small value is added to each feature, which equals to the minimum non-zero value of each feature divided by a constant factor, here chosen to be constant `e`.

3.3.7 Train-Test Splitting

The dataset was split into training and testing sets with an 80%:20% ratio. The training set was used for model training, while the testing set served as a basis for evaluating model accuracy and generalization performance.

3.3.8 Standardization

To better manage the scale of numerical features, standardization is utilized to scale them. The resulting features are more normally distributed compared with their raw states and are ready for feeding into the model (Fig. 4).

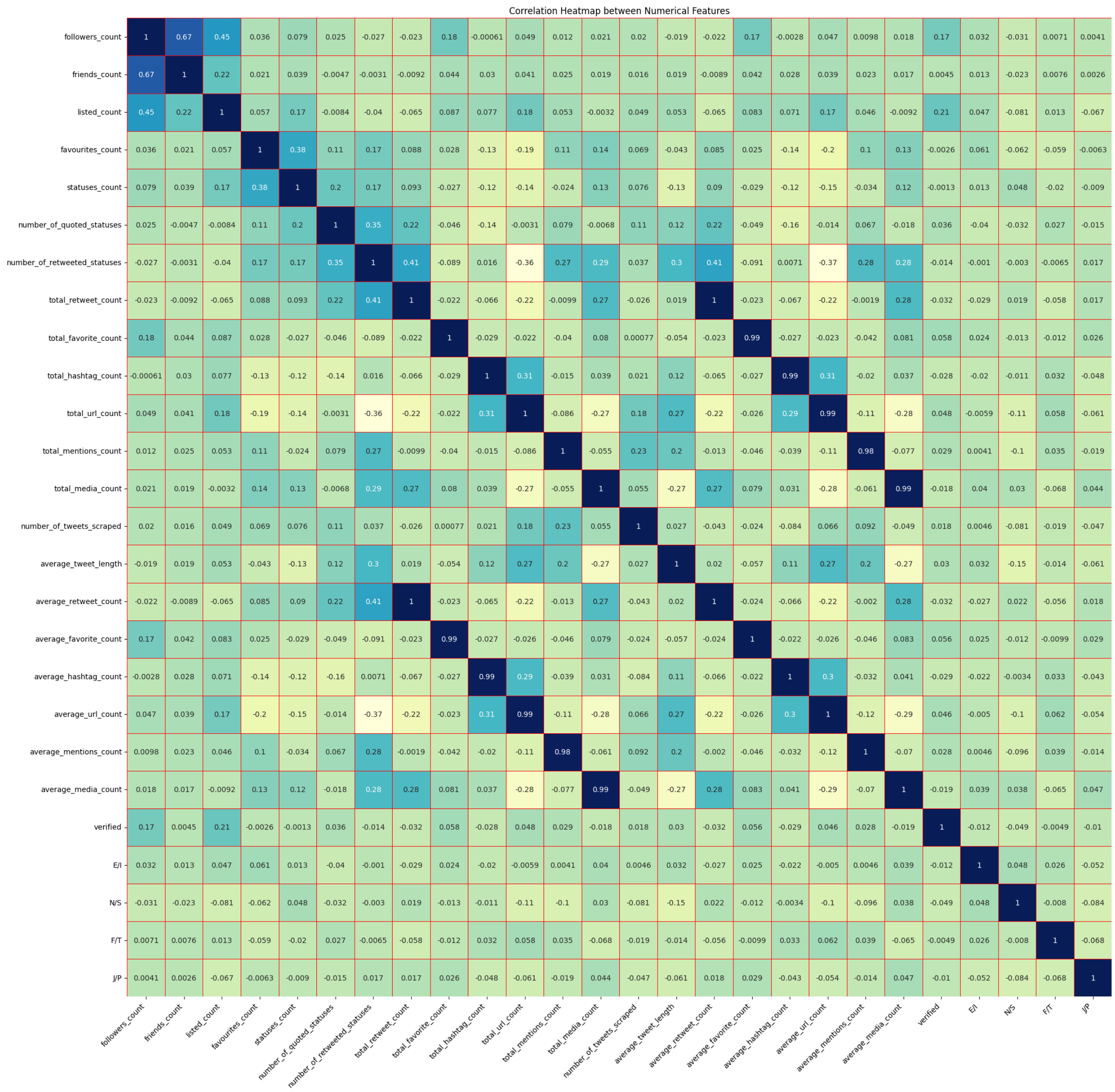

Figure 3: Heat map for the features. Notice certain features of sums and average values have very high correlations (such as `total_retweet_count` and `average_retweet_count`) and only the average ones are retained.

