MASTER IN BUSINESS ANALYTICS 2019/2020

Stance and Sentiment detection in Tweets

Research proposal

COMP41760_2019 Business Analytics Project

Z0151576

Table of Contents

1. Introduction with a problem statement

The rapid development and proliferation of social networks, messengers, online forums, news portals, and blogs unleashed an enormous potential for internet users to share their opinion, attitude and position towards various topics such as public policy, religion, and environment. Besides, the advent of digital technologies creates a new form of interaction between customers and business. For example, by reviewing customers' feedback from digital channels, retail company management may perform sentiment analysis to monitor public opinion, brand reputation and customer loyalty about launched products and services. However, because a text may consist of elements of ambiguity, hints, sarcasm, irony or implicit links, the primary challenge for business entities remains the correct identification and understanding the customer's stance towards a specific target. For example, observing the following tweet:

"Faithful He has been. Faithful He will be. #SemST"

It might be evident for a human annotator to identify the stance of given text as "Against" towards target "Atheism", but a difficult task to the computer. Moreover, there is also an existing problem for companies and government agencies for rumours and fake news recognition. To address these issues, the present report proposes the automatic stance and sentiment detection approach, capable of analysing and classifying the piece of text into various classes.

In particular, the stance and sentiment classification will be performed employing SemEval-2016 Task 6A stance dataset (Mohammad et al., 2016) with 4,163 tweets from Twitter microblogging platform. Besides, this research will use a Transformer based model (Vaswani et al., 2017), which leverages an attention mechanism (Bahdanau et al., 2014, Luong et al.,

2015) to incorporate target information in the process of stance and sentiment detection. Pre-trained on a large unlabeled text Transformer models allows performing several NLP-related tasks without training a neural network from scratch.

Access to information about the stance and sentiment in the text corpus triggers several research questions:

- How to explore the relationship between sentiment and stance in tweets?
- To what extent the sentiment information contributes to stance detection?
- Can Transformer-based models simultaneously detect stance and sentiment in the text?

Thus, the research objectives are developing the model capable of detecting stance and sentiment information from Twitter, classifying obtained information with respect to a given set of targets (for stance) and polarities (for sentiment), comparing the model performance with several baselines.

2. Literature review

The growing interest to stance detection problem from both research and business community triggered by the promising potential in diverse areas such as market analysis, an opinion survey, polling, retrieving and understanding customers reviews, rumours and fake news detection, fact-checking and social media monitoring (Küçük and Can, 2020).

Stance detection stands for Natural Language Processing (NLP) task for determining whether the author of the textual information is in support of, neutral or against towards a predefined target of interest (Mohammad et al., 2016). The target object can be the product, an organisation, a person, and a claim. Alternatively, the sentiment detection task aims to recognise whether the text has been written in the positive, neutral or negative way. The crucial difference between these two tasks is that stance detection depends on specific target information, which the author might not explicitly mention in the content (Du et al., 2017).

Early research related to stance detection problem conducted by Mohammad et al. (2016) and based on the SemEval-2016 shared task competition. The SemEval-2016 dataset consists of more than 4,000 tweets derived from Twitter and manually annotated with sentiment and stance labels. Consequently, several models for stance detection (Augenstein et al., 2016, Schuff et al., 2017, Zhou et al., 2017, Li and Caragea, 2019, Xu et al., 2018, Du et al., 2017, Dey et al., 2018, Sun et al., 2018, Wei et al., 2018, Siddiqua et al., 2019, Zhou et al., 2019, Lai et al., 2020) were proposed and evaluated on the basis of annotated SemEval-2016 dataset.

