Construction of MBTI Personality Estimation Model Considering Emotional Information

Ryota Kishima, Kazuyuki Matsumoto, Minoru Yoshida, Kenji Kita
Division of Science and Technology,
Graduate School of Sciences and Technology for Innovation,
Tokushima University, Japan
c612135039@tokushima-u.ac.jp
{matumoto;mino;kita}@is.tokushima-u.ac.jp

Abstract

The widespread use of social networking services (SNSs) such as Twitter and Facebook in recent years have rendered it easy to interact with various people all over the world. Analysis of the user personality and emotion using features such as the text information and image information on SNSs has been researched. This study analyzes the effectiveness of using emotional information as a feature when building a personality estimation model for Twitter users. The extensively used Myers-Briggs type indicator (MBTI) is used for analysis. A method is proposed for estimating the MBTI personality pattern by extracting emotional information from Twitter and adding it to the characteristics of the distributed expressions of sentences and words. Through evaluation experiments, it is determined whether the emotional information is effective in constructing the MBTI personality estimation model. The obtained results demonstrate that the accuracy is improved when emotional information is included as a feature, compared to the case without emotional information.

1 Introduction

In recent years, it has become possible for anyone to easily interact with many others all over the world because of the prevalence of social networking services (SNSs) such as Twitter and Facebook. Certain services recommend suitable friend candidates to the user based on information such as hobbies, occupation, age, and gender. However, it is difficult to comprehend the personality of the user based on this information alone. Therefore, we investigate the estimation of the user personality based on the comments posted on SNSs.

Bakry et al. (Bakry et al., 2019) described personality estimation based on user information from online social networks (OSNs). They not only summarized several machine learning algorithms, such as the SVM, Naive Bayes, gradient boosting, and neural networks, but also described the features and dataset scale used for training, considering each character estimation model. Some of the models have a high accuracy rate of approximately 90% but as the data set is small, there is no generalization. The other model accuracies are not high, at approximately 60%. Thus, the accuracy of personality estimation based on user information from SNSs is insufficient. However, it is crucial to improve the personality estimation accuracy for the realization of various services on SNSs.

Yamada et al. (Yamada et al., 2019) extracted features from user tweets and built a personality estimation model using a support vector machine (SVM). They utilized the Myers-Briggs type indicator (MBTI), which is an introspective self-report questionnaire. As shown in Table 1, it includes four indicators with two types each, and a total of 16 personality types.

In our research, we not only use the sentence vector extracted from a tweet sentence but also the emotional information estimated from it as a feature when determining the accuracy of the personality estimation model. Emotional information includes the emotions and impressions of human beings when watching a movie or listening to music, and is considered to be closely related to the personality. Okumura et al. (Okumura et al., 2015) analyzed the personality of a blogger based on the Big Five, which is an emotion judgment system that estimates nine emotions, such as “anger”, “sadness”, “joy”, “disappointment”, “regret”, “guilt”, “shame”, “fear”, and “safety” (Tsuchiya et al., 2009).

We attempt to acquire the Emotional information vector from a tweet using the emotion estimation model, and apply it as a new feature. We further confirm whether the estimation information is effective for constructing the personality estimation model. We compare and analyze the results when the personality estimation model is trained using the sentence vector extracted from a tweet sentence, and when the sentence vector included emotion information vector is trained as a feature in personality estimation. The emotional categories targeted in this research are based on the emotional system diagram proposed by Fisher (Fischer et al., 1989), which includes four emotions, namely, “joy”, “surprise”, “anger”, and “sorrow”.

The direction of energy Extraversion (E) or Introversion (I)
How to see things Sensing（S）or Intuition (N)
How to judge Thinking (T) or Feeling (F)
How to interact with the outside world Judging (J) or Perceiving (P)

Table 1: Four indicators and two types of MBTI.

2 Related Work

2.1 Personality Analysis Using Social Media

There are several other studies that estimate the personality from text information, in addition to the previously mentioned studies by (Yamada et al., 2019) and (Okumura et al., 2015).

Tanaka et al. (Tanaka et al., 2016) proposed a judgment method based on tweets on egogram diagnosis result using Naive Bayes classifier. Their method applied the word frequency vector as the feature to be learned and estimated the expression level of each state assuming that the five ego states of the egogram are independent.

