

Exploring Optimism and Pessimism in Twitter Using Deep Learning

Cornelia Caragea¹, Liviu P. Dinu², and Bogdan Dumitru²

¹Computer Science, Kansas State University
Manhattan, Kansas

²Faculty of Mathematics and Computer Science, University of Bucharest
Bucharest, Romania

ccaragea@ksu.edu, ldinu@fmi.unibuc.ro, bogdan27182@gmail.com

Abstract

Identifying optimistic and pessimistic viewpoints and users from Twitter is useful for providing better social support to those who need such support, and for minimizing the negative influence among users and maximizing the spread of positive attitudes and ideas. In this paper, we explore a range of deep learning models to predict optimism and pessimism in Twitter at both tweet and user level and show that these models substantially outperform traditional machine learning classifiers used in prior work. In addition, we show evidence that a sentiment classifier would not be sufficient for accurately predicting optimism and pessimism in Twitter. Last, we study the verb tense usage as well as the presence of polarity words in optimistic and pessimistic tweets.

1 Introduction

“You know, ever since we were little, I would get this feeling like... Like I’m floating outside of my body, looking down at myself... And I hate what I see... How I’m acting, the way I sound. And I don’t know how to change it. And I’m so scared... That the feeling is never gonna go away.”

The Edge of Seventeen¹

Much has been written about optimism and pessimism in psychological studies for decades (Scheier and Carver, 1992). These feelings are affected by one’s personality from an early age (as pinpointed above) and can strongly impact one’s psychological and physical health. For example, pessimism and negative attitudes impact negatively one’s mental health, can induce suicidal thoughts, and affect negatively not only the person in question, but also their family and friends (Peterson and Bossio, 2001; Achat et al., 2000;

Scheier et al., 2001). On the other hand, optimism reduces stress and promotes better physical health and overall well-being (Carver et al., 2010).

Despite that optimism and pessimism are under the scrutiny of many researchers (Rasmussen et al., 2009; Kumar et al., 2017), large scale analyses that explore optimism and pessimism in social media have just started to emerge (Ruan et al., 2016). However, Ruan et al. (2016) focused on identifying optimism and pessimism in Twitter using a simple “bag of words” representation with no emphasis on incorporating semantic information hidden in text. Often, a deeper understanding of the text that accounts for textual semantic similarities and the writer’s intention are required in order to correctly detect the characteristics of optimistic and pessimistic feelings in tweets. Towards this end, our contributions in this paper are as follows. First, we focus on the question: “Would a deep learning approach help to discover these characteristics better than traditional machine learning classifiers used in prior work?” To our knowledge, we take the first step towards exploring the performance of deep learning models for optimism/pessimism prediction in Twitter and identify the most promising deep learning models for this task. Identifying optimism and pessimism in Twitter has many applications including identifying suicidal/depressive people and providing better social support (e.g., emotional/empathetic support) that can improve people’s moods and attitudes (Yan and Tan, 2014; Biyani et al., 2014; Khanpour et al., 2018, 2017; Qiu et al., 2011).

Second, since it may seem intuitive that a positive sentiment is associated with optimism and a negative sentiment with pessimism, we address the question: “Would a sentiment classifier be sufficient to correctly identify optimism and pessimism in social media?” Figure 1 shows evidence that a sentiment tool would not suffice on accurately