In the original paper, Mohammad et al. (2016) performed stance detection by encoding words with Word2Vec Skip-gram model (Mikolov et al., 2013) and trained the model with a Support Vector Machine (SVM) classifier. The same stance problem was investigated employing Recurrent Neural Networks (RNN) with Long Short-Term Memory (LSTM) architecture

(Augenstein et al., 2016, Zarrella and Marsh, 2016). Schuff et al. (2017) extended SemEval-2016 stance dataset with emotion annotations to investigate the relationship between emotions and (sentiment, stance) pair, and achieved promising results with Bidirectional LSTM (BI-LSTM) architecture. Besides, studies for rumour stance detection from text corpus has been executed with Branch-LSTM (Kochkina et al., 2017) and XGBoost (Bahuleyan and Vechtomova, 2017) models within SemEval 2017 RumourEval competition. However, the majority of these studies perform stance detection relying on the extracted features from the text while ignoring the target specific information.

To overcome this limitation, Bahdanau et al. (2014) and Luong et al. (2015) proposed a "Attention" technique, which allows the model to focus on the specific salient parts of the text. Although this technique initially was proposed for machine translation purposes, the research community successfully applied it on stance and sentiment detection problems. For example, Zhou et al. (2017) proposed combined Gated Recurrent Unit (GRU) - Convolutional Neural Network (CNN) architecture that incorporated semantic-level attention mechanism and therefore better performs target-specific stance detection. A study conducted by Du et al. (2017) similarly proposed Target - specific Attentional Network (TAN) on both English and Chinese text corpora with BI-LSTM model for feature extraction. Dey et al. (2018) further extended the research by including a two-phase attention-embedded LSTM-based approach to capture the author's subjectivity first and then perform stance classification. Sun et al. (2018) presented a hierarchical attention network, which explored the mutual relationship between the whole text and the determined linguistic feature set. Later, Xu et al. (2018) introduced the concept of cross-target stance classification, where the stance detection model might be trained on the source target and then generalised to another related target. Wei et al. (2018) proposed end-to-end neural memory model that utilised a target-guided iteration process whereas

Siddiqua et al. (2019) developed the ensemble model with combined Bi-LSTM and nested LSTMs architectures, powered with a feed-forward attention mechanism.

Furthermore, Zhou et al. (2019) demonstrated Condensed CNN model by removing noisy tokens and making stance-relevant words more salient. Proposed CNN model was also equipped with self-attention and attention-based condensation modules to enhance embeddings representation in the text and subsequently eliminate redundant words. Sobhani et al. (2017) questioned previous studies for treating targets independently and proposed a multi-target stance detection approach, which assumed natural interdependency among targets. Finally, Li and Caragea (2019) presented a multi-task stance detection model based on sentiment and target related attention. Moreover, Li and Caragea (2019) also reformulated cross-entropy loss function, which using special constructed stance and sentiment lexicons, guided the attention mechanism.

While several stance and sentiment detection tasks have been performed with RNN, LSTM and GRU algorithms, the major downside of such methods are processing sequential input data words by words (Goodfellow et al., 2016). Therefore, training a large volume of the text might be a time consuming and expensive task. In contrast, CNN-based models capable of parallel text processing but fail to capture long-term dependencies in the text. Besides, although using Word2Vec embeddings technique to map input tokens into continuous vector space prevents from extensive feature engineering tasks, these pre-trained words usually static and fails to capture words with polysemous meanings (Samih and Darwish, 2020). To address these issues, Vaswani et al. (2017) introduced Transformer architecture, which entirely based on self-attention and positional encoding tools that replace recurrent and convolutional layers and facilitate parallel computations over text corpus (Vaswani et al., 2017). A self-attention mechanism refers to the process of computing token representation within a single sequence (Veličković et al., 2017), whereas positional encoding allows a Transformer model to memorise a word position during parallel training (Vaswani et al., 2017). The advent of Transformer architecture triggered the emergence of several language models. For instance, Bidirectional Encoder Representations from Transformers (BERT) represents trained on the large unlabeled datasets model, built using encoder blocks with bidirectional self-attention and dense feed-forward network layers inside (Devlin et al., 2018). Bidirectionality stands for the ability to capture information from both the left and right side of a sentence during the model training stage (Goodfellow et al., 2016). In contrast, OpenAI's Generative Pre-trained Transformer (GPT-2) model consist of stacked decoder blocks with masked self-attention layer and trained for the next word prediction task (Radford et al., 2019). Later, Liu et al. (2019) proposed BERT-based RoBERTa model, which trains more data with increased batch size and dynamic masking technique. Consequently, Lan et al. (2019) also extended BERT model presenting ALBERT transformer with parameter sharing mechanism.

