Personality Trait Detection
Using Bagged SVM over BERT Word Embedding Ensembles

Amirmohammad Kazameini[1], Samin Fatehi[1], Yash Mehta[2], Sauleh Eetemadi[1], Erik Cambria[3]

[1]School of Computer Engineering, Iran University of Science and Technology, Iran
[2]Gatsby Computational Neuroscience Unit, University College London, UK
[3]School of Computer Science and Engineering, Nanyang Technological University, Singapore
{a_kazemeini, sa_fatehi}@comp.iust.ac.ir, y.mehta@ucl.ac.uk,
sauleh@iust.ac.ir, cambria@ntu.edu.sg

Abstract

Recently, the automatic prediction of personality traits has received increasing attention and has emerged as a hot topic within the field of affective computing. In this work, we present a novel deep learning-based approach for automated personality detection from text. We leverage state of the art advances in natural language understanding, namely the BERT language model to extract contextualized word embeddings from textual data for automated author personality detection. Our primary goal is to develop a computationally efficient, high performance personality prediction model which can be easily used by a large number of people without access to huge computation resources. Our extensive experiments with this ideology in mind, led us to develop a novel model which feeds contextualized embeddings along with psycholinguistic features to a Bagged-SVM classifier for personality trait prediction. Our model outperforms the previous state of the art by 1.04% and, at the same time is significantly more computationally efficient to train. We report our results on the famous gold standard Essays dataset for personality detection.

1 Introduction and Related Work

An individual's personality has a great impact on their lives, affecting their life choices, well-being, health and even preferences and desires. Hence, there is a huge interest to develop models which can automatically identify an individual's personality with important practical applications such as in recommendation systems (Yin et al., 2018), job screening (Liem et al., 2018), social network analysis (Maria Balmaceda et al., 2014), etc. Our model makes binary predictions of the author's personality based on the famous Big-Five (Digman, 1990) personality measure, which are the following five traits: Extraversion (EXT), Neuroticism (NEU), Agreeableness (AGR), Conscientiousness (CON) and Openness (OPN).

Common author personality detection techniques usually involve extracting psycholinguistic features from text, such as Linguistic Inquiry and Word Count (LIWC) (Pennebaker et al., 2001), Mairesse features (Mairesse et al., 2007), and SenticNet (Cambria et al., 2018), which are then fed into traditional machine learning classifiers such as support vector machine (SVM) (Hearst et al., 1998), Naïve Bayes, etc. More recent work leverage deep learning and make use of pre-trained word embeddings like Word2Vec (Mikolov et al., 2013) and Glove (Pennington et al., 2014). Recently, (Mehta et al., 2020) reviewed the latest advances in deep learning-based automated personality detection from the viewpoint of different input modalities along with recent techniques for effective multimodal personality prediction.

The previous state of the art (Majumder et al., 2017) on the Essays dataset also make use of a deep learning based approach with a convolutional network on top of word embeddings extracted from Word2Vec. They also incorporate other inputs such as the Mairesse features, word count, average sentence length, etc. for their final prediction. Their model outperformed the previous best (Mohammad and Kiritchenko, 2015) by 0.55%, whereas our model outperforms (Majumder et al., 2017) by 1.04% and at the same time being significantly more computationally efficient to train.
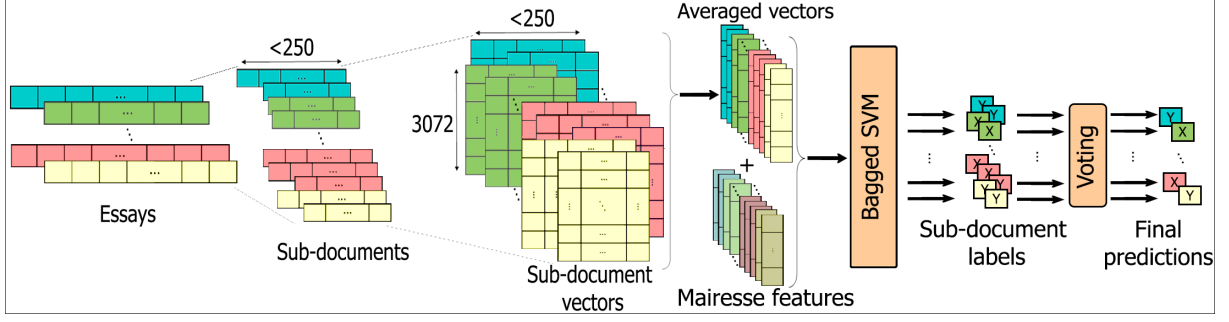
Figure 1: An overview of our deep learning-based Bagged-SVM model for automated personality detection

2 Proposed Method

Each of our inputs is an essay with a mean size of around 650 words. The maximum number of tokens BERT can process at a time is 512. Hence, to extract maximum information from the input, we break the essay into multiple chunks (sub-documents), with the maximum length of a chunk being 200 tokens. All these sub-documents of a particular essay are then annotated with the same personality label as that essay. We experimented with various methods of pre-processing the essays text before feeding it to the BERT tokenizer and use the best performing one.

