Research Papers on Stance Detection - A Brief Summary

Paper 1 – Topical Stance Detection for Twitter: A Two-Phase LSTM Model Using Attention.
Authors – Kuntal Dey, Ritvik Shrivastava, Saroj Kaushik
Publication Year – 2018

- **Problem statement -** The topical stance detection problem addresses detecting the stance of the text content with respect to a given topic.

- **Motivation -** Twitter, a hotbed of user generated content, has recently found traction among the researchers for the problem of topical stance detection.

- **Methodology**

  T-PAN, a two-phase LSTM-based model with attention embedding is proposed, for detecting user stance with respect to given topics on Twitter. First, tweets from data is classified into two: neutral and non-neutral, where non-neutral comprised of favour and against stances. Second, classified the tweets were labelled as non-neutral in the first phase, into two - favour and against stances. In each phase, the input sentences were encoded in form of a sequence of words using a bi-directional LSTM, and attention embedding. The impact of embedding topical attention, as well as, the impact of different LSTM architectures, on our approach was investigated.

- **Dataset**

  The benchmark training and test data provided by the SemEval 2016 stance detection task is being used. To perform data cleaning: net slang removal (for tweet normalization) using an online dictionary and stopword removal using a Stanford NLP resource for stopword removal are used.

- **Result**

  Best system (the T-PAN model) outperforms the state of the art that uses deep neural networks for topical stance classification. Out of the five given classes, it performs the best in one class, the TAN model outperforms this model in two classes and the SemEval tasks perform better than this model for the other two classes.

Target	NBOW	LSTM	LSTM$_E$	TOP Sem-Eval	TAN	*T-PAN*
Atheism	55.12	58.18	59.77	**61.47**	59.33	61.19
C.C.C.	39.93	40.05	48.98	41.63	53.59	**66.27**
Feminist Movement	50.21	49.06	52.04	**62.09**	55.77	58.45
Hillary Clinton	55.98	61.84	56.89	57.67	**65.38**	57.48
L.A.	55.07	51.03	60.34	57.28	**63.72**	60.21
Overall	60.19	63.21	66.24	67.82	68.79	**68.84**

- **Novelty**

  The problem is useful to solve certain problems in several practical scenarios, such as detecting user stance towards aspects of political, economic and social events, understanding stance-specific information propagation behaviour of users etc.

- **Strength**

  1. Model is easy to implement, reusable and practicable.
  2. Delivered the highest-known performance among all the deep learning approaches.

- **Weakness**

  1. Underperformed for some target classes.
  2. Accuracy lower than even baseline models for some classes.

Paper 2 – Automatic Stance Detection Using End-to-End Memory Networks.
Authors – Mitra Mohtarami, Ramy Baly, James Glass, Preslav Nakov, Lluís Màrquez, Alessandro Moschitti
Publication Year – 2018

- **Problem statement -** Creation of a novel end-to-end memory network for stance detection, which jointly (i) predicts whether a document agrees, disagrees, discusses or is unrelated with respect to a given target claim, and also (ii) extracts snippets of evidence for that prediction.

- **Motivation -** An unprecedented amount of false information has been flooding the Internet with aims ranging from affecting individual people's beliefs and decisions to influencing major events such as political elections. Consequently, manual fact checking has emerged with the promise to support accurate and unbiased analysis of public statements. As manual fact checking is a very tedious task, automatic fact checking has been proposed as an alternative.

- **Methodology**

  A novel memory network model enhanced with CNN and LSTM networks for stance detection is being utilised. Further a novel extension of the general architecture is proposed based on a similarity-based matrix, which is used at inference time, and showed that this extension offers sizable performance gains. Finally, it is shown that this model is capable of extracting meaningful snippets from the input text document, which is useful not only for stance detection, but more importantly can be useful for human experts who need to decide on the factuality of a given claim.

- **Dataset**

  The dataset provided by the Fake News Challenge is used, where each example consists of a claim–document pair with the following possible relationship: agree, disagree, discuss, unrelated. The data includes a total of 75.4K claim-document pairs, which link 2.5K unique articles with 2.5K unique claims, i.e., each claim is associated with 29.8 articles on average.