In the recent few years, the research community focused on Transformer-based models for stance and sentiment detection tasks. Popat et al. (2019) leveraged augmented BERT representations for stance classification and additionally claim and perspective capturing. Samih and Darwish (2020) extended Popat et al. (2019) research by joining BERT with unsupervised classification to detect less active Twitter users. Slovikovskaya (2019) added BERT, XLNet (Yang et al., 2019), and RoBERTa transformers to enhance results achieved on Fake News detection competition. Tian et al. (2020) proposed CNN and BERT based models for early rumour detection in Twitter by analysing users' comments at the early stage. Zhao et al. (2020) performed sentiment analysis and entity recognition on financial text corpus applying RoBERTa transformer.

However, there is a research gap related to simultaneous detection of stance and sentiment from social media with Transformer based models, where sentiment and stance information might be correlated and eventually facilitate detecting each other.

3. Research design and methods

3.1 Dataset

To evaluate the proposed models and answer research questions, we use the SemEval-2016 Task 6A dataset (Mohammad et al., 2016) with more than 4,000 English tweets from Twitter microblogging platform. The dataset is publicly available for upload from Roman Klinger's Homepage (2020) and split into training (with 2,914 tweets) and test (with 1,249 tweets) datasets. Figure 1 shows the first five rows of the training dataset. Each tweet has been annotated with stance labels (“favor”, “against”, “neither”), and sentiment labels (“positive”, “negative”, “neither”). The dataset contains five targets of interests: “Atheism”, “Climate Change is a Real Concern”, “Feminist Movement”, “Hillary Clinton”, and “Legalization of Abortion”. Finally, the dataset includes “Opinion Toward column”, which indicates whether a stance maker explicitly mentions a target or not.

	Tweet	Target	Stance	Opinion Towards	Sentiment
0	@tedcruz And, #HandOverTheServer she wiped cle...	Hillary Clinton	AGAINST	1. The tweet explicitly expresses opinion abo...	neg
1	Hillary is our best choice if we truly want to...	Hillary Clinton	FAVOR	1. The tweet explicitly expresses opinion abo...	pos
2	@TheView I think our country is ready for a fe...	Hillary Clinton	AGAINST	1. The tweet explicitly expresses opinion abo...	neg
3	I just gave an unhealthy amount of my hard-ear...	Hillary Clinton	AGAINST	1. The tweet explicitly expresses opinion abo...	neg
4	@PortiaABoulger Thank you for adding me to you...	Hillary Clinton	NONE	3. The tweet is not explicitly expressing opi...	pos

Figure 1. The first five rows of the training SemEval-2016 dataset.

Table 1 illustrates the number and distribution of stance and sentiment labels over five different targets on train and test datasets. According to Table 1, the percentage of "Against" stance label represents about 48% of the whole training dataset and varies across five targets from 3.8% ("Climate change is a concern") to 59.3% ("Atheism"). Sentiment distribution, except for "Atheism" target, demonstrates a predominance of negative labels across other targets.

		Stance			Sentiment		
	#Train	AGAINST	FAVOR	NONE	pos	neg	other
Atheism	513	59.3%	17.9%	22.8%	60.4%	35.1%	4.5%
Climate Change is a Real Concern	395	3.8%	53.7%	42.5%	31.6%	49.6%	18.7%
Feminist Movement	664	49.4%	31.6%	19.0%	17.9%	77.3%	4.8%
Hillary Clinton	689	57.0%	17.1%	25.8%	32.1%	64.0%	3.9%
Legalization of Abortion	653	54.4%	18.5%	27.1%	28.8%	66.2%	5.1%
Total	**2914**	**47.9%**	**25.8%**	**26.3%**	**33.0%**	**60.5%**	**6.5%**

