MBTI Personality Prediction Approach on Persian Twitter

Samin Fatehi[*], **Zahra Anvarian**[*], **Yasmin Madani**,
MohammadJavad Mehditabar, **Sauleh Eetemadi**
School of Computer Engineering, Iran University of Science and Technology, Iran
{sa_fatehi, yasmin_madani, m_mehditabar}@comp.iust.ac.ir,
zahra.anvarian97@gmail.com, sauleh@iust.ac.ir

Abstract

Automatic personality detection has many applications in different fields. In this paper, since there is no substantial study on Persian speakers' personality detection, we present an efficient and high-performance personality detection system in the Persian language. We collected text data from 3781 Twitter users, embedded them using fasttext, and fed them to a neural network that comprises LSTM and attention layers. Finally, our model demonstrated significant improvement in at least two traits. Also, we used LIME to interpret how a derivative of our model predicts individual personality.

1 Introduction

Personality traits are stable patterns of thoughts, feelings, and behaviors. Personality detection has many applications such as recommendation systems (Yin et al., 2018), job screening (Liem et al., 2018), etc.

In recent studies, transformers and feature combinations have improved automatic personality detection systems. Mehta et al. Combined BERT and psycholinguistic features for each Essay and Personality Cafe dataset, and fed them to MLP and SVM models. In another study, Christian et al. worked on MyPersonality and a dataset collected from Twitter in Bahasa Indonesia. They used a combination of pre-trained word embeddings and statistic features for personality detection. Although there is one collected Persian dataset in this field and the ParsBERT model was tuned on it (Abdollahpour et al., 2021), the model's result should be improved, which we accomplished in this paper.

In this paper, since there is no significant dataset in Persian, we collected a dataset from Twitter users on Myers Briggs Type Indicator (MBTI), which has four traits: Introvert–Extrovert, Intuitive–Sensing, Thinking–Feeling, and Judging–Perceiving (Martin, 1997). Then we developed a model base on fasttext embedding (Bojanowski et al., 2017), attention (Vaswani et al., 2017), and LSTM (Hochreiter and Schmidhuber, 1997) layers for predicting personality. Finally, we used LIME to interpret our model.

2 Methodology

In this section, We present our method to accumulate a significant and valid dataset in Persian and develop a model for personality detection.

2.1 Data Collection

Due to the lack of Persian datasets in this field, we chose Twitter and have exploited three different methods of collecting data. Ethical boundaries and users' privacy were considered at every process.

Bio Search: This part contains Iranian users with specified MBTI types in their bio. This dataset consists of users' publicly available tweets and MBTI labels which were attained by searching 16 labels in them.

Tweet Search: Through this approach, we have searched for 16 possible MBTI labels in users' recent tweets in which one of these types is mentioned. Every tweet was labeled manually whether the mentioned type was about the author user, other users, or just an opinion about personality types.

Questionnaire: We used a questionnaire to obtain more accurate and so-called golden data. We received a few more data from users as well as their gender, location, degree of education, and age for extra analysis. Also, we filtered users to have at least 150 tweets. Our questionnaire was ethical since all users

[*] These authors contributed equally.

completed it voluntarily; we explained utterly in the questionnaire that we would not utilize any users' private data, like usernames, and all users also were verified.
For data cleaning on all data (without retweets), firstly, we replaced URLs with [LINK], usernames with [USERNAME], emojis with [EMOJI], and smileys with [SMILEY]. We also filtered out non-Persian characters. Finally, we only preserve users with more than 100 tweets after performing the above-mentioned cleanings.

2.2 Proposed Model

We used a bidirectional LSTM as our model and also used the attention layer above it to weight effective tweets for determining labels more accurately. We used the fasttext, as our embedding model and created a vector for each user as $U^{700 \times 300}$. Each tweet's vector is produced by the mean of its words' vectors. This model is trained for each trait once; hence we have four classifications overall.

3 Experiment

In this section, we analyze the collected dataset and then discuss the results of our model.

3.1 Data Analysis

The precise statistics of the dataset are shown in Table 1.

Table 1: Statistics of data collection methods

	Bio-Tweets	Questionnaire	Total
# Tweets	5,978,246	161,626	6,139,872
# Users	3,781	95	3,876

Also, the distribution of 16 types in Twitter data implies introverts use Twitter more. It would be because extroverts prefer face-to-face communication, whereas social media and computer-mediated communication are ideal for introverts (Acar and Polonsky, 2007).

3.2 Experimental Results

To feed the data into a model, we used a stratified K-fold with k = 4 to decrease the impact of the dataset's unbalancedness as the majority baseline is shown in the table 2. As table 2 shows, our model performs 13 percent better than the majority baseline in P/J trait, and we can also see the improvement of our model in the T/F trait by about 7 percent. As shown in the table, our model does not learn how to recognize whether a person is intuitive or sensing in the N/S personality trait. In the previous study, Plank and Hovy tried to predict MBTI personality traits from Twitter texts, but their models could not achieve any progress on the N/S trait.