- **Result -** Given table reports the performance of all models on the test dataset.

Methods	Total Parameters	Trainable Parameters	Weighted Accuracy	Macro-F1	Accuracy
1. All-*unrelated*	–	–	39.37	20.96	72.20
2. All-*discuss*	–	–	43.89	7.47	17.57
3. CNN	2.7M	188.7K	40.66	24.44	41.53
4. LSTM	2.8M	261.3K	57.23	37.23	60.21
5. CNN+LSTM	4.2M	361.5K	42.02	27.36	48.54
6. LSTM+CNN	2.8M	281.5K	60.21	40.33	65.36
7. Gradient Boosting	–	–	75.20	46.13	86.32
8. sMemNN (dotProduct)	5.4M	275.2K	75.13	50.21	83.85
9. sMemNN	5.5M	377.5K	78.97	56.75	87.27
10. sMemNN (with TF)	110M	105M	**81.23**	**56.88**	**88.57**

  Further analysis of the output of the different systems (e.g., the confusion matrices) reveals the following general trends: (i) The unrelated examples are easy to detect, and most models show high performance for this class, (ii) The agree and the disagree examples are often mislabelled as discuss by the baselines, and (iii) The disagree examples are the most difficult ones for all models.

- **Novelty**

  As manual fact checking is a very tedious task, automatic fact checking has been proposed as an alternative. This is often broken into intermediate steps in order to alleviate the task complexity. One such step is stance detection, which is also useful for human experts as a stand-alone task.

- **Strength –**

  1. Major advantage of this model, is that it can explain its predictions.

- **Weakness**

  1. Prediction are explained only at instance level.
  2. A more coarse-grained implementation has to be done.

Paper 3 – MITRE at SemEval-2016 Task 6: Transfer Learning for Stance Detection.
Authors – Guido Zarrella, Amy Marsh
Publication Year – 2016

- **Problem statement –** Making a system for performing automatic stance detection in social media messages.

- **Motivation -** To the human observer messages like these contain an interpretable stance relevant to the topic of climate change. But to understand rhetorical devices like sarcasm, irony, analogy, and metaphor, a reader often uses personal experience to infer broader context. For machines, matters are additionally complicated by use of informal vocabulary, grammar, and spelling. Furthermore, training data is often expensive or difficult to collect in bulk. These challenges motivated our efforts to seek transfer learning of broad world knowledge through feature pre-training using large unlabelled datasets.

- **Methodology**

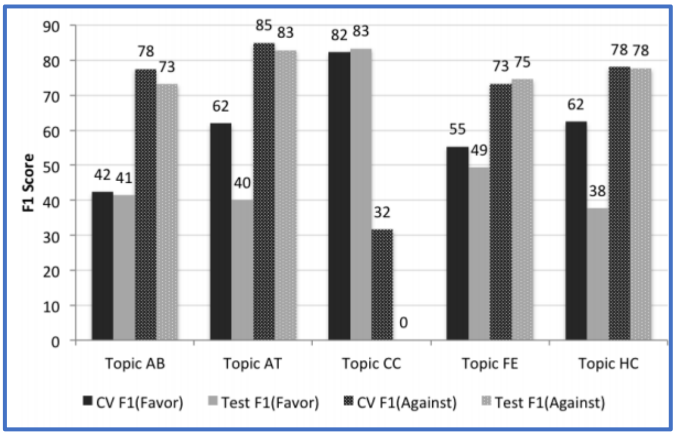
  A recurrent neural network initialized with features learned via distant supervision on two large unlabelled datasets was employed. We trained embeddings of words and phrases with the word2vec skip-gram method, then used those features to learn sentence representations via a hashtag prediction auxiliary task. These sentence vectors were then finetuned for stance detection on several hundred labelled examples.

- **Dataset**

  This approach was able to maximize the value of limited training data by transferring features from other systems trained on large, unlabelled datasets. Though, transfer learning does not completely eliminate the need for labelled in-domain training data.

- **Result**

  This submission achieved an average F1 score of 67.8 on the FAVOR and AGAINST classes of the held-out test set, which contained tweets from all five topics. This was the top scoring system among the 19 entries submitted to the supervised stance detection shared task.