¹<https://www.imdb.com/title/tt1878870/>

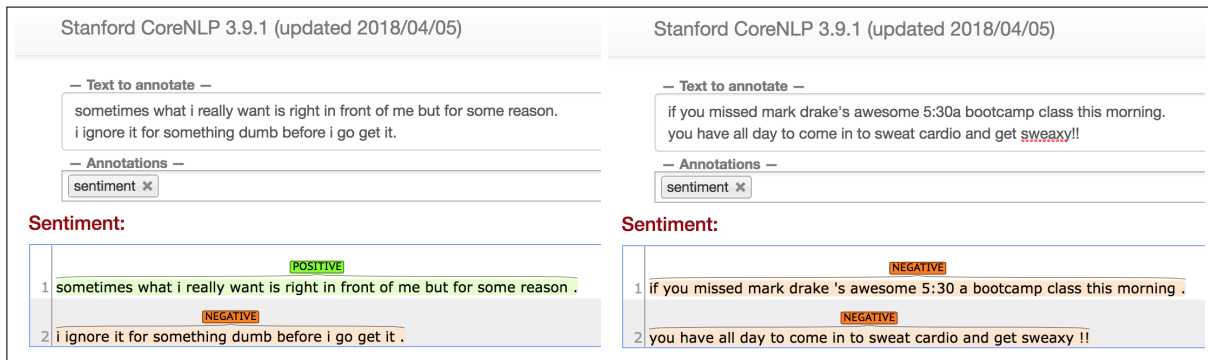


Figure 1: Sentiment prediction on tweets with pessimist (left) and optimist (right) connotations.

predicting tweets with pessimistic and optimistic connotations (left and right side of the figure, respectively). We answer the above question by investigating a spectrum of sentiment analysis tools and datasets for optimism/pessimism prediction.

Third, we perform a linguistic analysis, first of its kind, and study the usage of verb tenses (past, present, future) in optimistic and pessimistic tweets, as well as the presence of polarity words associated with both types of tweets.

2 Datasets

In this section, we first describe the optimism/pessimism Twitter dataset and then present two datasets used for sentiment analysis.

The Optimism/Pessimism Twitter dataset (OPT) was made available to us by Ruan et al. (2016). The total number of tweets in the dataset is 7,475. These tweets were sampled from data corresponding to 500 optimist and 500 pessimist users, and were manually annotated using Amazon Mechanical Turk. Precisely, each tweet was manually annotated by five independent annotators using a score between -3 (very pessimistic) and 3 (very optimistic). For our evaluation, we consider two different thresholds (0 and $1/-1$) on the above score and create two settings as follows. In the first evaluation setting, a tweet is labeled as pessimistic if its score is smaller than or equal to 0 , and optimistic, otherwise. In the second evaluation setting, a tweet is labeled as pessimistic if its score is smaller than or equal to -1 , and optimistic if its score is greater than or equal to 1 . A summary of this dataset is given in Table 1.

	Threshold 0	Threshold 1/-1
Number of tweets	7,475 (O: 4,679)	3,847 (O: 2,507)
Number of users	1,000	1,000

Table 1: Dataset statistics. ‘O’ stands for optimistic tweets.

The Stanford Sentiment Treebank (SST) (Socher et al., 2013) is a corpus for sentiment analysis that capture complex linguistic patterns. This dataset² is based on a dataset originally introduced by Pang and Lee (2005) and consists of 10,662 sentences from movie reviews downloaded from rottentomatoes.com. From these sentences, 215,154 phrases were extracted using the Stanford Parser (Klein and Manning, 2003) and labeled using Amazon Mechanical Turk such that each phrase was annotated by 3 human judges.

The Twitter Sentiment Analysis (TSA) dataset,³ available online for download, contains 1,578,627 tweets that are classified as 1 for positive sentiment and 0 for negative sentiment.

3 Experiments and Results

In experiments, we explore several deep learning models for optimism/pessimism prediction. The general training strategy is as follows: sentence embeddings are fed into a sentence encoder to obtain the sentence representation. The classifier consists of three fully connected layers topped by a softmax layer. Dropout was applied to the first layer only. We used several encoders as follows, based on: (1) Bidirectional Long Short Term Memory networks (BiLSTMs), which are a special type of Recurrent Neural Networks (RNNs) (Hochreiter and Schmidhuber, 1997); (2) Convolutional Neural Networks (CNNs), which consist of convolution and max pooling (Kim, 2014); and (3) Stacked Gated RNNs (Chung et al., 2015).

