Improving Social Emotion Prediction with Reader Comments Integration

Abdullah Alsaedi[1], Phillip Brooker[2], Floriana Grasso[1] and Stuart Thomason[1]
[1]*Department of Computer Science, University of Liverpool, UK*
[2]*Department of Sociology, Social Policy and Criminology, University of Liverpool, UK*
{*A.Alsaedi, P.D.Brooker, F.Grasso, S.Thomason*}*@liverpool.ac.uk*

Keywords: Social emotion prediction, emotion analysis

Abstract: Social emotion prediction is concerned with the prediction of the reader's emotion when exposed to a text. In this paper, we propose a comment integration method for social emotion prediction. The basic intuition is that enriching social media posts with related comments can enhance the models' ability to capture the conversation context, and hence improve the performance of social emotion prediction. We developed three models that use the comment integration method with different approaches: word-based, topic-based, and deep learning-based. Results show that our proposed models outperform popular models in terms of accuracy and F1-score.

1 INTRODUCTION

In recent years, social media platforms have provided an open environment for data scientists to explore and analyze users behavior (Ruths and Pfeffer, 2014). Many studies have been conducted on social media platforms such as Facebook, Twitter, Reddit, etc. that aim to identify the sentiment of users toward a certain event, company, or product (Fan and Gordon, 2014). Research into sentiment analysis has progressed rapidly in recent years, with some remarkable results. However, traditionally, most of the proposals made in this field use a too coarse-grained approach in which emotions are classified as positive, negative, or neutral. Some recent research has emerged attempting to integrate this with emotion analysis, to obtain a deeper understanding of the text based on predefined emotion models (Lerner and Keltner, 2000).

Unlike sentiment analysis, the analysis of emotion can be conducted for both subjective and objective text (Rao et al., 2012). Emotion analysis also deals with text from different perspectives: the writer's perspective, which represents the expressed emotion in the text; and the reader's perspective, which shows the emotions the text provokes in its readers (Lin et al., 2007). However, most existing work focuses on identifying the emotion from the writer's perspective (Guan et al., 2019).

Social emotion prediction is the consensus term to designate work on reader's emotion prediction, that is research concerned with the problem of predicting the emotion provoked to the reader after being exposed to the text. Predicting the social emotion is a challenging task as the reader's emotion is not declared in the text, but it is triggered by reading the text, and it is likely influenced by many factors, such as the readers' personal background and experiences.

One common way to approach the problem is to understand whether there is a relationship between reader's and writer's emotion, for instance by establishing if the reader experiences the same emotion that the writer is portraying in the text. This can be done by considering, as a single unit, a piece of text and any comments to that text. For example, in social media, one would consider posts followed by their related comments. The assumption is that the readers who wrote comments can be assumed to have been affected by the content of the post. Consequently, the written emotions in the comments could reflect the readers' emotion at reading the text. Previous studies show indeed a correlation between the emotions of the reader and writer of a comment (Yang et al., 2009; Liu et al., 2013).

Such comments integration approach was found to be effective for short-text topic modeling (Alvarez-Melis and Saveski, 2016) based on underlying assumptions on topic consistency between posts and comments. As the social emotion can be related to the topic of the post, the assumption is that the content of the comments could be informative and useful for social emotion prediction models to learn from, even without knowing the emotions experienced by the writers of the comments.

In this paper we propose three models for social emotion prediction using the comment integration approach. We integrate each post with its corresponding comments as a whole document, then train our models on these integrated documents of posts and comments. We report on experiments that show that, even with a basic architecture, by integrating the comments we can outperform established social emotion prediction models.

The rest of the paper is structured as follows. Section 2 explores related works. Section 3 presents our comments integration models. Section 4 discusses the experiments and analysis, and Section 5 presents the conclusion of our work.

2 RELATED WORK

In this section, we summarize related work in two areas: social emotion prediction and bi-perspective emotion analysis.

