AI Rational at CheckThat! 2022: Using transformer models for tweet classification

Aleksandar Savchev[a]

[a] *Sofia University, Bulgaria*

Abstract

This paper is an overview of the approach taken by team AI Rational in CheckThat! 2022 for Task1 in English, Bulgarian, Dutch and Turkish. Task 1 is about classifying COVID-19 tweets and has four subtasks: 1A Check-worthiness; 1B Verifiable factual claims detection; 1C Harmful tweet detection; 1D Attention-worthy tweet detection. This document will focus on the experiments done for 1A English where the team got first place out of 13 teams however the same techniques are done for the other languages and subtasks. This document will show our data preprocessing and data augmentation as well as the use of transformer models BERT, DistilBERT and RoBERTa for text classification and how we fined-tuned them for best results.

Keywords

Check-worthiness, COVID-19, transformer models

1. Introduction

In today's day and age with the mass spread of the internet, access to large quantities of information is easily achieved. With the increase of data, the amount of disinformation increases as well. People rely on trustworthy news sites where the articles are redacted and checked by publishers to satisfy their need for information. However, most people use social media sites as well where people can freely post information about different topics which is not fact-checked or can be malicious in nature. Vulnerable topics include but are not limited to finance, healthcare and politics. To combat the spread of misinformation in social media sites Task 1 in CheckThat! 2022 is about classifying tweeter tweets related to COVID-19 in English, Bulgarian, Turkish, Dutch and Arabic. Task 1 has four subtasks:

- **1A Check-worthiness of tweets:** Given a tweet, predict whether it is worth fact-checking.
- **1B Verifiable factual claims detection**: Given a tweet, predict whether it contains a verifiable factual claim.
- **1C Harmful tweet detection:** Given a tweet, predict whether it is harmful to society and why.
- **1D Attention-worthy tweet detection:** Given a tweet, predict whether it should get the attention of policy makers and why.

Models have been trained on all 4 of these subtasks in the languages: English, Bulgarian, Dutch and Turkish. However, this document will show the approach and experiments on subtask 1A English only. For the other subtasks and languages, the same approach was used.

Section 2 of this document will show the most successful models from previous editions of this task. The models used to solve the task: DistilBERT, BERT and RoBERTa. The method of fine-tuning these models. Section 2 also contains what kind of data preprocessing and data augmentation are used on the training data.

2. Usage of transformer models
2.1. Previous successful approaches for CheckThat!

In CheckThat! 2021 the most successful participants have used different transformer models as seen in [1] such models include BERT, RoBERTa, BETO, AraBERT, BERTurk. Some teams have applied data augmentation, and most have done simple data preprocessing. For deeper research papers from different participants were analyzed. In [2] the team has tried different transformer models. For their final submission they have used BERT pre-trained on tweets and have gotten first place in English and fourth in Spanish for task 1A which is the equivalent of clef 2022 task 1A. [7] have experimented with fine-tuning pretrained models BETO and RoBERTa. They have achieved fifth and first place in English and Spanish languages on task 1A.

2.2. Dataset

For task1 English the "train" corpus consists of 2122 entries. Each entry is a tweet related to COVID-19 which is labeled with 0 or 1 on whether it is worth fact-checking or not. The entries also consist of the tweet's id and URL. The organizers have also provided "dev" set of 195 entries and "dev_test" set of 574 entries.

For data preprocessing, all links in the data were replaced with "@link". Other kinds of text edits were not done since these models have their own tokenizer. For English, to increase the training set it was experimented with back translation (English tweets were translated to French and then translated back to English), tweets that remained the same were removed.

For training both entries from the "train" set and "dev" set were used, which total to 2317 entries. With back translation the train set increased to 4439 entries. For validation set the "dev_test" corpus was used.

2.3. Experiments done

Experiments for task1A in English were made with three different pre-trained transformer models: BERT [3], DistilBERT [4] (a distilled version of BERT) and RoBERTa [5]. All models used have been taken from huggingface[1]. Fine-tuning of parameters was done on DistilBERT model since it is the fastest one to train. The more notable hyper parameters are number of epochs 15, warm up ratio 6% and learning ratio 3e-5. These discoveries were used when training the different models. The same settings were used for the other subtasks and languages.

