GCDH at WNUT-2020 Task 2: BERT-Based Models for the Detection of Informativeness in English COVID-19 Related Tweets

Hanna Varachkina[1], Stefan Ziehe[2], Tillmann Dönicke[3], Franziska Pannach[3]

[1]Department for German Philology, [2]Institute for Informatics, [3]Göttingen Centre for Digital Humanities

Data

Tweets
informative & uninformative
Training set 7,000 tweets
Validation set 1,000 tweets

Our source code is available at https://gitlab.gwdg.de/tillmann.doenicke/wnut-2020.

Baselines

Naive Bayes (NBC) with binary features
Support vector machine (SVM) with binary, count features (frequency of n-gram) and tf·idf features
Majority of inf. & uninf. tweets have neutral sentiment and could not be uses as feature

Sentiment analysis

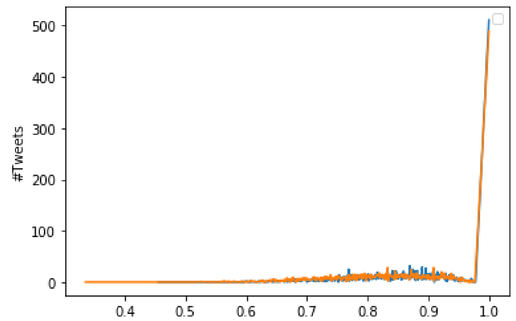

Fig. 1: Informative (orange) and uninformative (blue) tweets, 1 being perfectly neutral

BERT-based models

Fine-tuning

1. BERT large uncased for two epochs
2. COVID-Twitter-BERT (CT-BERT) for three epochs
   - sequence length: 96
   - training batch size: 32
   - evaluation batch size: 1024
   - Adam optimizer
   - a base learning rate: 2e-5
   - smaller learning rates for earlier layers, decay factor $\xi = 0.95$

Evaluation on the validation set

Model	F1	Precision	Recall	Accuracy
NBC	0.8012	0.7721	0.8326	0.8050
SVM	0.8009	0.8380	0.7669	0.8200
BERT	0.8928	0.8965	0.8891	0.8984
CT-BERT-1	0.9231	0.8988	0.9487	0.9248

Tab. 1: F1-score, Precision, Recall and Accuracy on the validation set

Best results

NBC when trained with uni- and bigrams as features and not masking numbers
SVM when trained with uni- and bigrams as features and masking numbers
BERT-based when trained on on the domain-specific Twitter data (CT-BERT)
CT-BERT-1 is the most performative model

CT-BERT comparison

Model	Training data	F1
CT-BERT-1	6,912 tweets	0.9091
CT-BERT-2	7,936 tweets	0.9036

Tab. 2: Model overview and F1-score on the final evaluation set

Evaluation on the test set

Model	Precision	Recall	Accuracy
CT-BERT-1	0.8919	0.9269	0.9125
CT-BERT-2	0.9036	0.9036	0.9090

Tab. 3: Precision, Recall and Accuracy on the final evaluation set

✉ hanna.varachkina@stud.uni-goettingen.de ✉ stefan.ziehe@cs.uni-goettingen.de ✉ tillmann.doenicke@uni-goettingen.de ✉ franziska.pannach@stud.uni-goettingen.de