- **Novelty**

  In stance detection we attempt to measure how an author's opinion is expressed in spontaneous, unstructured messages rather than the explicit prompts of formal opinion polls. Declarations of stance are often couched in figurative language that can be difficult for machines to unravel. To accomplish this challenging task a transfer learning-based approach was considered given that datasets are expensive and tough to find.

- **Strength**

  1. Transfer learning helped reduce the need of large labelled dataset.

- **Weakness**

  1. Model showed little amount of overfitting.
  2. Scores varied over across topics and classes.

Paper 4 – Adversarial Domain Adaptation for Stance Detection.
Authors – Brian Xu, Mitra Mohtarami, James Glass
Publication Year – 2019

- **Problem statement –** Aim is to predict the perspective (or stance) of a given document with respect to a given claim.

- **Motivation -** With the rise of social media and microblogs, there has been an increasing awareness of the negative influence of fake news and how it can unfairly influence public opinion on various events and policies. In order to counteract these effects, various organizations are now performing manual fact checking on suspicious claims. However, manual fact checking can't feasibly keep up with the sheer volume of fake claims. A fact-checking process for a given claim is a challenging multi-step process. One of the steps of the fact-checking process which is the stance detection task is the focus here. Stance detection aims to automatically determine the perspective (or stance) of a document to a claim as agree, disagree, discuss, or unrelated.

- **Methodology**

  Previously-proposed approaches for stance detection generally contain two components: a feature extraction component followed by a class label prediction component. In this, a model is presented for stance detection that augments the traditional models with a third component: a domain adaptation component. The domain adaptation component uses adversarial learning to encourage the feature extraction component to select common–rather than domain-specific–features when input data is from multiple different domains. This allows the model to better leverage source domain data for better prediction on data from the target domain.

- **Dataset**

  The Fake News Challenge (FNC) dataset is used as target dataset. This data is collected from a variety of sources such as rumour sites, e.g. snopes.com, and Twitter accounts. It contains around 50K claim-document pairs as training data with an imbalanced distribution over stance labels: 73% (unrelated), 18% (discuss), 7.3%(agree), 1.7%(disagree). Furthermore, the Fact Extraction and VERification (FEVER) dataset is used as source data. This dataset is collected from Wikipedia and contains around 145K claim-document pairs as training data with imbalanced distribution over stance labels: 55% (supported), 21% (refuted), 24% (Not Enough Information (NEI)).

- **Result**

  The hierarchy scheme can help the models to perform better across all the metrics. The best model is the combination of BOW, CNN+DA, and the hierarchy scheme. It outperforms the baselines on F1, especially on the most important classes: disagree, agree. The source FEVER data can improve the performance of our model for target FNC data through adversarial domain adaptation, when uses CNN model with hierarchy scheme or BOW+CNN model.

Model	Train Data	Weighted-Acc.	Acc.	Macro-F_1	F_1 (*agree, disagree, discuss, unrelated*)
BOW + CNN	FNC, FEVER	71.7	84.5	51.5	44.6 / 5.6 / 60.2 / 95.6
BOW + [CNN + DA]	FNC, FEVER	71.9	84.6	51.4	44.9 / 4.4 / 60.6 / 95.6
BOW + CNN (hierarchy)	FNC, FEVER	79.6	87.8	56.6	53.1 / 5.1 / 70.6 / 97.7
BOW + [CNN + DA] (hierarchy)	FNC, FEVER	**80.3**	**88.2**	**60.0**	**54.6 / 15.1 / 72.6 / 97.7**

- **Novelty**

  As annotating stances in different domains is a tedious and costly task, automatic methods based on machine learning are viable alternatives. In this paper, we focus on adversarial domain adaptation for stance detection where we assume there exists sufficient labelled data in the source domain and limited labelled data in the target domain. Extensive experiments on publicly available datasets showed the effectiveness of our domain adaption model in transferring knowledge for accurate stance detection across domains.

- **Strength**

  Domain adaption helps to improve the accuracies across datasets for the stance detection task.

- **Weakness**

  Needs validation on more datasets and results need to be verified and checked at each step.