2.1 Social emotion prediction

A shared task on Affective Text, proposed at the 2007 SemEval workshop (Strapparava and Mihalcea, 2007), served as starting point for the development of several systems for social emotion prediction using various approaches, such as word-based methods, topic-based methods, and deep learning-based methods (Alsaedi et al., 2021). In this section we summarise the more relevant ones to the scope of this paper.

Word-based systems assume that all words, even neutral ones, can be associated with a likelihood to provoke emotions (Strapparava and Mihalcea, 2008). SWAT (Katz et al., 2007), one of the top-performing word-based systems, uses a bag-of-words (BOW) model trained to label news headlines with social emotions. It utilizes Roget's Thesaurus for synonym and antonym expansion and scores each word to determine the average score for the headline. Another interesting approach is the one presented in (Chaumartin, 2007). The authors propose UPAR7, a rule-based system, depends on a syntactic parser that utilises resources from WordNet, SentiWordNet, and WordNet-Affect for social emotion prediction.

Topic-based systems suggest that social emotion should be linked to topics rather than words (Bao et al., 2009). Therefore, topic-based systems such as Emotion-Topic Model (ETM)(Bao et al., 2009)(Bao et al., 2011), Affective Topic Model (ATM) (Rao et al., 2014b), Sentiment Latent Topic Model (SLTM) (Rao et al., 2014a), etc. benefit from the machinery of topic modeling techniques and introduce an intermediate emotion layer to popular topic models such as Latent Dirichlet Allocation (LDA) (Blei et al., 2003). Similarly, Multi-label Supervised Topic Model (MSTM) (Rao et al., 2014a) utilizes the Supervised Topic Model (STM) (Blei and McAuliffe, 2010) that extended to model multi-labels rather than single label to be suitable for social emotion prediction.

State-of-the-art systems usually depend on deep learning-based methods, such as neural networks and word-embeddings, to obtain better understanding of sequences and overcome the data sparsity problem that occurs with other approaches. For example, in (Guan et al., 2019), a hierarchical model based on long short-term memory (LSTM) was proposed. The model utilizes the attention mechanism and attempts to capture the semantics of long texts using three different levels of embeddings: words, sentences and documents. Similarly, TESAN (Wang and Wang, 2020) combines semantic and topical features and feeds them into a unified deep learning model. The model consists of a neural topic model that learns the topical embeddings of documents, and a topic-enhanced self-attention mechanism to generate the document vector from the semantic and topical features. Both features are integrated in a final gate, achieving an improvement in performance.

2.2 Bi-perspective emotion analysis

Another set of works which are relevant to our research focus on the analysis of reader's emotion when compared and contrasted to the writer's emotion.

In (Yang et al., 2009) the readers' emotion from Yahoo!Kimo news, the Taiwan branch of Yahoo!, was analyzed in comparison to the writers' emotion from the Yahoo!Kimo blog corpus. The authors built a reader's emotion classifier trained on the news corpus, and applied it to the blog corpus, then analyzed the new corpus, annotated with both perspectives. They found that the valence, which is the degree of pleasantness, reveals that in blog topics, readers' and writers' emotions tend to agree on their polarities. However, the degree of influence is affected by the topic.

Tang and Chen (Tang and Chen, 2011) used data from the micro-blogging platform Plurk (https://www.plurk.com/portal/) that provides emotion tagging from both perspectives for posts and their corresponding comments. Users in Plurk have the ability to tag their own posts with emotion, hence self reporting the writer's emotion. Users replying to a post can also tagged their own replied, hence self reporting the reader's emotion. When analysing both perspectives, they conclude that predicting the reader's emotion is clearly more challenging than predicting the writer's emotion. The authors extended their work (Tang and Chen, 2012) by studying the emotion transition between writer and reader in the Plurk platform, and analyzed the linguistic features that signal the change of emotions between the two. They also proposed models to predict the transition and they suggested sentiment word mining as a useful tool for prediction performance.