We used SGD optimizer (Goodfellow et al., 2016) with a learning rate of 0.1 and no weight decay. At every tenth epoch we decreased the learn-

²<https://nlp.stanford.edu/sentiment/>

³<http://thinknook.com/twitter-sentiment-analysis-training-corpus-dataset-2012-09-22/>

Threshold	Tweet Level		User Level	
	0	1/-1	0	1/-1
BiLSTM	79.65	87.24	76.65	90.52
GRUStack	80.19	87.76	76.38	92.24
CNN	77.78	90.32	73.55	91.68
NB	74.20	84.10	71.30	80.10
SVM	67.80	83.30	64.70	81.80

Table 2: Accuracy of deep learning vs. traditional classifiers on the **OPT** dataset (shown as percentage).

ing rate by half. We used mini-batches of 40 samples. Dropout rate was set to 0.5 and the classifier’s last three layers have 300, 200, and 100 neurons. We used GloVe vectors (Pennington et al., 2014) trained on Common Crawl 840B⁴ with 300 dimensions as fixed word embeddings.

For sentence embedding, after a cleanup process, sentences were transformed into a list of words, then words were replaced with word embeddings (GloVe) and padding was used to align batch sentences to the same size.

3.1 Optimism/Pessimism Prediction

In our first experiment, we explore the above deep learning models on the **OPT** dataset and compare their performance with that of two traditional machine learning classifiers, Naïve Bayes (NB) and Support Vector Machines (SVM), which were used in the previous work for this task by Ruan et al. (2016). In this experiment, the **OPT** dataset is split in train-dev-test as 80-10-10(%), respectively. We repeated each experiment 5 times and averaged the results. Our deep learning implementation is built on top of TensorFlow (Abadi et al., 2015). For NB and SVM, we used their implementation available in scikit-learn (Pedregosa et al., 2011). Table 2 shows the accuracy of all these models at tweet and user level for the two thresholds 0 and 1/-1 (as discussed in Section 2).

We can see that overall, the deep learning models achieve a much higher performance compared with the work by Ruan et al. (2016), i.e., the NB and SVM classifiers on “bag of words,” for both tweet and user level with both thresholds, yielding an improvement in performance between 5%-10%. For example, at tweet level and 1/-1 threshold, CNN yields an accuracy of 90.32% as compared with NB, which achieves an accuracy of 84.10%. At user level and 1/-1 threshold, GRUStack yields an accuracy of 92.24%, as compared

⁴<https://nlp.stanford.edu/projects/glove/>

with 81.80% achieved by SVM. Not surprising, for both tweet and user level, when we use a threshold of 0, the performance of all models is smaller compared with that of models obtained when we use a 1/-1 threshold. Intuitively, this is true since most of the tweets with a human-annotated score between -1 and 1 are in the “gray” area that is harder to classify. Note that Ruan et al. (2016) considered the tweets with a score between -1 and 1 as being neutral.

3.2 Sentiment vs. Optimism/Pessimism

In our second experiment, we investigate the correlation between sentiment and optimism / pessimism, and argue that sentiment analyzers, that are trained to predict sentiment (Liu, 2012; Pang and Lee, 2008), fail to detect optimism and pessimism. Specifically, we train several sentiment classifiers on the large **SST** and **TSA** sentiment datasets (described in Section 2) and evaluate the performance of these classifiers on the optimism/pessimism categories from the **OPT** dataset.

Model	Train	Dev	Test	Acc%
LSTM	SST	SST	OPT	63.20
BiLSTM	SST	SST	OPT	63.60
CNN	SST	SST	OPT	59.60
CNN	TSA	TSA	OPT	67.60
RNN(char)	TSA	OPT	OPT	55.20
GRUStack	OPT	OPT	OPT	80.19

Table 3: Performance of sentiment classifiers on **OPT**.

Table 3 shows the performance of several deep learning models trained on either **SST** or **TSA** datasets and evaluated on the **OPT** dataset. Note that the Dev set was used for model selection. As can be seen from the table, the models trained on the sentiment datasets perform poorly on the optimism/pessimism dataset. For example, there is a drop in performance from 80.19% to 67.60% when training on **TSA** (with an even larger decrease when we train on **SST**).