Paper 5 – Towards Automatic Identification of Fake News: Headline-Article Stance Detection with LSTM Attention Models
Authors – Sahil Chopra, Saachi Jain
Publication Year – 2019

- **Problem statement -** We approach the problem of fake news via stance detection. Given an article as" ground truth", we attempt to classify whether a headline discusses, agrees, disagrees, or is unrelated to a given article.

- **Motivation -** In the wake of the 2016 Presidential Election, fake news has been a subject of increased discussion and debate. Accurate detection of fake news will allow for the elimination of deliberately deceptive news content, which will in turn promote a better-informed general public. As a result, there has been newfound interest in developing autonomous systems to identify fake news.

- **Methodology**

  Because unrelated samples comprised of over 73% of the data-set, baseline classifiers struggled to predict classes beyond the majority set, thus failing to capture the semantic differences between agree, disagree, and discuss. To address this issue, the problem is split into two more specific subproblems. In the first, it is tried to detect whether a headline and article are related, combining the agree, disagree, and discuss samples into an aggregate class related. In the second problem, given pairs that are already classified as related, we seek to label the pairs as agree, disagree, or discuss. We trained the two models separately on the train data, where the second problem is only trained on related samples from the training set. To produce the final predictions for the test set, we first feed the data to subproblem 1's model to filter out unrelated samples. We then send the remaining samples into the second model for further classification. The model involves two separate LSTMs, one for the headline and one for the article. The headline is fed through the first LSTM to extract the final hidden vector. This hidden state is used to initialize the LSTM of the article, thus "conditioning" the article LSTM on the headline.

- **Dataset**

  Here dataset used is the FNC-1 Dataset consisting of 1648 distinct headlines, 1683 distinct articles, and 49972 distinct headline-article pairings. The headlines had various lengths ranging from 10 to 220 words, while articles had lengths ranging from 25 to 5000 words. Additionally, The FNC-1 Dataset was very heavily biased towards unrelated headline-article pairs.

- **Result**

Model	F1 Score	S_2
Conditionally Encoded (CE) LSTM	.730	.7859
CE LSTM with Headline-to-Article Global Attention	.753	.8144
CE LSTM with Headline-to-Article Word-by-Word Attention	.768	.8263
CE LSTM with Bidirectional Global Attention	**.777**	.8324
Bidirectional CE LSTM with Bidirectional Global Attention (BiCE LSTM BiGA)	.761	**.8507**
5-Layer Bidirectional CE LSTM with Bidirectional Global Attention	.761	.8209
Bilateral Multi-Perspective Matching (Full, Maxpool Matching)	.760	.819

  When the entire pipeline is run, i.e. running the linear SVM classifier and feeding the headline-article pairs that had been classified as related into our neural networks, model performed well. Our BiCE LSTM BiGA scored SF NC of 0.8658, while our Bilateral Multi-Perspective Matching Model scored SF NC of 0.8501.

Model	S_{FNC}
Bidirectional CE LSTM with Bidirectional Global Attention (BiCE LSTM BiGA)	**.8658**
Bilateral Multi-Perspective Matching (Full, Maxpool Matching)	.8501

- **Novelty**

  Solutions to this stance detection problem will be leveraged as filters to limit the number of articles the fact checkers will have to examine by hand.

- **Strength –** Currently best performing model on the FNC-1 dataset.

- **Weakness** – Not tested on actual test data yet and all four attention layers aren't completely utilized yet.

Paper 6 – Stance Detection with Hierarchical Attention Network.
Authors – Qingying Sun, Zhongqing Wang, Qiaoming Zhu and Guodong Zhou
Publication Year – 2019

- **Problem statement –** Aim is to focus on stance detection task considering the impacts of different linguistic features.

- **Motivation -** Most of the previous works, model the sequence of words to learn document representation. However, much linguistic information, such as polarity and arguments of the document, is correlated with the stance of the document, and can inspire us to explore the stance. In addition, since the influences of different linguistic information are different, we propose a hierarchical attention network to weigh the importance of various linguistic information, and learn the mutual attention between the document and the linguistic information.