In (Liu et al., 2013), the authors studied the relationship between news and their comments, again on thee Yahoo!Kimo News. As the reader's emotion is provided by Yahoo!Kimo News, they manually annotated the comments with the writer's emotion. When comparing, the news reader's emotion and the comments writer's emotion, they found that both sets of emotions are strongly correlated in term of valence, but interestingly, not in fine-grained emotions such as happiness, anger, sadness, etc.

In (Buechel and Hahn, 2017), the authors examined the data annotation process for writer's emotion and reader's emotion, and they interestingly propose a third perspective in the process of analysis, that is the emotion from the point of view "of the text itself". They conducted experiments on two English corpora and suggest that the quality of the writer's emotion annotation was the highest. The work is notable and it is accompanied by a dataset, EmoBank, a large-scale bi-perspective emotion annotated corpus of 10k sentences, covering different genres.

Another important dataset is GoodNewsEveryone (Bostan et al., 2019), a corpus of English news headlines annotated with writer's emotion, writer's emotion intensity, reader's emotion and semantic roles according to the FrameNet (Baker et al., 1998) semantic frame. The dataset consists of 5000 English headlines annotated via crowdsourcing.

3 COMMENT INTEGRATION MODELS

In this section we present our models for social emotion prediction. We use for all models the comment integration approach, and we merge texts with the comments attached to that text in one single document. We conducted a number of experiments by varying different features, and we report the combination of features producing the best results for each of the three basic approaches: word-based, topic-based and deep learning based.

3.1 Comments integration word-based model

Word-based methods aim to detect the emotion in content through finding the reader's emotional trigger, as it is noticed that any words, even the neutral ones, can provoke emotions. Our simplest model combine posts and their comments and used a basic Naïve-Bayes classifier trained on the BOW representation of the documents resulting from the combination. A social emotion for the post is then attached to the whole document.

Figure 1 shows the architecture of the proposed word-based model.

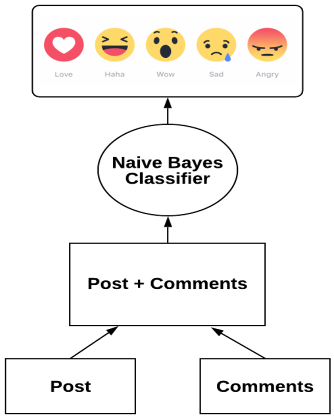

Figure 1: Architecture of the comments integration word-based model.

Results will be discussed in Section 4, but we anticipate here that, compared to the Naïve-Bayes model trained on posts content only, our model improved the performance in terms of accuracy, precision and recall. The model also achieved a significant improvement compared to other popular models, including topic-based models.

3.2 Comments integration topic-based model

In this model we used the merged document of post and related comments as an input to an LDA model to extract the topics distribution for each document. We used EmoLex (Mohammad and Turney, 2013), which is an emotion lexicon annotated with classes from Plutchick's emotion model (Plutchik, 1980). EmoLex was used to extract the emotion-word frequency in each document in order to produce a vector of emotion frequency. Both the LDA output and the emotion frequency vector were then used as an input to a maximum entropy classifier. Figure 2 shows the architecture of the proposed topic-based model.

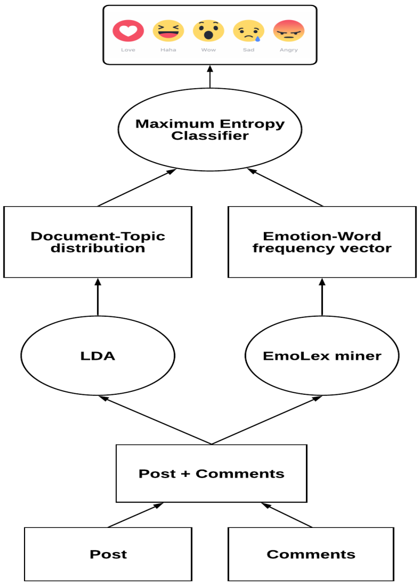

Figure 2: Architecture of the comments integration topic-based model.

