

BERT of all trades, master of some

Author1, Author2, Author3

Affiliation1, Affiliation2, Affiliation3

Address1, Address2, Address3

author1@xxx.yy, author2@zzz.edu, author3@hhh.com

{author1, author5, author9}@abc.org

Abstract

This paper describes our results for TRAC 2020 competition held together with the conference LREC 2020. Competition consisted of 2 subtasks where participants had to classify aggression in texts and decide if it is gendered or not for 3 languages: Benghali, English and Hindi. We used a single BERT-based system with two outputs for all tasks simultaneously. Our model took the first place in English gendered text classification and the second place in Bengali gendered text classification.

Keywords: aggression, classification, BERT, neural network, Transformer, NLP

1. Introduction

This paper is devoted to our system’s solution for TRAC 2020 competition held together LREC 2020 conference. TRAC competition consisted of 2 sub-tasks in 3 languages: Bengali, English and Hindi. In the first sub-task contestants needed to make a system that labeled texts into three classes: ‘Overtly Aggressive’, ‘Covertly Aggressive’ and ‘Non-aggressive’. The dataset contained 5000 texts in total.

We used a single BERT-based system with two outputs for all tasks simultaneously. Our model took the first place in English gendered text classification and the second place in Bengali gendered text classification.

2. Related Work

TRAC
Offensive val
Potapova
2 Kaggle challenges.
Text classification approaches.
Single model in machine translation by Google

3. TRAC-2 dataset

TRAC 2020 competition contained 5000 texts in 3 languages: Bengali, English and Hindi.

The authors of the competition split texts in all languages into training, validation and test datasets.

4. BERT model with multiple outputs

In this task we wanted to experiment with a single model that works with multiple languages at once. We could have used an embedding-based approach [Word2Vec, FastText]. However, pre-trained language models are usually trained for one language at a time. Fortunately, it is possible to overcome this using multilingual language models such as BERT.

BERT (Devlin et al., 2018) is a Transformer-based model (Vaswani et al., 2017). It was trained using Wikipedia texts. All texts were tokenized using byte-pair encoding (BPE) which allows to limit the vocabulary size compared

to Word2vec and other word vector models. The training consisted in predicting a random masked token in the sentence and the next sentence. We did not fine tune the language model. Information about the text language was not included into the model. The model had to infer it itself. We did not perform any text augmentation or pre-processing besides standard byte-pair encoding. We used a multilingual uncased BERT model provided by Hugging Face (Wolf et al., 2019). We used PyTorch framework to create our model.

Half precision training was used vi Apex library.

We used the same training, validation and test datasets as they were provided by the organizers.

5. Results

Task	F1 (weighted)	Accuracy	Rank
Bengali-A	0.7716	0.7811	4
Bengali-B	0.9297	0.9293	2
English-A	0.7568	0.7683	3
English-B	0.8716	0.8708	1
Hindi-A	0.7761	0.7683	4
Hindi-B	0.8381	0.8392	3

Table 1: Results for all tasks

6. Conclusion

This paper describes our results for TRAC 2020 competition held together with the conference LREC 2020. Competition consisted of 2 subtasks where participants had to classify aggression in texts and decide if it is gendered or not for 3 languages: Benghali, English and Hindi. We used a single BERT-based system with two outputs for all tasks simultaneously. Our model took the first place in English gendered text classification and the second place in Bengali gendered text classification.

7. References

Devlin, J., Chang, M.-W., Lee, K., and Toutanova, K. (2018). BERT: Pre-training of Deep Bidirectional Transformers for Language Understanding. oct.

- Vaswani, A., Shazeer, N., Parmar, N., Uszkoreit, J., Jones, L., Gomez, A. N., Kaiser, Ł., and Polosukhin, I. (2017). Attention is all you need. In *Adv. Neural Inf. Process. Syst.*, volume 2017-Decem, pages 5999–6009.
- Wolf, T., Debut, L., Sanh, V., Chaumond, J., Delangue, C., Moi, A., Cistac, P., Rault, T., Louf, R., Funtowicz, M., and Brew, J. (2019). HuggingFace’s Transformers: State-of-the-art Natural Language Processing. *ArXiv*, abs/1910.0.