Sarcasm Detection with Stance detection

Prakhar Srivastava
prakhar.srivastava@ut.ee

Hasan Tanvir
hasantanvir79@gmail.com

Alo Aasmäe
aloaas@gmail.com

Motivation

Figurative language detection is not only challenging for computers, it is equally difficult for human to understand the conveyed information with respect to the lexical meaning. Given the difficulty, we approached solving it with semantic NLP.

Background

Limited number of researches from rule-based techniques to state-of-the art transformer based pretrained models, were performed over time. Use of several machine learning techniques, such as SVM, RNN, decision tree, CNN paved the way towards application of transformer based models. There are combination of models, RCNN-RoBERTa, that showed interesting result.

In our project, we applied GloVE+RNN including stance detection and in second approach BERT for binary classification.

Workflow

Training BERT: The training of the transformer based bert model, was straight-forward. At first the textual data was tokenized, the dataset was splitted between train and test, the model was built and trained. FInally, the model was tested using the test dataset.

Training GloVE & Stance: For our second approach, we have used GloVe for feature extraction and trained it with LSTM(RNN). The dataset is preprocessed using the traditional and NLTK method. The GloVe model is used to obtain vector representation by mapping words into semantic meaningful space.
For Stance, we have used the 'Udacity profanity online model' as a meter for sarcastic text.

Implementation

Training BERT: The dataset was divided in train-dev-test (80/10/10 split). The textual data was encoded using pretrained bert-base-uncased tokenizer, the tokenized sentences were later converted to tensors.

Then a classifier layer was added on top of the BERT encoder that includes 4 dense layers and the output layer. For the hidden layers, we used a combination of rectified linear, hyperbolic tangent, sigmoid activation function. The model was trained for 109,712,129 trainable parameters for 15 epochs and tested using the development dataset and finally tested on the test dataset.

Training GloVE: After the data preprocessing, the dataset was splitted into train-validation (80/20) split.

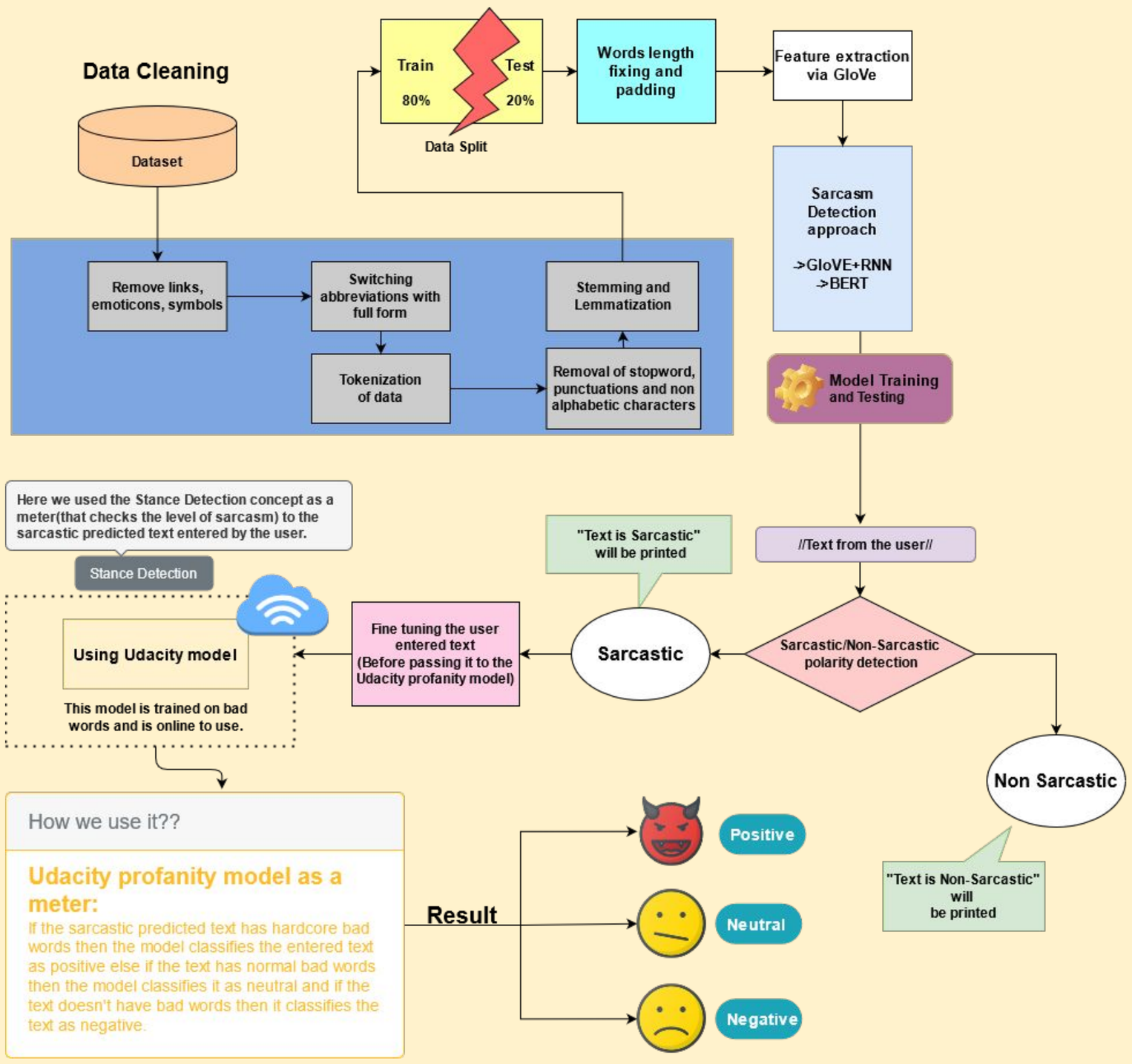

Figure 1. Sarcasm Detection with Stance detection flowchart.