3.3 Comments integration deep learning-based model

The deep learning model utilizes a Term Frequency–Inverse Document Frequency (TF-IDF)(Salton and Buckley, 1988) representation that gives more importance to words that are more relevant to the document. We found that the use of a TF-IDF representation of the merged document of posts and comments improved the performance significantly. The model consists of a multi-layer neural network with one hidden layer and a softmax activation in the output layer. Figure 3 shows the architecture of the proposed deep learning-based model.

4 EXPERIMENTS

4.1 Dataset

The evaluation was conducted using FacebookR (Krebs et al., 2018), a dataset of Facebook posts with their social emotion and comments. Facebook introduced the "reactions" feature in 2016 to enable its users to express their emotion toward posts, representing their social emotion. The reactions set includes

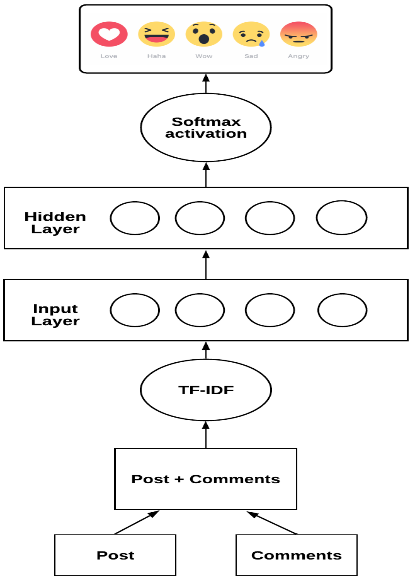

Figure 3: Architecture of the comments integration deep learning-based model.

Like, Love, Care, HAHA, Wow, Sad and Angry.

FacebookR is scraped from customer service pages of popular supermarkets from the United States and the United Kingdom. To the best of our knowledge, this is the only available English dataset that provides comments as well as posts with the social emotion labels. The dataset consists of over 70,000 posts. Of these, we considered those that excluded the "like" reaction. As pointed out in (Krebs et al., 2018), the like reaction makes the dataset very imbalanced and inaccurate, as Facebook users tend to use "like" for any positive emotion or, sometimes, to show that they have read the post. The resulting dataset when excluding the "like" reaction consists of 8103 posts.

As manual labeling for social emotion is challenging and usually produces low quality labels, FacebookR is scraped from real data that is annotated by Facebook users themselves. Also, based on statistical analysis conducted by the dataset publishers, user labeling seems consistent and the reactions that appear together tend to agree in a high degree. We use the dataset to predict the top social emotion for posts. Table 1 shows the number of posts with the reaction of the top number of votes in the dataset.

There are on average six comments for each post and among 8103 posts only 150 are without any comment. Table 2 shows the comments statistics for FacebookR and Table 3 shows the documents length before and after merging the posts with their comments.

Top reaction	Number of posts
ANGRY	2276
HAHA	2253
LOVE	1648
WOW	1237
SAD	686

Table 1: The number of posts with their top social emotion.

Measure	Number of comments
mean	6.31
std	37.56
min	1
max	2605

Table 2: Comments statistics for FacebookR posts.

Number of words	Posts	Posts merged with comments
mean	111.66	274.62
std	125.23	624.15
min	1	2
max	2093	38004

Table 3: Number of words in posts versus posts merged with their comments.