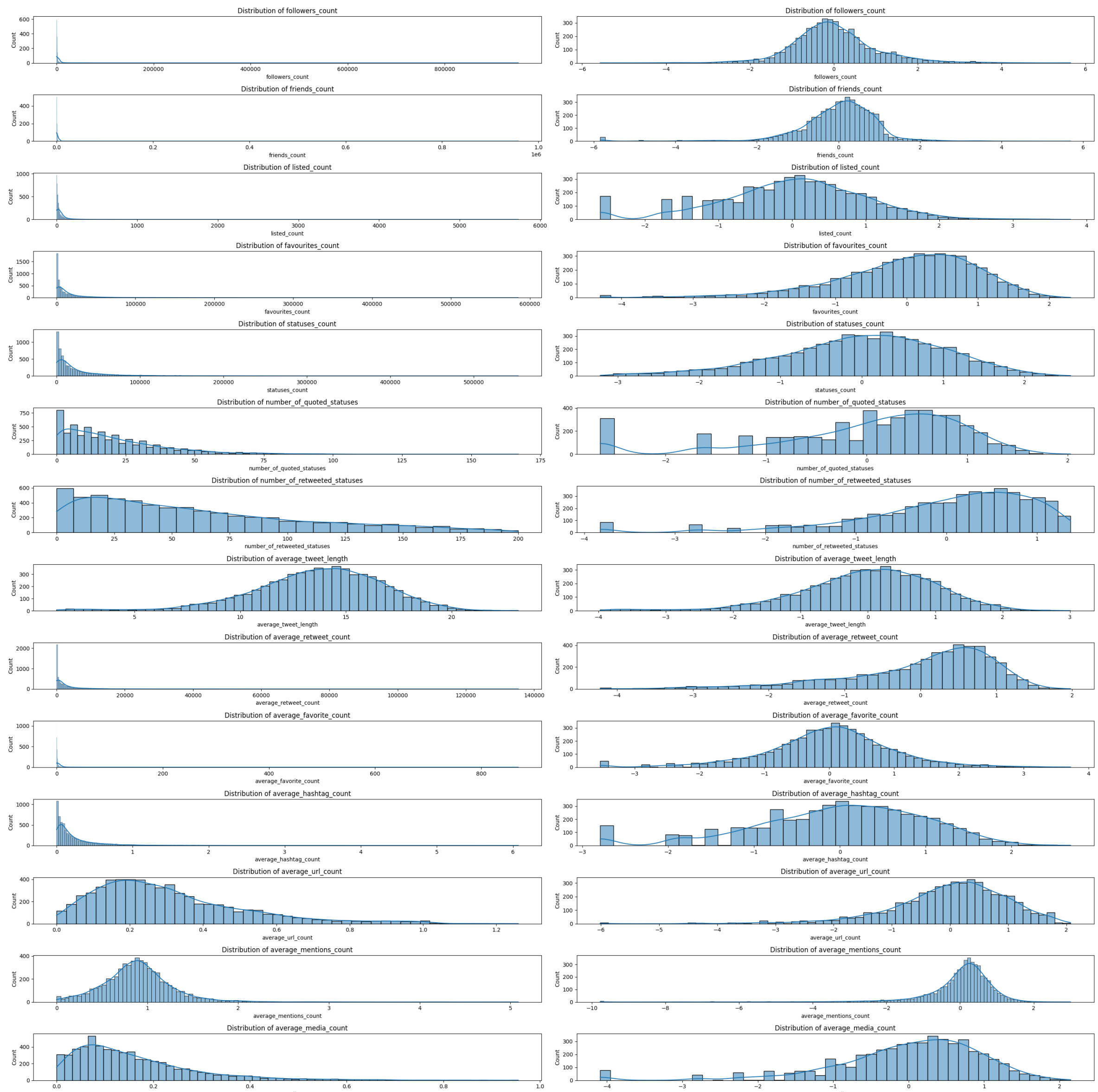

Figure 4: Distribution of selected numerical features. Left side displays the distribution before processing, and the right one is after applying logarithmic algorithm (except `average_tweet_length`) and standardization (for all numerical features).

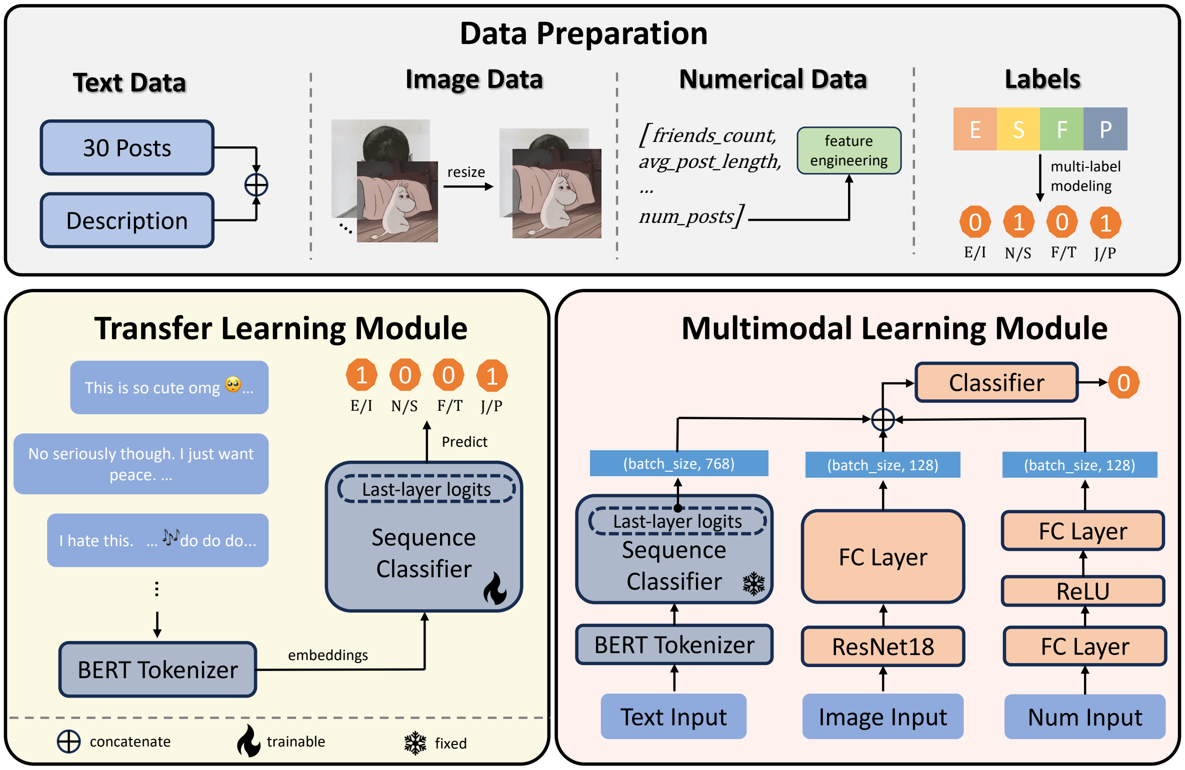

Figure 5: The whole model structure of our proposed method. It contains a data preparation stage, and two separate training stages. The Transfer Learning Module is the fine-tuning of BERT model and the Multimodal Learning Module is to train with three types of data together to predict the target label.

3.4 Model Structure

The architecture of our MBTI personality prediction model is designed to integrate textual, visual, and numerical data into a comprehensive framework, enabling a nuanced understanding of personality traits. The structure of the framework is shown in Fig. 5.