The **SST/TSA** sentiment classifiers are trained to predict the sentiment as negative, neutral, or positive. To calculate the accuracy in Table 3, an optimistic tweet predicted as positive by the sentiment classifier counts as a correct prediction, whereas an optimistic tweet predicted as either neutral or negative by the sentiment classifier counts as an incorrect prediction (similarly for pessimistic tweets). This analysis is done at tweet level for the threshold of 0.

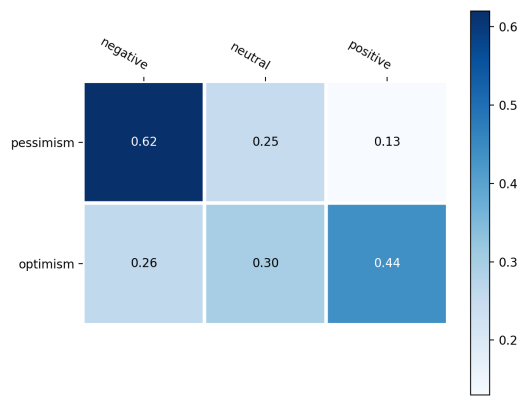


Figure 2: Optimism and pessimism classified as positive, negative, and neutral sentiment.

Figure 2 shows the normalized number of examples from optimism and pessimism categories classified as positive, negative and neutral, using the CNN model trained on TSA. Precisely, we show how many tweets from the set of optimistic (or pessimistic) tweets in the **OPT** dataset are predicted as negative, neutral or positive by the **TSA** sentiment classifier. The numbers on each row sum up to 1. As we can see from the figure, although pessimism is more correlated with a negative sentiment, 13% of the pessimistic tweets are classified as positive (with similar results on the optimism category).

3.3 Linguistic Analysis

In this section, we perform a linguistic analysis and study the usage of verb tenses in optimistic and pessimistic tweets, as well as the presence of polarity words associated with both types of tweets. This analysis is done at tweet level with 1/-1 threshold. The reason for using the 1/-1 threshold is that we wanted to study the usage of verb tenses and polarity words in tweets that are clear optimistic or clear pessimistic (far from the decision boundary).

3.3.1 Verb Tenses in Optimism/Pessimism

For this analysis, we used the part of speech tagger spaCy⁵ and assigned the verbs to their corresponding tenses according to the Penn Treebank Project; that is, the tags VBD and VBN correspond to past tense, VBG, VBZ, VBP correspond to present tense, whereas an MD tag followed by VB (possibly with a negation between them) corresponds to the future tense.

⁵<http://textanalysisonline.com/spacy-pos-tagging>

As mentioned, a tweet was considered optimist if its manually annotated score was above 1 and pessimist if the score was below -1 . The numbers of tweets with past, present, and future tenses in the optimistic category are: 1,474, 7,444, and 561, respectively, whereas these numbers in the pessimistic category are: 1,276, 5,311, and 325, respectively.

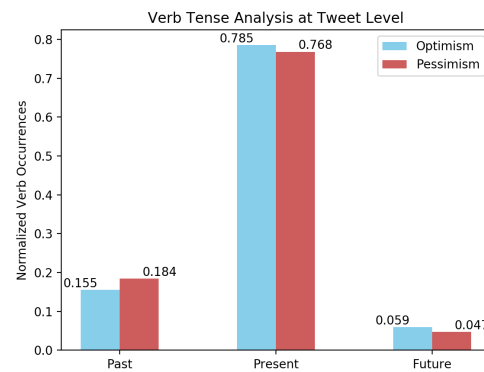


Figure 3: Verb tenses in optimist/pessimist tweets.

Figure 3 shows the normalized verb occurrences at past, present and future tenses in optimistic and pessimistic tweets. As can be seen from the figure, the present tense is the most prevalent for both categories, although there are more present tense verbs in the optimistic category compared with the pessimistic one. We can also observe that more past tense verbs occur in the pessimistic category and less future tense verbs in the pessimistic one.