4.2 Baselines

We compare our models with the following popular models from the literature:

- SVM: Support vector machine classifier.
- NB: Naive Bayes classifier.
- SWAT: One of the best performing models proposed in SemEval-2007 task 14, which depends on scoring the document words.
- ET (Bao et al., 2011): A straightforward model that depends on a Naive Bayes classifier to learn the word-emotion associations.
- ETM: The first topic-based model that introduced an emotion layer to LDA for social emotion prediction.
- ATM: A topic-based model that generates social emotion lexicon and predicts unlabelled documents.
- MSTM: A supervised topic model that learns the association of word and topic, then predicts the emotions from each topic.
- SLTM: A topic-based model that associates words and emotions to topics to predict the social emotion from unseen documents.

In addition, we perform a comparison with the following deep learning models:

- LSTM (Hochreiter and Schmidhuber, 1997): A long short-term memory model, which is a popular type of recurrent neural network, usually used for learning from ordered sequences such as sentences.
- AttBiLSTM (Wang and Yang, 2020): A widely-used bidirectional long short-term memory network with an attention mechanism.

4.3 Experimental design

We arbitrarily set the number of topics K to 20 for the following models: ETM, ATM, MSTM, SLTM and our topic-based model. For LSTM, AttBiLSTM, and our deep learning model we used the Adam optimizer (Kingma and Ba, 2015) for training with a learning rate of 0.001, batch size 100, 20 epochs, and early stopping on validation loss to prevent overfitting (SJÖBERG and LJUNG, 1995). Both LSTM and AttBiLSTM consist of one layer with 32 units, and use a pre-trained glove (Pennington et al., 2014) word-embedding with 200-dimensions trained on a Twitter dataset as the language on Twitter and Facebook share similar characteristics.

4.4 Evaluation metrics

We use $Acc@1$ as a metric for model evaluation in our experiments, as it has been proved to represent the correct prediction for the social emotion with the highest number of ratings and is considered the most important metric according to (Bao et al., 2009; Bao et al., 2011).

To determine $Acc@1$, the correctly predicted documents are counted against the whole number of documents as follows:

$$Acc_d@1 \begin{cases} 1 & \text{if } e_{pr} = EM_{top} \\ 0 & \text{otherwise} \end{cases}$$

$$Acc@1 = \frac{\sum_{d \in D} Acc_d@1}{|D|}$$

Put simply, if there are many reactions for a post, we will consider the reaction with the top ratings as the social emotion of that post. If two or more reactions have the same number of ratings we will consider either of them.

4.5 Prediction performance

Here we evaluate our models' ability to predict each social emotion by measuring the precision, recall, and

F1-score. Table 4 shows the evaluation for our word-based model, topic-based model, and deep learning-based model.

Model	Emotion	Precision	Recall	F1
Word-based	ANGRY	0.47	0.83	0.60
	HAHA	0.56	0.55	0.56
	LOVE	0.74	0.60	0.66
	SAD	0.15	0.02	0.03
	WOW	0.32	0.07	0.11
Topic-based	ANGRY	0.45	0.67	0.54
	HAHA	0.45	0.62	0.52
	LOVE	0.59	0.56	0.58
	SAD	0.00	0.00	0.00
	WOW	0.19	0.01	0.02
Deep Learning-based	ANGRY	0.50	0.77	0.61
	HAHA	0.51	0.75	0.61
	LOVE	0.82	0.61	0.70
	SAD	0.00	0.00	0.00
	WOW	0.33	0.00	0.01

Table 4: Evaluation of the proposed models.

We can notice that the models performances vary depending on the predicted emotion. As the dataset is unbalanced, the prediction for some emotions is affected negatively, such as the SAD and WOW emotions, especially for the topic-based and the deep learning-based models. Interestingly, the F1 for the LOVE emotion was always the highest, even though the number of training samples for the ANGRY and HAHA are higher which needs further investigation.

4.6 Comparison with baselines

In this subsection we compare our social emotion prediction models to the baselines mentioned in 4.2. All models were evaluated on the FacebookR dataset and two evaluation metrics have been used in the comparison: Accuracy of the top social emotion (*Acc@*1), and F1-score. We focus on comparing models with similar methods and divide the models into word-based models, topic-based models, and deep learning-based models. Table 5 summarises the performance of all models.