3.4.1 Data Preparation Stage

The model begins with a detailed data preparation stage. Text data, including up to 30 user posts and their profile descriptions, is combined into a single representation to capture linguistic tendencies and thematic content. Image data, primarily profile avatars, is resized for compatibility with downstream image processing tasks, while numerical data, such as follower counts, average post length, and other engagement metrics, is subjected to feature engineering to extract meaningful patterns. The MBTI personality types are represented as four binary labels (E/I, N/S, F/T, P/J), allowing the model to independently classify each dimension.

3.4.2 Transfer Learning Module

The first major component of the architecture is the Transfer Learning Module, which focuses on processing textual data. This module fine-tunes a BERT-based sequence classifier, leveraging its pre-trained contextual embeddings to extract semantic information from user posts and descriptions. By tokenizing the input text, BERT generates embeddings that summarize linguistic features relevant to personality prediction. The model is specifically tuned for the MBTI classification task, with only the final layers trained to adapt BERT's general language knowledge to this specific application. The logits from the final BERT layer predict the four MBTI dimensions and serve as input to the subsequent multimodal module, ensuring that rich textual features are preserved and integrated.

3.4.3 Multimodal Learning Module

The Multimodal Learning Module is the core of the model, where text, images, and numerical data are fused to generate a comprehensive prediction. Text logits from the Transfer Learning Module are integrated into this module in the form of high-dimensional embeddings. The profile avatars, resized during data preparation, are processed through ResNet18 [HZRS15], a convolutional neural network that extracts visual features indicative of personality traits. These features, reduced by fully connected layers, provide information about the user's preferences and the emotional expression conveyed visually. Simultaneously, numerical data, such as social engagement metrics, are passed through their own fully coupled layer with ReLU activation, transforming the raw values into higher-dimensional representations compatible with other modalities.

The results of the three modalities are combined to form a unified functional representation. This multimodal coupling allows the model to capture additional information about different types of data: textual data can highlight cognitive or emotional patterns, visual data can suggest aesthetic preferences, and numerical data provides contextual information about social interactions and activity levels. The concatenated representation passes through additional fully connected layers before being classified into the four binary MBTI labels. By approaching each dimension as an independent classification task, the model avoids the limitation of assuming interdependence between labels, focusing instead on each function separately.

3.4.4 Design Choice

This architecture exemplifies a carefully balanced integration of advanced transfer learning techniques and multimodal fusion strategies. The choice to use BERT and ResNet18 reflects the importance of leveraging pre-trained models to maximize performance with limited labeled data. The ReLU activation for numerical features introduces nonlinearity, capturing complex relationships in user behavior metrics.

Trainable-fixed Strategy. We choose to firstly train the BERT model solely based on textual inputs, and fix the BERT model in the Multimodal training stage. This is to reduce training time and simplifies the multimodal model. If we combines the BERT fine-tuning together with other two training streams, it will be difficult to simultaneously learn the insights of all three modalities and the model will be complex, thus making the training process difficult and time-consuming to converge. Learning the textual features in advance and then combining the other two types of inputs can equip the model with a comprehensive understanding of the relationship between the inputs and target MBTI labels with more ease.

Last-layer Logits Strategy. In the Multimodal Learning Module, the results of the pre-trained sequence classifier are not used directly; instead, the last layer logits are extracted to be further concatenated with other inputs' features. Although the output of the sequence classifier can be further fed into a fully connected layer to map to a higher-level feature vector, this step is not performed since it increases the model's number of parameters and slows down the training process. The last-layer logits are considered already representative and full of information, so the 768-length vector is directly used for concatenation with others.

Overall, the model design ensures that each modality contributes effectively to the final prediction, enabling a holistic and accurate assessment of MBTI personality types based on diverse social media data.

3.5 Experiments

This section includes the experiment environment setup, computing resources (hardware and software) preparation, machine learning methods and parameter settings, experiment results, and the insights obtained.