- **Methodology**

  A hierarchical attention model is proposed, which stands for the mutual attention between the document and the linguistic factors. The model contains two parts: linguistic attention part and hyper attention part. The former helps learn flexible and adequate document representation with different linguistic feature set, and the latter helps adjust the weight of different feature sets.

- **Dataset**

  Two datasets are used to evaluate the performance of the proposed system. H&N14 is collected by Hasan and Ng, and SemEval16 is from SemEval-2016 Share Task. H&N14 is collected from an English online debate forum with four targets: “Abortion”, “Gay Rights”, “Obama”, and “Marijuana”. SemEval16 is the dataset for stance detection from English tweets, and each tweet corresponds to a special target: “Atheism”, “Climate Change is a Real Concern” (“Climate”), “Feminist Movement” (“Feminist”), “Hillary Clinton” (“Hillary”), and “Legalization of Abortion” (“Abortion”).

- **Result**

  HAN is the proposed model, which uses Hierarchical Attention Neural (HAN) model to learn the mutual attention between document and linguistic information.

Model	Abortion	GayRights	Obama	Marijuana	$MacF_{avg}$	$MicF_{avg}$
SVM	59.48	**59.52**	63.02	55.02	59.26	60.52
LSTM	60.72	56.07	60.14	55.58	58.13	59.45
TAN	**63.96**	58.13	63.00	56.88	60.49	62.35
HAN	63.66	57.36	**65.67**	**62.03**	**62.18**	**63.25**

Table 3: Comparison with baselines on H&N14 dataset.

Model	Atheism	Climate	Feminism	Hillary	Abortion	$MacF_{avg}$	$MicF_{avg}$
SVM	62.16	42.91	56.43	55.68	60.38	55.51	67.01
LSTM	58.18	40.05	49.06	61.84	51.03	52.03	63.21
TAN	59.33	**53.59**	55.77	**65.38**	63.72	59.56	68.79
HAN	**70.53**	49.56	**57.50**	61.23	**66.16**	**61.00**	**69.79**

Table 4: Comparison with baselines on SemEval16 dataset.

- **Novelty**

  Stance detection system can potentially have a positive social impact and are of practical interest to non-profits, governmental organizations, and companies.

- **Strength**

  1. An attention model is used to measure the importance of different linguistic features, and learn the mutual attention between the document and the linguistic information.
  2. Model achieves better performance compared with the state-of-the-art models on two datasets.

- **Weakness**

  1. Accuracies are comparatively low.

Paper 7 – Multi-Target Stance Detection via a Dynamic Memory-Augmented Network.
Authors – Penghui Wei, Junjie Lin, and Wenji Mao
Publication Year – 2018

- **Problem statement -** Multi-target stance detection, in contrast to general stance detection, aims at jointly detecting stances towards multiple related targets.

- **Motivation –** Most of the existing studies consider different target entities separately. However, in many scenarios, stance targets are closely related, such as several candidates in a general election and different brands of the same product. As stance expression regarding a target can provide additional information to help identify the stances towards other related targets, modelling expressions regarding multiple targets jointly is beneficial for improving the overall performance compared to single-target scheme.

- **Methodology**

  The problem of multi-target stance detection is formulized as follows, given a tweet and a set of K pre-chosen targets, predict the stances towards these targets. We denote the tweet as a word sequence. Stance labels include FAVOR, AGAINST and NEITHER. The overall architecture of the model, DMAN consisting of three modules: (1) Representation module, learning target-specific tweet representations using attentive network; (2) Memory module, equipping a shared external memory to capture and store stance-indicative information for enhancing tweet representations; (3) Classification module, classifying stance towards each target.

- **Dataset**

  Experiments are conducted on a benchmark Twitter dataset. Four candidates of the 2016 US election, i.e., Donald Trump, Hillary Clinton, Ted Cruz and Bernie Sanders, are the pre-defined targets. Each tweet is annotated with two stance labels regarding two targets (called a target pair). There are three target pairs in this dataset.

Table 1: Statistics of instances in the dataset.