The table shows that our proposed models perform best among models with the same method in term of accuracy, and also in F1 except for the topic-based models, where ETM is better and achieved a comparable accuracy to our topic-based model. Our deep learning model outperforms all of other models in terms of accuracy and F1. The word-based model outperforms SWAT and ET both in accuracy and F1. However, the ET model performs much better than SWAT, which has a very low F1. Interestingly, the performance of our word-based model was better than all of the topic-based models, including our model, and that might be related to the topics distribution the extent of overlapping topics in our dataset.

When comparing the topic-based models we can see that their performances vary significantly. ATM, SLTM, and MSTM accuracies are low compared to ETM and our model. However, with regard to F1 score, the differences between them are much bigger since ATM, SLTM, and MSTM have very low F1.

Deep learning models provide a relatively close accuracy for F1. There are no significant differences between the accuracy and F1 score, which sometimes appears in some of the word-based and topic-based models. However, our deep learning-based model boosts the accuracy by 10% compared to the LSTM model. The F1 score also increased by 0.08. The AttBiLSTM model improved the performance by 2% over the LSTM model the F1 score raised by 0.02.

All of the deep learning models were learning quickly before suffering from overfitting after the first few epochs. We tried to avoid that by applying the early stopping technique to force the training to stop when there was no improvement. Figure 4 presents the accuracy of all deep learning models over 20 epochs. It is noteworthy that the models' accuracy began to reduce after the first few epochs, which is a sign of overfitting.

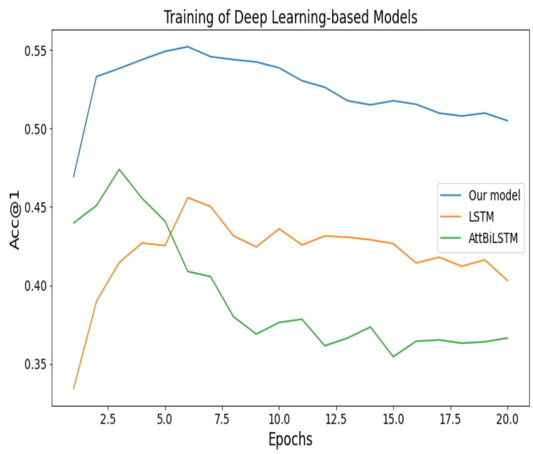

Figure 4: Accuracy of deep-based models with different number of epochs.

5 CONCLUSION

In this paper, we have proposed a comment integration method for social emotion prediction. We developed word-based, topic-based, and deep learning-based models that use our proposed methods and

Method	Model	Accuracy	F1
Word-based	SVM	48.41	0.44
	NB	50.08	0.46
	SWAT	18.73	0.06
	ET	43.41	0.34
	Our Word-based model	**52.59**	**0.48**
Topic-based	ETM	46.21	**0.44**
	ATM	21.16	0.06
	SLTM	27.43	0.22
	MSTM	19.40	0.06
	Our Topic-based model	**48.12**	0.42
Deep Learning-based	LSTM	45.34	0.41
	AttBiLSTM	47.22	0.43
	Our Deep learning-based model	**55.55**	**0.49**

Table 5: A summary of models' performances in terms of accuracy and F1.

compared them to popular social emotion prediction methods. Experiments show that models that use the comments integration method outperform popular models in terms of $Acc@1$ and F1. We found that merging social media comments with their related posts added valuable data about the readers' emotions and enhanced the ability to predict the social emotion for posts. In the future, we will utilize the comment integration methods to improve the models from two aspects. On the one hand, we will develop more advanced models that depend on deep learning-based methods and neural topic models. On the other hand, we will attempt to improve the prediction performance for emotions with low precision and recall.