For the other subtasks the same experiments were done where for English the final submissions were with RoBERTa while for the other languages XLM-RoBERTa [6] (a cross-lingual version of RoBERTa) was used. For Dutch and Turkish data augmentation was not used while for Bulgarian the test set was increased with back translation, translating Bulgarian to English and then back to Bulgarian to have more data.

2.4. Results from experiments

Table 1 shows the results of the different experiments on task 1A English with the given train and test set from the organizers. As expected, the best results are with RoBERTa model. Data augmentation helped improve the metrics slightly.

Table 1
Experiments

Model	Accuracy	F1(for positive class)
BERT	0.8118	0.8833
DistilBERT	0.8101	0.8836
RoBERTa	0.8432	0.9010
BERT + Data augmentation	0.8118	0.8838
DistilBERT+Data augmentation	0.8153	0.8862
RoBERTa + Data augmentation	0.8536	0.9070

3. Placements

Table 2 shows the placements in the official leaderboards[2]. The rankings are written in format X/N where X is the place taken by the model and N is the number of participants in that subtask.

Table 2
Placements

Subtask	English	Bulgarian	Dutch	Turkish
1A	1/13	3/5	2/5	-
1B	4/9	1/2	1/2	2/4
1C	2/10	1/2	2/2	3/4
1D	2/5	1/2	1/2	1/2

4. Conclusion and future work

This paper presents the experiments done with transformer models BERT, DistilBERT, RoBERTa and data augmentation technique back translation on task1A in English. The results achieved are satisfactory. For future work experiments with transformer models pre-trained on tweets is advised and experimenting with different data augmentation techniques to increase the data set may yield better results.

5. Acknowledgements

https://docs.google.com/spreadsheets/d/1LMjU7nrl2R7iuAE023kwip2lxQw9jG2V6hEHPChBuHE[2]

6. References

[1] Shaden Shaar, Maram Hasanain, Bayan Hamdan, Zien Sheikh Ali, Fatima Haouari, Alex Nikolov, Mucahid Kutlu, Yavuz Selim Kartal, Firoj Alam, Giovanni Da San Martino, Alberto Barrón-Cedeño, Rubén Míguez, Javier Beltrán, Tamer Elsayed and Preslav Nakov, Overview of the CLEF-2021 CheckThat! Lab Task 1 on Check-Worthiness Estimation in Tweets and Political Debates http://ceur-ws.org/Vol-2936/paper-28.pdf

[2] Juan R. Martinez-Rico, Juan Martinez-Romo and Lourdes Araujo NLP&IR@UNED at CheckThat! 2021: Check-worthiness estimation and fake news detection using transformer models http://ceur-ws.org/Vol-2936/paper-44.pdf

[3] Jacob Devlin, Ming-Wei Chang, Kenton Lee, Kristina Toutanova, BERT: Pre-training of Deep Bidirectional Transformers for Language Understanding https://arxiv.org/pdf/1810.04805.pdf

[4] Victor Sanh, Lysandre Debut, Julien Chaumond, Thomas Wolf , DistilBERT, a distilled version of BERT: smaller, faster, cheaper and lighter https://arxiv.org/pdf/1910.01108.pdf

[5] Yinhan Liu, Myle Ott, Naman Goyal, Jingfei Du, Mandar Joshi, Danqi Chen, Omer Levy, Mike Lewis, Luke Zettlemoyer, Veselin Stoyanov, RoBERTa: A Robustly Optimized BERT Pretraining Approach https://arxiv.org/pdf/1907.11692.pdf

[6] Naman Goyal, Vishrav Chaudhary, Guillaume Wenzek, Francisco Guzmàn, Edouard Grave, Myle Ott, Luke Zettlemoyer, Veselin Stoyanov, Unsupervised Cross-lingual Representation Learning at Scale https://arxiv.org/pdf/1911.02116.pdf

[7] Robiert Sepúlveda-Torres, Estela Saquete GPLSI team at CheckThat! 2021: Fine-tuning BETO and RoBERTa http://ceur-ws.org/Vol-2936/paper-52.pdf