Matsumoto et al. (Matsumoto et al., 2017) extracted the feature vector as a sentence unit using sentence2vec and constructed an egogram estimation model through a multi-layer neural network. They demonstrated a higher personality match rate compared to that using only sentence2vec by including word intimacy, which is an index that shows the familiarity of a word based on a real value ranging from 1-7. However, the contribution of this value to each personality pattern is not clear because word intimacy does not indicate the familiarity of the word to the user.

Minamikawa et al. (Minamikawa et al., 2010) proposed a personality estimation method based on blogs, using text mining. They used egograms for personality diagnosis, provided information gain scores to the Bag-of-Words features and performed personality estimation using the Naive Bayes classifier and SVM. The experimental results demonstrated, that the more the features used, the more accurate was the personality estimation.

The study by Guntuku et al. (Guntuku et al., 2017) aimed to determine how and to what extent personality differences, measured using the Big Five model, are related to online image posting and liking. The purpose of their study was to determine how and to what extent personality differences, as measured by the Big Five model, are related to online image posting and liking. In their study, they used large-scale image content analysis to extract interpretable semantic concepts from 1.5 million Twitter images posted by 4,000 users, and analyzed the differences in each personality trait. The prediction results show that image content can predict personality traits, and that the posted image and the ”like We showed that fusing the signals from both images provides a significant performance improvement.

Most of the studies on personality estimation based on text information use word-based features but few have attempted to improve the accuracy consider the relationship between the personality and sensibility. However, the research results of (Okumura et al., 2015) indicate that emotional information can be extracted from the remarks on blogs and SNSs, which not only express the feelings of the writer, but also serve as a clue to their personality. In addition, it is shown that social media image content is related to personality. In this study, we not only analyze words but also extract the emotional features from sentences composed of multiple words and confirm whether they are effective as features for personality estimation.

2.2 Emotion Estimation from Tweets

Several studies estimate emotions based on the text posted on Twitter. For example, Matsumoto et al. (Matsumoto et al., 2016) proposed a method for constructing an emotion estimation model for each attribute of a Twitter user. Although they considered the occupation and gender of the user, modelling by attribute may not be more effective than a generalized model that does not consider the attributes; moreover, the features need to be effective for emotion estimation regardless of the attributes.

Fujino et al. (Fujino et al., 2018) attempted to improve the accuracy of emotion estimation with a classification model using a neural network adapted to gender. A certain degree of accuracy improvement was observed but the effectiveness could not be confirmed.

Jamal et al. (Jamal et al., 2021) proposed an IoT-based framework for sentiment classification of tweets using S-LSTM model with improved Term Frequency Inverse Document Frequency (TFIDF) and Long-Short Term Memory (LSTM) after filtering the tweets. Furthermore, they use Adaptive Synthetic (ADASYN) to resolve class imbalance and improve the accuracy. The evaluation experiments proved that their method outperforms the state-of-the-art methods.

These studies demonstrate that the distribution of the vocabulary and labels in the target data is important for highly accurate emotion estimation from the posted text on Twitter, and that differences in the user attributes may not directly affect the accuracy of emotion estimation. In this study, the influence of vocabulary distribution is reduced by using a model with a distributed representation of sentences learned from a large-scale corpus as a feature and by adopting an oversampling method to deal with the imbalance in the emotion labels.

3 Proposed Method

3.1 Tweet Collection

The accounts of Twitter users who tweeted the personality diagnosis results with the hashtag “#16personalities” were identified using the Twitter API. 16personalities [1] is a web service that provides a personality diagnosis test service based on the MBTI, and the diagnosis results can be shared directly on Twitter. Figure 1 is an example of a tweet shared directly from 16Personalities on Twitter

Figure 1: Example of a Tweet shared from 16Personalities.

Further, we collected approximately 100-200 tweets from each user account on dates close to the shared personality diagnosis result tweet.

3.2 Feature Extraction from Tweets

All the collected tweets were preprocessed to remove noise, as follows:

- Remove URL
- Remove emoticon
- Remove hashtag
- Convert full-width alphabets and numbers to half-width characters
- Replace numbers with zero

Python’s neologdn package[2] for text normalization was used, which normalizes Japanese text, and easily and rapidly performs a wide variety of text normalization.