We split the text into a sequence of sentences at the period and question mark characters and remove all characters other than ASCII letters, digits, quotations and exclamation marks. We expand all shortenings (e.g., "you're" becomes "you are") which increases the maximum length of a sub-document from 200 to 250 tokens. After this initial pre-processing step, the sub-documents are then fed into the pre-trained $BERT_{BASE}$ model. For each layer of BERT, we average the contextual token representations of that layer. Then, we concatenate the last four layer representations and concatenate this with the corresponding 84 Mairesse features for the essay. This is then considered as the feature vector for the document, which is of the dimension $\mathbb{R}^{3156}$.

In the classification phase, we feed the document feature vector to a SVM which predicts a binary label corresponding to a particular personality trait. To enhance the performance further, we use ten SVM classifiers to perform the prediction in parallel like the bagging classifier (Breiman, 1996). The estimator trains on all the features on the total number of the sub-document stack with replacement and the final predicted model for a document is done by majority voting.

3 Evaluation

Our model achieves a 1.04% increase in performance in comparison to the previous state of the art along with being significantly more computationally efficient. To put it in perspective, (Majumder et al., 2017) train their model on an Intel Core i7-4720 HQ CPU with the optimal configuration and it takes about 50 hours to complete. Our fine-tuning model only takes about 7 minutes to train. Table 1 gives a comparison of our model, BB-SVM, with others. We modify various parts of the BB-SVM model and discuss their effect on performance in the following section.

	Personality Traits					
Model Name	EXT	NEU	AGR	CON	OPN	Average
Majority Baseline	51.72	50.2	53.10	50.79	51.52	51.43
Mairesse	55.13	58.9	55.35	55.28	59.57	56.84
Previous state of the art	58.09	57.33	**56.71**	56.71	61.13	57.99
BB-SVM	**59.30**	**59.39**	56.52	**57.84**	**62.09**	**59.03**†

Table 1: A comparison of the performance of our model (BB-SVM) with others (†: Statistically significant at p ≤ 0.05)

Model Id	Word Embedding	Sentence Feature Extraction	Document Feature Extraction	Classifier	Average Accuracy
M8	W2V	-	Mean	Bagging-SVM	57.38
BB-SVM	BERT (4 last layers)	-	Mean	Bagging-SVM	59.03

Table 2: A comparison of average accuracy of all the 5 traits with different word embeddings

3.0.1 Word Embedding

We compare the model performance with context-independent word embeddings such as Word2Vec. Table 2 shows a comparison of the results. Also, results reported by (Devlin et al., 2018) suggest that concatenating the last four layers of BERT gives the best representation for a word. The comparison of BB-SVM results with inputs of different BERT layers is shown in Figure 2.
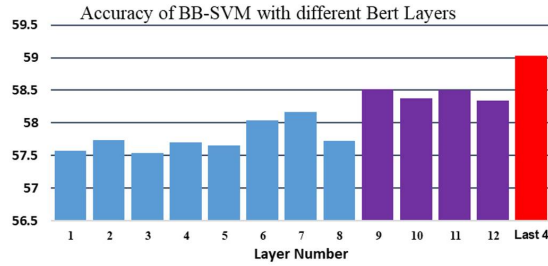

Figure 2: Accuracy of BB-SVM model with different BERT layer inputs compared to the accuracy of this model with the concatenation of the last four layers

Model Id	Word Embedding	Sentence Feature Extraction	Document Feature Extraction	Classifier	Average Accuracy
DocBERT	BERT	-	-	MLP	57.11
M11	BERT (layer 11)	Mean	CNN+Max	MLP	57.42
M12	BERT (layer 11)	Mean	CNN+GRU	MLP	57.42
M3	BERT (layer 11)	-	Mean	SVM	58.49
M14	BERT (layer 11)	-	Mean	Bagging-SVM	58.51

Table 3: A comparison of average accuracy of all the 5 traits with different base classifiers

3.0.2 Fine-tuning Network and Feature Extraction

In the classification phase, we experimented with a SVM and a multi-layer perceptron for making the final personality trait predictions. We found that using a SVM results in better performance. We also experiment with feeding sub-document features to DocBERT (Adhikari et al., 2019), followed by averaging sub-document predictions to obtain the document's prediction. However, this did not improve the results. Table 3 shows a comparison of the results. We train the model by applying Bagging (using ten simultaneous SVM classifiers) and in line with previous studies (Kim et al., 2002), Bagging improved the classification accuracy for the task of personality detection as well (table 4).