		Stance			Sentiment		
	#Test	AGAINST	FAVOR	NONE	pos	neg	other
Atheism	220	72.7%	14.5%	12.7%	59.1%	35.5%	5.5%
Climate Change is a Real Concern	169	6.5%	72.8%	20.7%	29.6%	51.5%	18.9%
Feminist Movement	285	64.2%	20.4%	15.4%	19.3%	76.1%	4.6%
Hillary Clinton	295	58.3%	15.3%	26.4%	25.8%	70.2%	4.1%
Legalization of Abortion	280	67.5%	16.4%	16.1%	20.4%	72.1%	7.5%
Total	**1249**	**57.2%**	**24.3%**	**18.4%**	**29.5%**	**63.3%**	**7.2%**

Table 1. Data distribution of SemEval-2016 Task 6A dataset.

3.2 Research design approach

The present study follows a quantitative type of research approach since stance and sentiment detection tasks include a series of experiments with Transformer models, investigations of correlation between stance and sentiment information and comparison achieved results with strong baselines. Specifically, the present quantitative study is intended to perform a multiclass classification with SemEval-2016 Task 6A dataset employing Transformer based models. Besides, all research-related computations will be performed using the Python programming language and the Jupyter Notebook environment. Transformer models will be implemented in the TensorFlow Python library (Abadi et al., 2016).

3.3 Data preprocessing

The first step in preprocessing the data is to remove non-English and Unicode characters from the "Tweet" column of the SemEval-2016 dataset. The next step is to lowercase and split the dataset into training and validation sets. Finally, the preprocessed text needs to be tokenised and converted into numerical representation for further processing using Transformers.

3.4 Data analysis methods

The present research project employs Transformer based models for data analysis purposes, which have been trained in an unsupervised manner on a large text corpus and might then be applied to perform supervised tasks. Specifically, we will experiment with the BERT Transformer, which was pre-trained on a large unlabeled text from BookCorpus and English Wikipedia (Devlin et al., 2018) and received an in-depth representation of words in a bidirectional manner. Therefore, although initially BERT was pre-trained for masked language modelling and next sentence prediction tasks, it could be fine-tuned for stance and sentiment classification tasks by placing additional output layers in the existing architecture. For computational simplicity, we choose the smaller Uncased BERT-Base architecture with 12 hidden layers, 12 attention heads, 768-dimensional hidden layers and 110M parameters.

Firstly, preprocessed tweets from training and validation sets need to be encoded with Bert Tokenizer tool, which splits the word into tokens, adds special symbols such as CLS at the beginning and SEP at the end of the sentence and converts tokens to integer numbers. Then, since our goal is to perform stance detection using sentiment information, the BERT architecture design will consider auxiliary modules for sentiment detection stacked with base modules for stance classification. Finally, the output layer will be supplemented with feed-forward dense stacked with softmax layer to obtain final class labels.

The results will be evaluated on the test dataset with a macro and micro average of F1 score. The selection of these evaluation metrics is motivated by their common use in the research community for similar tasks.

The final model performance will be compared with strong baseline models, which utilised both deep learning and transformer-based models for stance and sentiment detection.

3.5 Ethical issues and commercial sensitivity

Since the present study is fully research-oriented, and the dataset has been downloaded from a public source, this project does not address any ethical issues and commercial sensitivity.

4. Reference list

Abadi, M., Agarwal, A., Barham, P., Brevdo, E., Chen, Z., Citro, C., Corrado, G.S., Davis, A., Dean, J., Devin, M. and Ghemawat, S. (2016) 'Tensorflow: Large-scale machine learning on heterogeneous distributed systems', *arXiv preprint arXiv:1603.04467*.

Augenstein, I., Rocktäschel, T., Vlachos, A. and Bontcheva, K. (2016) 'Stance detection with bidirectional conditional encoding', *arXiv preprint arXiv:1606.05464*.