3.5.1 Environment Setup

The hardware and software of the experiments contain a MacBook Pro with VSCode and Colab. The data preparation is done on MacBook local enironment with basic python package. Feature engineering and model training are done on Colab, with its L4 GPUs. Important packages we import in Colab include pandas [pdt20], sklearn [BLB+13], torch [PGM+19] and transformer from hugging face [WDS+20].

3.5.2 Machine Learning Methods & Parameter Settings

For the first training stage, Transfer Learning Module, the machine learning methods we use is BERT for sequence classification. We fine-tune the BERT sequence classification model to obtain rich textual features that are beneficial for final stage predictions. For the second training stage, we define our own customized Multimodal model, and we

Task	Accuracy%	Precision	Recall	F1-score
E-I	62.09	0.75	0.48	0.59
N-S	94.85	0.99	0.79	0.88
T-F	93.77	0.94	0.92	0.93
J-P	93.35	0.92	0.95	0.93

Table 1: Results of the first time experiments.

apply basic methods like MLP and CNN to do binary classifications. The loss functions for both stage are binary cross-entropy loss function.

Most hyperparameters are set as their defaults. For the first training stage, we train for 3 epochs for each target label with batch size 16. For the second stage's numerical input stream, the first Full-connected layer has output shape (batch size, 64), and the second full-connected layer has output shape (batch size, 128). We train the second stage for 2 epochs per target label with batch size 32.

3.5.3 Results

The result of our first experiments for 4 target binary label is shown in Table. 1.

For the last three binary classification tasks, our model achieves satisfactory performance, with N-S prediction achieving the highest accuracy among all dimensions. However, it is notable that the N-S task has the lowest recall and F1 score, which can probably be attributed to the imbalance in the data set. In particular, the dataset contains about three times more N-labeled samples than S-labeled samples, making it difficult for the model to effectively learn features related to the minority class. This imbalance skews the model's predictions towards the majority class, resulting in reduced performance for metrics like Recall and F1-score, which are sensitive to the correct classification of minority samples.

In contrast, the E-I classification task shows a different trend, with relatively poor performance. During training, we observed that the model did not converge well for this task, exhibiting unstable behavior.

A plausible explanation for this phenomenon is overfitting. During training, the model may initially identify patterns in the data associated with the E-I labels, but as training progresses, it adapts to noise or spurious correlations present in the data set. This overfitting leads to a loss of generalization, which explains the degradation of performance after the first epoch. Another contributing factor may be the inherent weakness of the relationship between E-I labels and input data. If the features associated with the E-I dimension are not strongly represented in the dataset, the model may struggle to establish a strong predictive relationship, resulting in erratic learning behavior and unstable performance.