The model was built with three layers, including the embedding (GloVe), LSTM, and a hidden layer. In total 5,731,600 parameters, of which 67,840 are trainable After training for 10 epochs, the results are good, with validation accuracy reaching 0.91, but not outperforming the model that uses BERT.

As an add-on to our Sarcasm detection, we have used the concept of Stance Detection as a meter to measure the quality of sarcastic predicted text entered by the user. For this meter, we have used an online model 'Udacity profanity model' which is trained on bad words. This is how it works: If the sarcastic predicted text has hardcore bad words then the model classifies the entered text as positive else if the text has normal bad words then the model classifies it as neutral and if the text doesn't have bad words then it classifies the text as negative.

Note: Positive, neutral and negative is printed by the model based on its classification along with the emoticons as shown in Figure 1.

	Precision	Recall	F1-score
Not sarcastic (0)	0.91	0.95	0.93
Sarcastic (1)	0.94	0.90	0.92
Weighted average	0.93	0.92	0.92

Table 1. Classification report on validation data after training model with BERT for 15 epochs.

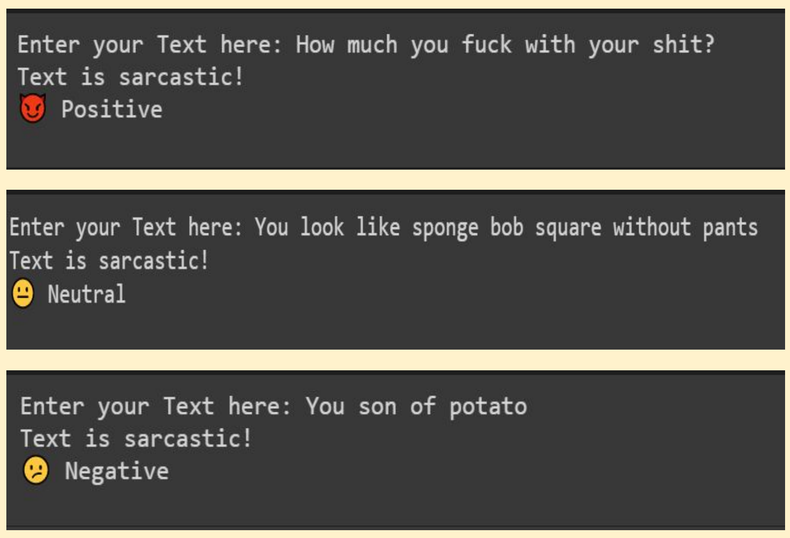

Figure 2. Sample predictions for sarcasm detection with GloVe.

Results

The results were surprisingly similar for both models, with the BERT model slightly outperforming GloVe model, but both having high accuracy (0.92 compared to 0.91). Figure 2 displays the result of user interaction by inputting a text and receiving a classification of sarcasm with a stance de-tection. “You son of potato” and “I love NLP” was predicted as sarcastic and not-sarcastic using the BERT model as well.

Conclusion

Both of the pretrained models, showed high accuracy in sarcasm detection. The results are better than traditional machine learning or rule-based methods. BERT outperformed GloVE in accuracy although the difference is very small.

References
Aditya Joshi, Pushpak Bhattacharyya, and Mark J. Carman. 2017. Automatic Sarcasm Detection: A Survey. ACM Comput. Surv. 50, 5, Article 73 (November 2017), 22 pages. DOI:https://doi.org/10.1145/3124420
Potamias, R.A., Siolas, G. & Stafylopatis, A . A transformer-based approach to irony and sarcasm detection. Neural Comput & Applic 32, 17309–17320 (2020). https://doi.org/10.1007/s00521-020-05102-3
P. Shrikhande, V. Setty and D. A. Sahani, Sarcasm Detection in Newspaper Headlines, 2020 IEEE 15th International Conference on Industrial and Information Systems (ICIIS), RUPNAGAR, India, 2020, pp. 483-487, DOI: 10.1109/ICIIS51140.2020.9342742
Dilek Küçük and Fazli Can. 2020. Stance Detection: A Survey. ACM Comput. Surv. 53, 1, Article 12 (May 2020), 37 pages. DOI:https://doi.org/10.1145/3369026
Kumar, A. and Anand, V., 2020, July. Transformers on sarcasm detection with context. In Proceedings of the Second Workshop on Figurative Language Processing (pp. 88-92).

Dataset

	article_link	headline	is_sarcastic
0	https://www.huffingtonpost.com/entry/versace-b...	former versace store clerk sues over secret 'b...	0
1	https://www.huffingtonpost.com/entry/roseanne-...	the 'roseanne' revival catches up to our thorn...	0
2	https://local.theonion.com/mom-starting-to-fea...	mom starting to fear son's web series closest ...	1
3	https://politics.theonion.com/boehner-just-wan...	boehner just wants wife to listen, not come up...	1
4	https://www.huffingtonpost.com/entry/jk-rowlin...	j.k. rowling wishes snape happy birthday in th...	0