While there are some common verbs such as “be,” “have,” and “do,” that appear most frequently in both optimistic and pessimistic categories at all three tenses, there are some verbs that are more specific to one category than the other. Examples of such verbs and their frequencies from both categories at the present tense are shown in Table 4. As we can see, optimism is characterized more by verbs with a positive connotation.

[optimism]	[pessimism]
(‘be’, 1662)	(‘be’, 1198)
(‘have’, 738)	(‘have’, 549)
(‘do’, 316)	(‘do’, 342)
(‘love’, 266)	(‘hate’, 122)
(‘thank’, 105)	(‘f-k’, 88)
(‘look’, 98)	(‘kill’, 35)
(‘want’, 94)	(‘try’, 34)
(‘hope’, 70)	(‘cry’, 33)

Table 4: Verbs in tweets occurring at present tense.

3.3.2 Polarity Words in Optimism/Pessimism

Next, we analyze the association of polarity words from the positive and negative lexicons constructed by Hu and Liu (2004), in both tweet categories: optimism and pessimism. Instead of using the presence or absence of the words from the two lexicons in tweets, we calculated the cosine similarity between the word embeddings of the words in the two lexicons with the words in the tweets. If the similarity is above 0.8, then we consider the word from the corresponding lexicon to be present in the tweet (or synonym with a word in tweet). Using the cosine similarity between the word embeddings of words in lexicons with words in tweets captures not only the exact match between the words (a cosine similarity of 1 for exact match), but also incorporates the semantic information that exists in the text.

Although this word similarity computation relaxes the exact match/presence of a word in a tweet and aims at incorporating semantic similarity, a high similarity between antonyms may occur since word embeddings are known to not differentiate well between synonyms and antonyms, which tend to appear in similar contexts.

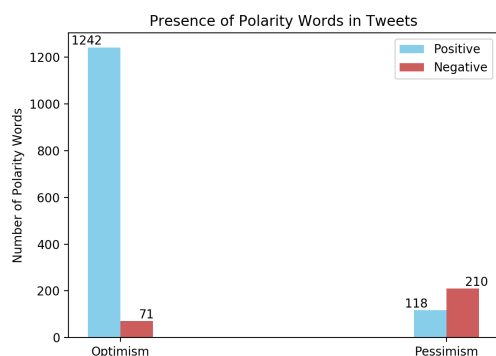


Figure 4: Polarity words in optimist/pessimist tweets.

Figure 4 shows the number of polarity words in optimistic and pessimistic tweets. As shown in the figure, more positive words appear in optimistic tweets compared with negative words (1,242 vs. 71), while there is not a substantial difference between the numbers of positive and negative words in pessimistic tweets (118 vs. 210).

Table 5 shows the top most frequent polarity words associated with optimism and pessimism. As we can see, words with a negative polarity (e.g., bad) although not very frequent, still appear in optimistic tweets. This supports our intuition that a sentiment model is not enough to accurately predict pessimism and optimism in Twitter.

Optimism	
[positive]	[negative]
('great', 289)	('bad', 13)
('loved', 279)	('worried', 11)
('wonderful', 155)	('lost', 6)
('glad', 76)	('scared', 5)
('kind', 45)	('terrible', 4)
('thrilled', 29)	('disappointed', 3)
Pessimism	
[positive]	[negative]
('great', 29)	('bad', 36)
('kind', 27)	('lost', 18)
('loved', 13)	('scared', 16)
('wonderful', 9)	('alone', 13)
('surprised', 8)	('terrible', 12)
('glad', 6)	('terrified', 12)

Table 5: Top frequent polarity words.

4 Concluding Remarks

In this paper, we explored deep learning models for optimism and pessimism prediction in Twitter and showed that these models substantially outperform traditional classifiers such as Naïve Bayes and Support Vector Machines. To our knowledge, this work is the first computational study that explores optimism and pessimism using deep learning. We also showed that a sentiment classifier would not be sufficient for accurately predicting optimism and pessimism. This topic is less explored despite its importance in many applications such as identifying suicidal/depressive people.