Table 2: Accuracy of the proposed method

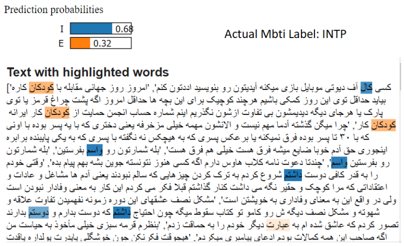
Model	I/E	N/S	T/F	P/J	Avg
Majority Baseline	66.61%	74.72%	59.22%	51.89%	63.11%
Proposed Model	66.6%	69.16%	66.86%	64.22%	66.71%

3.3 Visualization and Interpretability

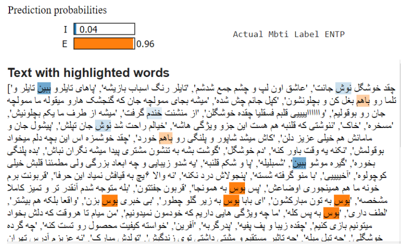
Finally, we used LIME (Ribeiro et al., 2016) to interpret the model with the structure of our main model, but with just an embedding layer instead of fasttext. The results will have a slight drop in accuracy, but they can give us good insight.
Figure 1b indicates that the model found it is important that an extroverted person has used words such as "kiss" (بوسه) or "together" (باهم), which seems rational as extroverted people are more show their feelings and interested in group works. In Figure 1a by contrast, according to models' anticipations, an introverted user with the "INTP" MBTI type uses more words with personal pronouns. Most verbs end in (م), which means "I" or "my," and it shows our introverted sample had focused on himself rather than other people.

4 Conclusion

This paper presents an automated personality detection system for Persian speakers. Since there is no Persian personality dataset, we collected data by three methods from Twitter: searching keywords, Bios, and questionnaires, and we have 3876 users who declared their MBTI personality types (in their tweets or bios) and 95 users who completed the questionnaire. After extracting features from text using fasttext, we fed vectors to a neural network made up of LSTM, attention, and dense layers.



(a) INTP



(b) ENTP

Figure 1: Sample highlighted data using LIME

References

Mohammad Mahdi Abdollahpour, Zahra Anvarian, Samin Fatehi, and Sauleh Eetemadi. 2021. Parstset: A persian dataset for personality prediction on twitter. In *The Fourth Annual West Coast NLP (WeCNLP) Summit*.

Adam S Acar and Maxim Polonsky. 2007. Online social networks and insights into marketing communications. *Journal of Internet Commerce*, 6(4):55–72.

Piotr Bojanowski, Edouard Grave, Armand Joulin, and Tomas Mikolov. 2017. Enriching word vectors with subword information. *Transactions of the association for computational linguistics*, 5:135–146.

Hans Christian, Derwin Suhartono, Andry Chowanda, and Kamal Z Zamli. 2021. Text based personality prediction from multiple social media data sources using pre-trained language model and model averaging. *Journal of Big Data*, 8(1):1–20.

Sepp Hochreiter and Jürgen Schmidhuber. 1997. Long short-term memory. *Neural computation*, 9(8):1735–1780.

Cynthia CS Liem, Markus Langer, Andrew Demetriou, Annemarie MF Hiemstra, Achmadnoer Sukma Wicaksana, Marise Ph Born, and Cornelius J König. 2018. Psychology meets machine learning: Interdisciplinary perspectives on algorithmic job candidate screening. In *Explainable and Interpretable Models in Computer Vision and Machine Learning*, pages 197–253. Springer.

Charles R Martin. 1997. *Looking at type: The fundamentals*. Center for Applications of Psychological Type.

Yash Mehta, Samin Fatehi, Amirmohammad Kazameini, Clemens Stachl, Erik Cambria, and Sauleh Eetemadi. 2020. Bottom-up and top-down: Predicting personality with psycholinguistic and language model features. In *2020 IEEE International Conference on Data Mining (ICDM)*, pages 1184–1189. IEEE.

Barbara Plank and Dirk Hovy. 2015. Personality traits on twitter—or—how to get 1,500 personality tests in a week. In *Proceedings of the 6th workshop on computational approaches to subjectivity, sentiment and social media analysis*, pages 92–98.

Marco Túlio Ribeiro, Sameer Singh, and Carlos Guestrin. 2016. "why should I trust you?": Explaining the predictions of any classifier. *CoRR*, abs/1602.04938.

Ashish Vaswani, Noam Shazeer, Niki Parmar, Jakob Uszkoreit, Llion Jones, Aidan N Gomez, Łukasz Kaiser, and Illia Polosukhin. 2017. Attention is all you need. *Advances in neural information processing systems*, 30.

Han Yin, Yue Wang, Qian Li, Wei Xu, Ying Yu, and Tao Zhang. 2018. A network-enhanced prediction method for automobile purchase classification using deep learning. In *PACIS*, page 111.