We also tried 2 different ways to extract the document features. In the traditional approach, the document features are directly constructed from word features. A different approach is to first construct sentence features using the word features and then construct the document features from these sentence features. Table 5 shows a comparison of the results.

Model Id	Word Embedding	Sentence Feature Extraction	Document Feature Extraction	Classifier	Average Accuracy
M13	BERT (4 last layers)	-	Mean	SVM	58.76
BB-SVM	BERT (4 last layers)	-	Mean	Bagging-SVM	59.03

Table 4: A comparison of average accuracy of all the 5 traits with and without applying bagging

Model Id	Word Embedding	Sentence Feature Extraction	Document Feature Extraction	Classifier	Average Accuracy
M9	BERT (4 last layers)	Mean	Mean	Bagging-SVM	57.91
BB-SVM	BERT (4 last layers)	-	Mean	Bagging-SVM	59.03

Table 5: A comparison of average accuracy of all the 5 traits with different features extraction methods

4 Conclusion and Future Work

In this paper, we presented a computationally efficient deep learning-based model which outperformed the state of the art on the famous stream of consciousness Essays dataset. We hope that our model can be useful for research teams which do not have access to large computational resources. We believe a promising direction of future research would be to make more interpretable deep learning models which can provide valuable insights into the main psychological features driving these predictions and in turn also help advance psychological studies. Currently, the availability of quality personality datasets is quite limited. If an individual's personality can predicted with a little more reliability, there is scope for integrating automated personality detection in almost all human-machine interaction agents such as voice assistants, robots, cars, etc.

References

Ashutosh Adhikari, Achyudh Ram, Raphael Tang, and Jimmy Lin. 2019. Docbert: Bert for document classification. *arXiv preprint arXiv:1904.08398*.

Leo Breiman. 1996. Bagging predictors. *Machine learning*, 24(2):123–140.

Erik Cambria, Soujanya Poria, Devamanyu Hazarika, and Kenneth Kwok. 2018. SenticNet 5: Discovering conceptual primitives for sentiment analysis by means of context embeddings. In *AAAI*, pages 1795–1802.

Jacob Devlin, Ming-Wei Chang, Kenton Lee, and Kristina Toutanova. 2018. Bert: Pre-training of deep bidirectional transformers for language understanding. *arXiv preprint arXiv:1810.04805*.

John M Digman. 1990. Personality structure: Emergence of the five-factor model. *Annual review of psychology*, 41(1):417–440.

Marti A. Hearst, Susan T Dumais, Edgar Osuna, John Platt, and Bernhard Scholkopf. 1998. Support vector machines. *IEEE Intelligent Systems and their applications*, 13(4):18–28.

Hyun-Chul Kim, Shaoning Pang, Hong-Mo Je, Daijin Kim, and Sung-Yang Bang. 2002. Support vector machine ensemble with bagging. In *International Workshop on Support Vector Machines*, pages 397–408. Springer.

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.

François Mairesse, Marilyn A Walker, Matthias R Mehl, and Roger K Moore. 2007. Using linguistic cues for the automatic recognition of personality in conversation and text. *Journal of artificial intelligence research*, 30:457–500.

Navonil Majumder, Soujanya Poria, Alexander Gelbukh, and Erik Cambria. 2017. Deep learning-based document modeling for personality detection from text. *IEEE Intelligent Systems*, 32(2):74–79.

Jose Maria Balmaceda, Silvia Schiaffino, and Daniela Godoy. 2014. How do personality traits affect communication among users in online social networks? *Online Information Review*, 38(1):136–153.

Yash Mehta, Navonil Majumder, Alexander Gelbukh, and Erik Cambria. 2020. Recent trends in deep learning based personality detection. *Artificial Intelligence Review*, 53:2313–2339.

Tomas Mikolov, Kai Chen, Greg Corrado, and Jeffrey Dean. 2013. Efficient estimation of word representations in vector space. *arXiv preprint arXiv:1301.3781*.

Saif M Mohammad and Svetlana Kiritchenko. 2015. Using hashtags to capture fine emotion categories from tweets. *Computational Intelligence*, 31(2):301–326.

James W Pennebaker, Martha E Francis, and Roger J Booth. 2001. Linguistic inquiry and word count: Liwc 2001. *Mahway: Lawrence Erlbaum Associates*, 71(2001):2001.

Jeffrey Pennington, Richard Socher, and Christopher Manning. 2014. Glove: Global vectors for word representation. In *Proceedings of the 2014 conference on empirical methods in natural language processing (EMNLP)*, pages 1532–1543.

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.