Results Visualization. We also provide sample tests to demonstrate how our model

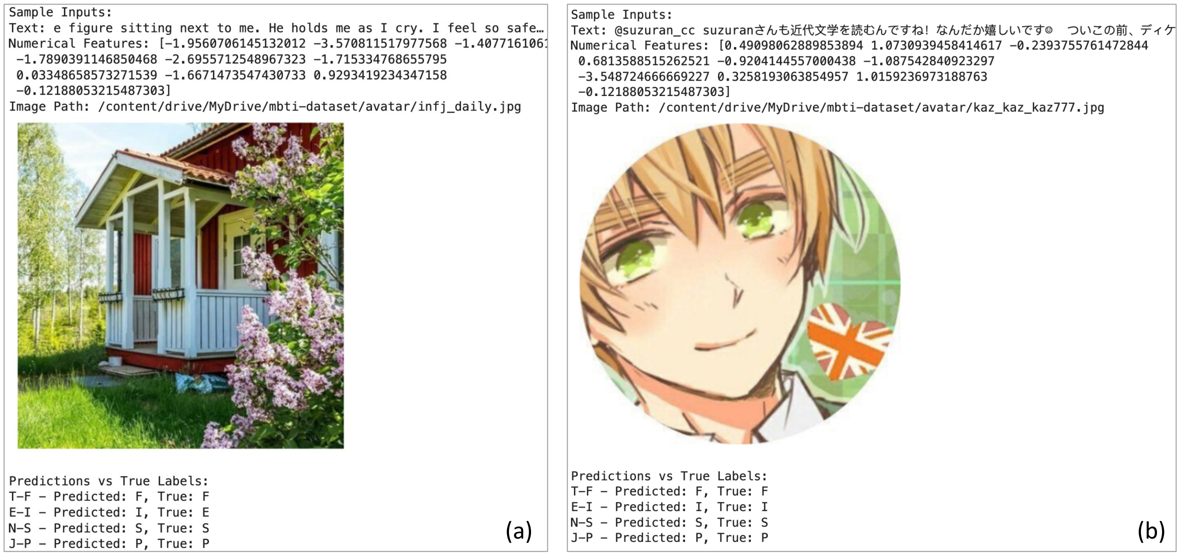

Figure 6: Demonstration of the inputs and outputs of our prediction model. (a) Predict 3 labels correctly. (b) Predict 4 labels correctly.

predicts a person's MBTI from his/her profile information. The reuslts are shown in Fig. 6.

We saved our model weights, including the internal fine-tuned BERT models and the final multimodal models of the 4 tasks. One can obtain all the models and also the datasets from this Google Drive.

3.5.4 Discussion

Unlike other MBTI dimensions, the traits of extroversion (E) and introversion (I) are less likely to be clearly reflected in an individual's language style or superficial online behaviors. According to the MBTI framework, these traits are primarily related to how individuals gain energy and find satisfaction in real life. For example, E types generally feel energized by interacting with others, while I types tend to recharge through solitude. However, these differences are difficult to observe through social media profiles.

Interestingly, some introverts may actively express themselves online to compensate for limited social interaction in real life, while others may feel too reserved or hesitant to engage digitally. This variability highlights the complexity of capturing the online behavioral patterns of Types E and I. However, other types may be more strongly influenced by language styles, avatar choices, and online activities.

These results suggest that improving the E-I classification task merits further exploration. Possible solutions include mitigating overfitting through data augmentation, applying class balancing strategies, or exploring architectures that better accommodate subtle relationships between features. Furthermore, recognizing the inherent difficulty in extracting E-I features from the current dataset reinforces the need for richer, more focused data that can better reflect the essence of this dimension. Understanding these limitations provides valuable insights into the model's performance and the broader applicability of MBTI prediction from online behaviors.

4 Conclusion

In this project, we explored the prediction of Myers-Briggs Type Indicator (MBTI) personality types using a multimodal approach that integrates textual, visual, categorical, and numerical data derived from social media profiles. Using advanced machine learning techniques, including a Transfer Learning Module and a Multimodal Learning Module, we sought to capture the complex relationships between user behaviors and MBTI labels.

Our experiments demonstrated that the multi-label classification strategy, which decomposes the MBTI labels into four binary tasks — Extroversion/Introversion (E/I), Intuition/Sensing (N/S), Feeling/Thinking (F/T), and Judging/Perceiving (J/P) — yielded high accuracy in predicting the N-S, F-T, and J-P dimensions. However, challenges remain in distinguishing between E-I, indicating the need for improved modeling and data augmentation.

The findings in this project demonstrate the ability to integrate different types of data to understand human characteristics and behavior online. This research combines the fields of psychology and machine learning and demonstrates the transformative potential of personalized applications in digital environments.

One of the main strengths of our approach is the use of the BERT model which has been optimized for text computing. BERT's ability to capture nuanced language features greatly improves the accuracy of personality predictions, as it leverages pre-trained contextual embeddings to extract semantic insights from user-generated content.

However, it is important to point out a potential limitation in the balance of feature importance. While the textual data benefited from the state-of-the-art BERT model, the image and numerical data were processed using comparatively simpler methods. This disparity may lead to a feature importance imbalance, where the textual features could disproportionately influence the model's predictions. Consequently, this might overshadow the contributions of image and numerical data, potentially limiting the model's holistic understanding of personality traits.

Addressing this imbalance in future work could involve exploring more advanced techniques for image and numerical data processing, such as employing more tailor-made neural networks for avatar images and more sophisticated feature engineering for numerical data. By ensuring that all data modalities are given equal consideration in terms of complexity and processing power, a more balanced and comprehensive model can be expected to achieve a better performance.