Further, MeCab, a morphological analysis engine, was applied to divide all the tweet sentences

[1] https://www.16personalities.com/

[2] https://github.com/ikegami-yukino/neologdn

denoised word-by-word. mecab-ipadic-neologd (Sato., 2015) was used as the MeCab dictionary. This dictionary includes new words and named entities, and is suitable for the morphological analysis of tweet sentences because unique expressions and new words appear at a high rate in the tweet on Twitter. We extracted a total of four features: Word2Vec, Sentence BERT, the vector of the intermediate layer of the neural network that outputs the emotional information vector, and the emotional information vector from the processed data. The method for extracting each vector is presented below.

3.3 Word2Vec

We provided distributed representation to the divided tweet text using the learned Word2Vec model (Hotlink, Inc.). This model trains CBOW using the Japanese Wikipedia as the training data and has 100 dimensions. The distributed representation of each split word of the tweet sentence was given the distributed representation of the model. The sum of each distributed representation was then divided by the number of appear words, and used as the distributed representation of the tweet. Equation (1) was applied for calculating the distributed representation of the tweet using Word2Vec.

$$Vec_t = \frac{1}{n}\sum_{i=1}^{n} wv_i \tag{1}$$

3.4 Sentence BERT

We extracted the vector from the tweet text using the learned Japanese-Sentence BERT model (NS Solutions Corporation). This model was created for Japanese text based on the Sentence-BERT (Nils and Iryna., 2019). The Japanese BERT model developed by Tohoku University, Japan[3] was used. The data set used for training is not open to the public.

3.5 Emotion Vector

We built an emotion estimation model using a neural network based on a corpus with emotion tags to obtain the emotion vectors as features. The corpus was manually selected and constructed using an extended corpus by Fujino et al. (Fujino et al., 2019), based on tweets to which five emotional labels were manually attached. The emotional label is based on the emotional system diagram proposed by (Fischer et al., 1989), which includes “ joy ”, “ surprise ”, “ anger ”, and “ sorrow ”. If it doesn’t apply to any label, it is set to “neutral”. This model improved the accuracy of the test data by approximately 80%, due to oversampling using SMOTE (Chawla et al., 2002) during training because of the label imbalance in the corpus. The emotion vector represents the type of emotion categories (5 types) estimated from the input sentence and is output as a 5-dimensional real-value vector.

In addition, the 544-dimensional vector extracted from the hidden layer of the trained network described above was also used as a feature. The hidden layers in neural networks can often function as meaningful features. Using these two features as the emotional information, we compared the average of Word2Vec’s distributed representation, the Sentence-BERT vector, and the combined vector. Furthermore, we experimentally confirmed whether the emotional information was useful for constructing a personality estimation model. The average of the distributed representation of a 100-dimensional Word2Vec and the vector of a 768-dimensional Sentence BERT was input to the emotion estimation model, which finally, output five emotion labels (joy, surprise, anger, sadness, neutral) using the SoftMax function. Figure 2 depicts the architecture of the emotion estimation model. The vector were extracted from the layer which is surrounded by the red line of the network.

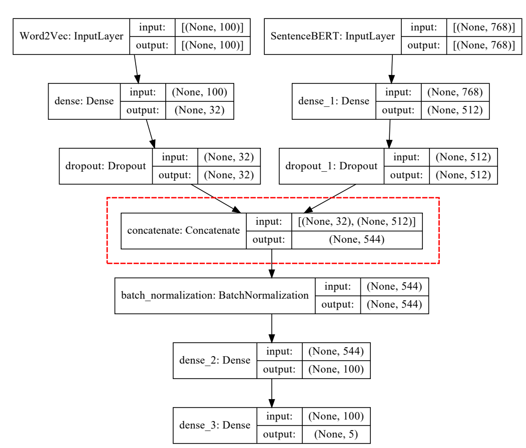

Figure 2: Architecture of the emotion estimation model.

[3]https://github.com/cl-tohoku/bert-japanese

3.6 Construction of the Personality Estimation Model

We built a classification model with LightGBM (Guolin et al., 2017), which takes the average features extracted from the user tweets as the input and outputs the personality types. LightGBM is a gradient boosting machine learning framework based on the decision tree algorithm. In gradient boosting, the calculation cost increases with the increase in the number of records and features of the training data; however, LightGBM uses gradient-based one-side sampling (GOSS) and exclusive feature bundling (EFB) to reduce the number of records and features, and efficient training is performed.