Target Pair	#total	#train	#dev	#test
Trump-Clinton	1722	1240	177	355
Trump-Cruz	1317	922	132	263
Clinton-Sanders	1366	957	137	272
Total	4455	3119	446	890

- **Result**

Performance Comparison of Different Models

Method	Target pair			Overall
	Trump Clinton	Trump Cruz	Clinton Sanders	
Seq2Seq	56.60[1]	53.12[1]	54.72[1]	54.81
DMAN	**60.30**	**53.64**	**56.25**	**56.73**
−F	57.94	53.12	54.38	55.15
−DM	55.99	53.22	54.50	54.57
−MT	59.07	52.02	52.90	54.67
−DM, −MT	55.80	51.59	52.26	53.22
−DM, −MT, −ATT	53.55	51.17	46.83	50.52

- **Novelty**

  Automatically identifying user stances by analysing the online contents they posted has widespread applications in real-life scenarios. For instance, it is beneficial for governments to understand public reactions about policy changes. With the booming of social media, stance analysis in social networking platforms like Twitter has become an important research topic in opinion mining.

- **Strength**

  1. DMAN learns target-specific representations by target-based attentive network, and then utilizes an external memory to capture and store stance-indicative information of multiple targets for effective stance detection.
  2. Experimental results show effectiveness of the model.

- **Weakness** – Highly dependent on a accurate multi-target labelled dataset.

Paper 8 – Stance Classification with Target-Specific Neural Attention Networks.
Authors – Jiachen Du, Ruifeng Xu, Yulan He, Lin Gui
Publication Year – 2019

- **Problem statement –** Incorporating a target-specific information into stance classification by following a novel attention mechanism

- **Motivation**

  With the rapidly development of Internet and Web 2.0, more and more people post their opinions online. To detect sentiment or retrieve opinions from online text, sentiment analysis and opinion mining has become a hot research topic in natural language processing. Various techniques have been proposed to identify the polarity of a given text. However, in many practical applications, we are interested to learn the position of an author to a specific target or topic rather than the polarity of the whole text. For example, during the US general election, we would like to find out from someone's online posts whether she supports Trump or not. This is referred to as target-specific stance detection.

- **Methodology**

  The performance of stance detection could be potentially improved by considering both text content features and target related features. Motivated by this, an RNN-based model is proposed which can concentrate on salient parts in text corresponding to a given target, named as Target-specific Attention Neural Network (TAN). It consists of two main components: a recurrent neural network (RNN) as the feature extractor for text and a fully-connected network as the target-specific attention selector. These two components are combined by an element-wise multiplication operation in the classification layer.

- **Dataset**

  Semeval-2016 Task 6 released the first dataset for stance detection from English tweets. In this dataset, more than 4,000 tweets are annotated for whether one can deduce favourable or unfavourable stance towards one of five targets "Atheism", "Climate Change is a Real Concern", "Feminist Movement", "Hillary Clinton", and "Legalization of Abortion". Task has two subtasks including subtask-A supervised learning and subtask-B unsupervised learning. In this evaluation, we only use the dataset of subtask-A in which the targets provided in the test set can all be found in the training set.

- **Result**

Target	NBOW	LSTM	$LSTM_E$	TOP	TAN
E_1	55.12	58.18	59.77	***61.47***	59.33
E_2	39.93	40.05	48.98	41.63	***53.59***
E_3	50.21	49.06	52.04	***62.09***	55.77
E_4	55.98	61.84	56.89	57.67	***65.38***
E_5	55.07	51.03	60.34	57.28	***63.72***
Overall	60.19	63.21	66.24	67.82	***68.79***

Performance comparison of different models

  It is observed that this model performs the best among all methods. In specific, TAN outperforms the first-place system of NLPCC-2016's by 1.8%. The top performing system of NLPCC is a relatively strong baseline that used carefully chosen hand-crafted features and optimal parameters tuned by grid search. Further analysis demonstrate that TAN is a language-independent model that perform consistently well across different languages.

- **Novelty**

  The main contribution of this model is to learn target-augmented embeddings for text and use attention mechanism to extract target-specific parts in text to improve classification performance.

- **Strength**

  The impressive capability of the model to extract the important parts which are helpful to improve stance detection.

- **Weakness**

  Transfer learning is not used, thus limiting the performance of the model.