Division of Contribution

In this project, ZHANG Hongyi mainly contributes to data exploration and processing, while Liu Runsheng is mainly responsible for the model building and training. Regarding the report, abstract, introduction, related work, model structure, and experiments are mainly written by Liu, and dataset obtaining, problem formulation, data exploration, feature processing, and conclusion are mainly written by Zhang. Both authors share their ideas and opinions and cooperate together across the whole project and contributes approximately equally (i.e., 50% each).

References

[BLB+13] Lars Buitinck, Gilles Louppe, Mathieu Blondel, Fabian Pedregosa, Andreas Mueller, Olivier Grisel, Vlad Niculae, Peter Prettenhofer, Alexandre Gramfort, Jaques Grobler, Robert Layton, Jake VanderPlas, Arnaud Joly, Brian Holt, and Gaël Varoquaux. API design for machine learning software: experiences from the scikit-learn project. In *ECML PKDD Workshop: Languages for Data Mining and Machine Learning*, pages 108–122, 2013.

[DCLT19] Jacob Devlin, Ming-Wei Chang, Kenton Lee, and Kristina Toutanova. BERT: Pre-training of deep bidirectional transformers for language understanding. In Jill Burstein, Christy Doran, and Thamar Solorio, editors, *Proceedings of the 2019 Conference of the North American Chapter of the Association for Computational Linguistics: Human Language Technologies, Volume 1 (Long and Short Papers)*, pages 4171–4186, Minneapolis, Minnesota, June 2019. Association for Computational Linguistics.

[HZRS15] Kaiming He, Xiangyu Zhang, Shaoqing Ren, and Jian Sun. Deep residual learning for image recognition. December 2015.

[kag22] Myers-Briggs Personality Type Dataset. https://www.kaggle.com/datasets/datasnaek/mbti-type/code, 2022.

[NDD+22] Hrutik Naik, Shrumi Dedhia, Ashwini Dubbewar, Meet Joshi, and Vandana Patil. Myers briggs type indicator (mbti) - personality prediction using deep learning. In *2022 2nd Asian Conference on Innovation in Technology (ASIANCON)*, pages 1–6, 2022.

[pdt20] The pandas development team. pandas-dev/pandas: Pandas, February 2020.

[PGM+19] Adam Paszke, Sam Gross, Francisco Massa, Adam Lerer, James Bradbury, Gregory Chanan, et al. Pytorch: An imperative style, high-performance deep learning library. *Advances in Neural Information Processing Systems*, 32:8024–8035, 2019.

[Rai22] Sanket Rai. Twitter mbti personality types dataset. https://www.kaggle.com/ds/701530, 2022.

[Uni22] Stanford University. Predicting myers-briggs type indicator with text. https://web.stanford.edu/class/archive/cs/cs224n/cs224n.1184/reports/6839354.pdf, 2022.

[Var20] Rajshree Varma. Twitter mbti personality types dataset. https://www.kaggle.com/code/rajshreev/mbti-personality-predictor-using-machine-learning/notebook, 2020.

[Ved22] Vedant. Mbti personality prediction using machine learning. https://github.com/Vedant-02/MBTI-Personality-Prediction-using-Machine-Learning, 2022.

[WDS+20] Thomas Wolf, Lysandre Debut, Victor Sanh, Julien Chaumond, Clement Delangue, Anthony Moi, Pierric Cistac, Tim Rault, Rémi Louf, Morgan Funtowicz, Joe Davison, Sam Shleifer, Patrick von Platen, Clara Ma, Yacine Jernite, Julien Plu, Canwen Xu, Teven Le Scao, Sylvain Gugger, Mariama Drame, Quentin Lhoest, and Alexander M. Rush. Transformers: State-of-the-art natural language processing. In *Proceedings of the 2020 Conference on Empirical Methods in Natural Language Processing: System Demonstrations*, pages 38–45, Online, October 2020. Association for Computational Linguistics.