4 Evaluation Experiment

We conducted an experiment to evaluate the effectiveness of personality estimation through the proposed method using the collected data. We had identified approximately 8,000 accounts in which the MBTI personality diagnosis results had been tweeted. Figure 3 displays a graph of the number of collected users with respect to the MBTI personality type. It can be observed that the number of INFP type users is extremely high and the number of users belonging to the extraversion (E) type is smaller than that of the introversion (I). Approximately 100-200 tweets were collected from these accounts, starting from the tweets on the personality diagnosis results, and a total of approximately 780,000 tweets were obtained, and contain the MBTI diagnosis result tweet.

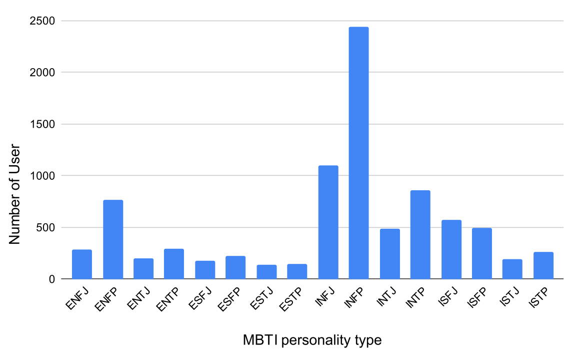

Figure 3: Number of users with respect to the MBTI personality type.

We divided the collected user accounts into training and test user accounts at a ratio of 8:2. Table 2 lists the number of users and the number of collected training and test data tweets.

	Users	Tweets
Training Data	6,463	856,541
Test Data	1,623	217,118
Total	8,086	1,017,118

Table 2: Breakdown of the training and test data.

With these data, LightGBM was used to train a model that binary classified each of the four indicators of the MBTI personality type into two types. As LightGBM includes many hyper parameters, we optimized them using Optuna (Akiba et al., 2019), which is an automatic hyper parameter optimization framework designed for machine learning. The accuracy of the personality estimation model was compared for four patterns: the mean of the distributed representation of Word2Vec (Word2Vec), the mean of the vectors of Sentence BERT (Sentence BERT), Word2Vec + Sentence BERT, and All features (Word2Vec + Sentence BERT + Emotional Vector).

In addition, we used the ROC curve AUC as the evaluation index. The ROC curve AUC is a performance measurement for the classification model. The ROC curve is plotted with the false positive rate (FPR) on the x-axis and the true positive rate (TPR) on the y-axis. The ROC curve is plotted with a false positive rate (FPR) on the x-axis and a true positive rate (TPR) on the y-axis, and the higher the area under the curve (AUC), the more accurate the model. The TPR and FPR can be calculated for equations 2 and 3.

$$TPR = \frac{TP}{TP + FN} \tag{2}$$

$$FPR = \frac{FP}{TN + FP} \tag{3}$$

5 Results and Discussion

Table 3 shows the AUC of the ROC curve of the binary classification personality estimation model, which classifies each of the four indicators of the MBTI personality type into two types. For

feature	W2V	SBERT	W2V + SBERT	ALL (Contain Emotion)
IE	0.647	0.645	0.661	**0.734**
NS	0.666	0.669	0.681	**0.738**
TF	0.643	0.659	0.666	**0.722**
JP	0.571	0.604	0.606	**0.694**
Average	0.632	0.644	0.654	**0.722**

Table 3: AUC of the ROC curve of the binary classification personality estimation model.

Word2Vec and Sentence BERT, the AUC is almost the same; on combining Word2Vec and Sentence BERT, the AUC value is slightly improved in comparison. Compared to (Word2Vec + Sentence BERT + Emotional vector) and Word2Vec + Sentence BERT, the AUC for All is higher for all the MBTI indicators. The IE, NS, and TF indicators show improvement in the AUC by approximately 5-7%, and whereas the JP indicator show significant improvement in the AUC by approximately 10%. The average AUC of each indicator shows an improvement of approximately 7%.

The emotion vector does not improve the AUC using only the five-dimension vector of the emotion estimation model but improve the AUC using a 544-dimension vector of the hidden layer. The higher the number of feature dimensions, the better the AUC of the personality estimation model.

The JP indicator reflects the processing in the outside world, where judging (J) people prefer the judgment process (thinking or emotion) when dealing with the outside world and perceiving (P) people prefer the cognitive process (sense or intuition). Moreover, there is a relationship between the personality and emotional information because the AUC is improved not only for the JP indicator but also the other indicators. In addition, the personality and sensibilities such as joy and anger are likely to appear frequently in the text because it is possible to easily tweet intuitive personal impressions on several factors such as one's physical condition, social situation, and the weather.