Bahdanau, D., Cho, K. and Bengio, Y. (2014) 'Neural machine translation by jointly learning to align and translate', *arXiv preprint arXiv:1409.0473*.

Bahuleyan, H. and Vechtomova, O. (2017) 'UWaterloo at SemEval-2017 Task 8: Detecting stance towards rumours with topic independent features', In *Proceedings of the 11th International Workshop on Semantic Evaluation (SemEval-2017)* (pp. 461-464).

Borges, L., Martins, B. and Calado, P. (2019) 'Combining similarity features and deep representation learning for stance detection in the context of checking fake news', *Journal of Data and Information Quality (JDIQ)*, *11*(3), pp.1-26.

Dai, Z., Yang, Z., Yang, Y., Carbonell, J., Le, Q.V. and Salakhutdinov, R. (2019) 'Transformer-xl: Attentive language models beyond a fixed-length context', *arXiv preprint arXiv:1901.02860*.

Dey, K., Shrivastava, R., and Kaushik, S. (2018) 'Topical Stance Detection for Twitter: A Two-Phase LSTM Model Using Attention', *ArXiv.org,* ArXiv.org, Jan 9, 2018.

Du, J., Xu, R., He, Y. and Gui, L. (2017) ‘Stance classification with target-specific neural attention networks’, International Joint Conferences on Artificial Intelligence.

Elbayad, M., Besacier, L. and Verbeek, J. (2018) ‘Pervasive attention: 2d convolutional neural networks for sequence-to-sequence prediction’, *arXiv preprint arXiv:1808.03867*.

Goodfellow, I., Bengio, Y. and Courville, A. (2016) *Deep learning*. MIT press.

Kochkina, E., Liakata, M. and Augenstein, I. (2017) ‘Turing at semeval-2017 task 8: Sequential approach to rumour stance classification with branch-lstm’, *arXiv preprint arXiv:1704.07221*.

Küçük, D. and Can, F. (2020) ‘Stance detection: A survey’, *ACM Computing Surveys (CSUR)*, *53*(1), pp.1-37.

Lai, M., Cignarella, A., Hernández Farías, D., Bosco, C., Patti, V., and Rosso, P. (2020). ‘Multilingual stance detection in social media political debates’, *Computer Speech & Language, 63*, Computer Speech & Language, September 2020, Vol.63.

Lan, Z., Chen, M., Goodman, S., Gimpel, K., Sharma, P. and Soricut, R. (2019) ‘Albert: A lite bert for self-supervised learning of language representations’, *arXiv preprint arXiv:1909.11942*.

Li, Y. and Caragea, C. (2019) ‘November. Multi-Task Stance Detection with Sentiment and Stance Lexicons’, In *Proceedings of the 2019 Conference on Empirical Methods in Natural*

Language Processing and the 9th International Joint Conference on Natural Language Processing (EMNLP-IJCNLP) (pp. 6300-6306).

Liu, Y., Ott, M., Goyal, N., Du, J., Joshi, M., Chen, D., Levy, O., Lewis, M., Zettlemoyer, L. and Stoyanov, V. (2019) 'Roberta: A robustly optimized bert pretraining approach', *arXiv preprint arXiv:1907.11692*.

Luong, M.T., Pham, H. and Manning, C.D. (2015) 'Effective approaches to attention-based neural machine translation', *arXiv preprint arXiv:1508.04025*.

Mikolov, T., Sutskever, I., Chen, K., Corrado, G.S. and Dean, J. (2013) 'Distributed representations of words and phrases and their compositionality', In *Advances in neural information processing systems* (pp. 3111-3119).

Mohammad, S., Kiritchenko, S., Sobhani, P., Zhu, X. and Cherry, C. (2016) 'Semeval-2016 task 6: Detecting stance in tweets', In Proceedings of the 10th International Workshop on Semantic Evaluation (SemEval-2016) (pp. 31-41).