REFERENCES

Alsaedi, A., Brooker, P., Grasso, F., and Thomason, S. (2021). A survey of social emotion prediction methods. In *Proceedings of the 10th International Conference on Data Science, Technology and Applications*. SCITEPRESS-Science and Technology Publications.

Alvarez-Melis, D. and Saveski, M. (2016). Topic Modeling in Twitter: Aggregating Tweets by Conversations. In *Proceedings of the Tenth International AAAI Conference onWeb and Social Media*.

Baker, C. F., Fillmore, C. J., and Lowe, J. B. (1998). The berkeley framenet project. In *36th Annual Meeting of the Association for Computational Linguistics and 17th International Conference on Computational Linguistics, Volume 1*, pages 86–90.

Bao, S., Xu, S., Zhang, L., Yan, R., Su, Z., Han, D., and Yu, Y. (2009). Joint emotion-topic modeling for social affective text mining. In *Proceedings - IEEE International Conference on Data Mining, ICDM*, pages 699–704. IEEE.

Bao, S., Xu, S., Zhang, L., Yan, R., Su, Z., Han, D., and Yu, Y. (2011). Mining Social Emotions from Affective Text. *IEEE Transactions on Knowledge and Data Engineering*, 24(9):1658–1670.

Blei, D. M. and McAuliffe, J. D. (2010). Supervised topic models. *arXiv preprint arXiv:1003.0783*.

Blei, D. M., Ng, A. Y., and Jordan, M. I. (2003). Latent dirichlet allocation. *Journal of machine Learning research*, 3(Jan):993–1022.

Bostan, L., Kim, E., and Klinger, R. (2019). GoodNewsEveryone: A Corpus of News Headlines Annotated with Emotions, Semantic Roles, and Reader Perception. arXiv:1912.03184v3 [cs.CL].

Buechel, S. and Hahn, U. (2017). EMOBANK: Studying the impact of annotation perspective and representation format on dimensional emotion analysis. In *15th Conference of the European Chapter of the Association for Computational Linguistics, EACL 2017*, volume 2, pages 578–585.

Chaumartin, F. R. (2007). UPAR7: A knowledge-based system for headline sentiment tagging. In Agirre, E., Màrquez, L., and Wicentowski, R., editors, *Proceedings of the 4th International Workshop on Semantic Evaluations (SemEval-2007)*, pages 422–425. ACL.

Fan, W. and Gordon, M. D. (2014). The power of social media analytics. *Communications of the ACM*, 57(6):74–81.

Guan, X., Peng, Q., Li, X., and Zhu, Z. (2019). Social Emotion Prediction with Attention-based Hierarchical Neural Network. In *2019 IEEE 4th Advanced Information Technology, Electronic and Automation Control Conference (IAEAC)*, pages 1001–1005.

Hochreiter, S. and Schmidhuber, J. (1997). Long short-term memory. *Neural Comput.*, 9(8):1735–1780.

Katz, P., Singleton, M., and Wicentowski, R. (2007). SWAT-MP: The SemEval-2007 systems for task 5 and task 14. In *ACL 2007 - SemEval 2007 - Proceedings of the 4th International Workshop on Semantic Evaluations*, pages 308–313.

Kingma, D. P. and Ba, J. (2015). Adam: A Method for Stochastic Optimization. In Bengio, Y. and LeCun, Y., editors, *3rd International Conference on Learning Representations, ICLR 2015, San Diego, CA, USA, May 7-9, 2015, Conference Track Proceedings*.

Krebs, F., Lubascher, B., Moers, T., Schaap, P., and Spanakis, G. (2018). Social emotion mining techniques for facebook posts reaction prediction. *ICAART 2018 - Proceedings of the 10th International Conference on Agents and Artificial Intelligence*, 2:211–220.

Lerner, J. S. and Keltner, D. (2000). Beyond valence: Toward a model of emotion-specific influences on judgement and choice. *Cognition & emotion*, 14(4):473–493.