5.1 Data Bias

The collected MBTI data included numerous INFP personality types and the data was highly biased. The data collected in the study by Yamada et al. (Yamada et al., 2019) was also biased and the number of INFP personality types was large as in this study. Table 4 depicts the percentage of users for each MBTI indicator in our study and the study by Yamada. Yamada's study mentions that the data distribution has negligible effect because the methods are compared when evaluating the ROC curve through the AUC as well. Moreover, it is considered that the impact is relatively less because our study uses the same evaluation indicator as that of Yamada.

Data Set	I-E	N-S	T-F	J-P
Yamada's	78.0-22.0	67.4-32.6	46.9-53.1	34.4-65.6
This Research	73.0-27.0	72.9-27.1	31.0-69.0	37.7-62.3

Table 4: Percentage of MBTI users with respect to each indicator.

5.2 Comparison with a Latest Research

We compare this study with the latest research on personality estimation. Amirhosseini et al. (Amirhosseini et al., 2020) conducted a similar study of personality estimation based on the MBTI. They use the MBTI dataset from Kaggle and use XGBoost (Chen et al., 2016) for training a personality estimation model. They also use TF-IDF as a feature. The accuracy of this model is 78.17% for IE, 86.06% for NS, 71.78% for TF and 65.70% for JP. This model IE indicators approximately 5% and NS approximately 13% higher than our model. But TF indicators are approximately 0.5% and NS approximately 7% less than our model. This indicates that emotional information contributes to the improvement of the accuracy of NS and TF indicators.

6 Conclusion

In this study, we proposed a method to estimate the MBTI personality pattern by extracting emotional information from Twitter and adding it to the characteristics of the distributed expressions of sentences and words. Evaluation experiments were performed to determine whether the emotional information was effective in constructing the MBTI personality estimation model, through the proposed method. The obtained results confirmed an improvement in the accuracy of the personality estimation model when emotional information was included as a feature, compared to the case without emotional information.

In the future, we intend to undertake the following. First, we intend to increase the data for training the personality estimation model in order to enhance model generalization. Next, we intend to build a personality estimation model using other methods, such as neural networks. The model trained by other methods will be compared to the model trained by LightGBM to build a more accurate personality estimation model. In addition, we intend to further investigate the type of emotional information that relates to the personality because the correlation between the personality and emotional information has not been fully investigated.

Finally, I want to create a visualization tool for understanding personality. Gou et al. (Gou et al., 2013) developed PersonalityViz, an interactive visualization tool for understanding one's personality traits derived from social media. The system uses the Linguistic Inquiry and Word Count (LIWC) text analysis tool and LIWC/Big Five personality correlation to calculate a person's Big Five personality from his or her tweets. The system provides an interactive visual interface that allows users to explore their personality traits in chronological order, and by examining the visual evidence, it is possible to understand how personality traits are derived from related tweets. In addition to their functionality, my visualization tool intends to add the ability to recommend compatible users.

Acknowledgments

This work was supported by the 2021 SCAT Research Grant and JSPS KAKENHI Grant Number JP20K12027, JP21K12141.

References

Bakry, M. R., Nasr, M. M. and Al-sheref, F. K. 2019. *A Survey of Psychological Personality Classification Approaches*, Future Computing and Informatics Journal: Vol. 4 : Iss. 2 , Article 3.

Yamada, K., Sasano, Y. and Takeda, K 2019. *Incorporating Textual Information on User Behavior for Personality Prediction*, In Proc. of the 57th Annual Meeting of the Association for Computational Linguistics: Student Research Workshop, pp.177–182.

Okumura, N., Kanamaru, Y. and Okumura, M. 2015. *Personality Estimation of Blog Authors based on Emotion Judgment and Big Five (in Japanese)*, The 29th Annual Conference of the Japanese Society for Artificial Intelligence, pp.4J15-4J15.

Tsuchiya, S., Yoshimura, E., Ren, F., Watabe, H. 2009. *Emotion Judgment Based on Relationship between Speaker and Sentential Actor*, In International Conference on Knowledge-Based and Intelligent Information and Engineering Systems, pp.62–69.

Fischer, K. W., Shaver, P. and Carnochan, P. 1989. *A skill approach to emotional development: From basic-to subordinate-category emotions*, In W. Damon (Ed.), Child development today and tomorrow, pp.107—136.