Popat, K., Mukherjee, S., Yates, A. and Weikum, G. (2019) 'STANCY: Stance classification based on consistency cues', *arXiv preprint arXiv:1910.06048*.

Radford, A., Wu, J., Child, R., Luan, D., Amodei, D. and Sutskever, I. (2019) 'Language models are unsupervised multitask learners', *OpenAI Blog*, *1*(8), p.9.

Roman Klinger's Homepage. (2020). *Stance Sentiment Emotion Corpus (SSEC)*. [online] Available at: http://www.romanklinger.de/ssec/ [Accessed 13 June 2020].

Samih, Y. and Darwish, K. (2020) ‘A Few Topical Tweets are Enough for Effective User-Level Stance Detection’ *arXiv preprint arXiv:2004.03485*.

Schuff, H., Barnes, J., Mohme, J., Padó, S. and Klinger, R. (2017) ‘Annotation, modelling and analysis of fine-grained emotions on a stance and sentiment detection corpus’, In *Proceedings of the 8th Workshop on Computational Approaches to Subjectivity, Sentiment and Social Media Analysis* (pp. 13-23).

Siddiqua, U.A., Chy, A.N. and Aono, M. (2019) ‘Tweet Stance Detection Using an Attention based Neural Ensemble Model’, In *Proceedings of the 2019 Conference of the North American Chapter of the Association for Computational Linguistics: Human Language Technologies, Volume 1 (Long and Short Papers)* (pp. 1868-1873).

Slovikovskaya, V. (2019) ‘Transfer Learning from Transformers to Fake News Challenge Stance Detection (FNC-1) Task’, *arXiv preprint arXiv:1910.14353*.

Sobhani, P., Inkpen, D. and Zhu, X. (2017) ‘A dataset for multi-target stance detection’, In *Proceedings of the 15th Conference of the European Chapter of the Association for Computational Linguistics: Volume 2, Short Papers* (pp. 551-557).

Sun, Q., Wang, Z., Zhu, Q. and Zhou, G. (2018) 'Stance detection with hierarchical attention network', In *Proceedings of the 27th International Conference on Computational Linguistics* (pp. 2399-2409).

Tian, L., Zhang, X., Wang, Y. and Liu, H. (2020) 'Early detection of rumours on Twitter via stance transfer learning', In *European Conference on Information Retrieval* (pp. 575-588). Springer, Cham.

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 *Advances in neural information processing systems* (pp. 5998-6008).

Veličković, P., Cucurull, G., Casanova, A., Romero, A., Lio, P. and Bengio, Y. (2017) 'Graph attention networks', arXiv preprint arXiv:1710.10903.

Wei, P., Mao, W. and Zeng, D. (2018) 'A target-guided neural memory model for stance detection in Twitter', In *2018 International Joint Conference on Neural Networks (IJCNN)* (pp. 1-8). IEEE.

Xu, C., Paris, C., Nepal, S. and Sparks, R. (2018) 'Cross-target stance classification with self-attention networks', arXiv preprint arXiv:1805.06593.

Yang, Z., Dai, Z., Yang, Y., Carbonell, J., Salakhutdinov, R.R. and Le, Q.V. (2019), 'Xlnet: Generalized autoregressive pretraining for language understanding', In *Advances in neural information processing systems* (pp. 5754-5764).

Zarrella, G. and Marsh, A. (2016) ‘Mitre at semeval-2016 task 6: Transfer learning for stance detection’, *arXiv preprint arXiv:1606.03784*.

Zhao, L., Li, L. and Zheng, X. (2020) ‘A BERT based Sentiment Analysis and Key Entity Detection Approach for Online Financial Texts’, *arXiv preprint arXiv:2001.05326*.

Zhou, S., Lin, J., Tan, L. and Liu, X. (2019) ‘Condensed Convolution Neural Network by Attention over Self-attention for Stance Detection in Twitter’, In *2019 International Joint Conference on Neural Networks (IJCNN)* (pp. 1-8). IEEE.