Lin, K. H.-Y., Yang, C., and Chen, H.-H. (2007). What emotions do news articles trigger in their readers? In *Proceedings of the 30th annual international ACM SIGIR conference on Research and development in information retrieval*, pages 733–734.

Liu, H., Li, S., Zhou, G., Huang, C. R., and Li, P. (2013). Joint modeling of news reader's and comment writer's emotions. In *ACL 2013 - 51st Annual Meeting of the Association for Computational Linguistics, Proceedings of the Conference*, volume 2, pages 511–515.

Mohammad, S. M. and Turney, P. D. (2013). Crowdsourcing a word-emotion association lexicon. *Computational Intelligence*, 29(3):436–465.

Pennington, J., Socher, R., and Manning, C. D. (2014). Glove: Global vectors for word representation. In *Proceedings of the 2014 conference on empirical methods in natural language processing (EMNLP)*, pages 1532–1543.

Plutchik, R. (1980). Psychoevolutionary Theory of Basic Emotions. *American Scientist*.

Rao, Y., Li, Q., Mao, X., and Wenyin, L. (2014a). Sentiment topic models for social emotion mining. *Information Sciences*, 266:90–100.

Rao, Y., Li, Q., Wenyin, L., Wu, Q., and Quan, X. (2014b). Affective topic model for social emotion detection. *Neural Networks*, 58:29–37.

Rao, Y., Quan, X., Wenyin, L., Li, Q., and Chen, M. (2012). Buildingword-emotion mapping dictionary for online news. In *SDAD@ ECML/PKDD*, pages 28–39.

Ruths, D. and Pfeffer, J. (2014). Social media for large studies of behavior. *Science*, 346(6213):1063–1064.

Salton, G. and Buckley, C. (1988). Term-weighting approaches in automatic text retrieval. *Information processing & management*, 24(5):513–523.

SJÖBERG, J. and LJUNG, L. (1995). Overtraining, regularization and searching for a minimum, with application to neural networks. *International Journal of Control*, 62(6):1391–1407.

Strapparava, C. and Mihalcea, R. (2007). SemEval-2007 task 14: Affective text. In *ACL 2007 - SemEval 2007 - Proceedings of the 4th International Workshop on Semantic Evaluations*.

Strapparava, C. and Mihalcea, R. (2008). Learning to identify emotions in text. In *Proceedings of the ACM Symposium on Applied Computing*, pages 1556–1560.

Tang, Y.-j. and Chen, H.-h. (2011). Emotion Modeling from Writer / Reader Perspectives Using a Microblog Dataset. In *Proceedings of the Workshop on Sentiment Analysis where AI meets Psychology (SAAIP)*, pages 11–19.

Tang, Y. J. and Chen, H. H. (2012). Mining sentiment words from microblogs for predicting writer-reader emotion transition. In *Proceedings of the 8th International Conference on Language Resources and Evaluation, LREC 2012*, pages 1226–1229.

Wang, C. and Wang, B. (2020). An End-to-end Topic-Enhanced Self-Attention Network for Social Emotion Classification. In *The Web Conference 2020 - Proceedings of the World Wide Web Conference, WWW 2020*, volume 2, pages 2210–2219.

Wang, Z. and Yang, B. (2020). Attention-based bidirectional long short-term memory networks for relation classification using knowledge distillation from bert. In *2020 IEEE Intl Conf on Dependable, Autonomic and Secure Computing, Intl Conf on Pervasive Intelligence and Computing, Intl Conf on Cloud and Big Data Computing, Intl Conf on Cyber Science and Technology Congress (DASC/PiCom/CBDCom/CyberSciTech)*, pages 562–568.

Yang, C., Lin, K. H. Y., and Chen, H. H. (2009). Writer meets reader: Emotion analysis of social media from both the writer's and reader's perspectives. In *Proceedings - 2009 IEEE/WIC/ACM International Conference on Web Intelligence, WI 2009*, volume 1, pages 287–290.