Tanaka, S., Matsumoto, K., Yoshida, M. and Kita, K. 2016. *Personality Analysis Method of Twitter User for Information Recommendation (in Japanese)*, The 30th Annual Conference of the Japanese Society for Artificial Intelligence, pp.3B4NFC05b3–3B4NFC05b3.

Matsumoto, K., Tanaka, S., Yoshida, M., Kita, K. and Ren, F. 2017. *Ego-state Estimation from Short Texts Based on Sentence Distributed Representation*, International Journal of Advanced Intelligence (IJAI), pp.145–161.

Minamikawa, A. and Yokoyama, H. 2010. *Personality Estimation from Weblog Data Using Text Mining (in Japanese)*, JSAI SIG-DMSM, SIGDMSM-A903-15.

Guntuku, S. C., Lin, W., Carpenter, J., Ng, W. K., Ungar, L.H., Preoţiuc-Pietro, D. 2017. *Studying Personality through the Content of Posted and Liked Images on Twitter*, In Proc. of the 2017 ACM on Web Science Conference, pp.223–227.

Matsumoto, K., Yoshida, M., Kita, K., Wu, Y. and Iwasa, F. 2016. *Effect of user's attribute on emotion estimation from Twitter*, In 2016 2nd IEEE International Conference on Computer and Communications (ICCC), pp.1186–1190.

Fujino, N., Matsumoto, K., Yoshida, M. and Kita, K. 2018. *Emotion Estimation Adapted to Gender of User Based on Deep Neural Networks*, International Journal of Advanced Intelligence, pp.121–133.

Jamal, N., Xianqiao, C., AL-turjman, F. and Ullah, F 2021. *A Deep Learning–based Approach for Emotions Classification in Big Corpus of Imbalanced Tweets*, Transactions on Asian and Low-Resource Language Information Processing, Volume 20, Issue 3, pp.1—16.

Matsuno, S., Mizuki, S. and Sakaki, T. 2019. *Constructing of the word embedding model by Japanese large* *scale SNS + Web corpus (in Japanese)*, The 33rd Annual Conference of the Japanese Society for Artificial Intelligence, pp.4Rin113–4Rin113.

Chawla, N. V., Bowyer, K. W., Hall, L. O., and Kegelmeyer, W. P. 2002. *SMOTE: Synthetic Minority Over-sampling Technique*, Journal of artificial intelligence research, pp.321–357.

Ke, G, Meng, Q., Finely, T., Wang, T., Chen, W., Ma, W., Ye, Q., Liu, TY. 2017. *LightGBM: A Highly Efficient Gradient Boosting Decision Tree*, Advances in neural information processing systems, 30, 3146-3154.

Sato, T. 2015. *Neologism dictionary based on the language resources on the Web for Mecab*, `https://github.com/neologd/mecab-ipadic-neologd`

Fujino, N., Matsumoto, K., Yoshida, M. and Kita, K. 2019. *[Examination of improving emotion estimation accuracy by corpus expansion using Word Mover's Distance] Word Mover's Distance wo mochiita copasu kakutyo ni yoru kanjo suitei seido kojo no kento (in Japanese)*, Proceedings of the Twenty-fifth Annual Meeting of the Association for Natural Language Processing, pp.129–131.

Reimers, N. and Gurevych, I. 2019. *Sentence-BERT: Sentence Embeddings using Siamese BERT-Networks*, arXiv preprint arXiv:1908.10084.

Akiba, T., Sano, S., Yanase, T., Ohta, T. and Koyama, M. 2019. *Optuna: A Next-generation Hyperparameter Optimization Framework*, In Proc. of the 25th ACM SIGKDD international conference on knowledge discovery & data mining, pp.2623–2631.

Amirhosseini, M. H. and Kazemian, H. 2020. *Machine Learning Approach to Personality Type Prediction Based on the Myers-Briggs Type Indicator*, Multimodal Technologies and Interaction.

Chen, T. and Guestrin, C. 2016. *XGBoost: A Scalable Tree Boosting System*, In Proc. of the 22nd ACM SIGKDD International Conference on Knowledge Discovery and Data Mining, pp.785—794.

Gou, L., Mahmud, J., Haber, E., Zhou, M. 2013. *PersonalityViz: a visualization tool to analyze people's personality with social media*, In Proc. of the companion publication of the 2013 international conference on Intelligent user interfaces companion, pp.45–46.