Interesting future directions are: understanding how one's age is correlated with optimism / pessimism; if one user is characterized by a mixture of topics, is that user optimist (pessimist) across all these topics? Thus, decomposing a user's textual data into topic and correlating this with optimism and pessimism may be interesting to explore; last, studying how optimism and pessimism are affected by sarcasm.

As we started our study with a pessimistic quote from the movie "The Edge of Seventeen," we end our study with a quote from the same movie, with a positive sentiment and full of optimism:

"Life's about taking risks. Don't be afraid to put yourself out there."

Acknowledgments

All authors contributed equally. LP Dinu was supported by UEFISCDI, project #53BG/2016. We thank our reviewers for their constructive comments and feedback.

References

- Martín Abadi, Ashish Agarwal, Paul Barham, Eugene Brevdo, Zhifeng Chen, Craig Citro, Greg S. Corrado, Andy Davis, Jeffrey Dean, Matthieu Devin, Sanjay Ghemawat, Ian Goodfellow, Andrew Harp, Geoffrey Irving, Michael Isard, Yangqing Jia, Rafal Jozefowicz, Lukasz Kaiser, Manjunath Kudlur, Josh Levenberg, Dandelion Mané, Rajat Monga, Sherry Moore, Derek Murray, Chris Olah, Mike Schuster, Jonathon Shlens, Benoit Steiner, Ilya Sutskever, Kunal Talwar, Paul Tucker, Vincent Vanhoucke, Vijay Vasudevan, Fernanda Viégas, Oriol Vinyals, Pete Warden, Martin Wattenberg, Martin Wicke, Yuan Yu, and Xiaoqiang Zheng. 2015. TensorFlow: Large-scale machine learning on heterogeneous systems. Software available from tensorflow.org.
- Helen Achat, Ichiro Kawachi, Avron Spiro, Deborah A DeMolles, and David Sparrow. 2000. Optimism and depression as predictors of physical and mental health functioning: the normative aging study. *Annals of Behavioral Medicine*, 22(2):127–130.
- Prakhar Biyani, Cornelia Caragea, Prasenjit Mitra, and John Yen. 2014. Identifying emotional and informational support in online health communities. In *Proceedings of COLING 2014, the 25th International Conference on Computational Linguistics: Technical Papers*, pages 827–836, Dublin, Ireland. Dublin City University and Association for Computational Linguistics.
- Charles S. Carver, Michael F. Scheier, and Suzanne C. Segerstrom. 2010. Optimism. *Clinical psychology review*, 30(7):879–889.
- Junyoung Chung, Caglar Gulcehre, Kyunghyun Cho, and Yoshua Bengio. 2015. Gated feedback recurrent neural networks. In *Proceedings of the 32nd International Conference on International Conference on Machine Learning - Volume 37, ICML'15*, pages 2067–2075.
- Ian Goodfellow, Yoshua Bengio, and Aaron Courville. 2016. *Deep Learning*. MIT Press. <http://www.deeplearningbook.org>.
- Sepp Hochreiter and Jürgen Schmidhuber. 1997. Long short-term memory. *Neural Computation*, 9(8):1735–1780.
- Minqing Hu and Bing Liu. 2004. Mining and summarizing customer reviews. In *Proceedings of the Tenth ACM SIGKDD International Conference on Knowledge Discovery and Data Mining*, KDD '04, pages 168–177, New York, NY, USA. ACM.
- Hamed Khanpour, Cornelia Caragea, and Prakhar Biyani. 2017. Identifying empathetic messages in online health communities. In *Proceedings of the Eighth International Joint Conference on Natural Language Processing, IJCNLP 2017, Taipei, Taiwan, November 27 - December 1, 2017, Volume 2: Short Papers*, pages 246–251.
- Hamed Khanpour, Cornelia Caragea, and Prakhar Biyani. 2018. Identifying emotional support in online health communities. In *Proceedings of the Thirty-Second AAAI Conference on Artificial Intelligence, New Orleans, Louisiana, USA, February 2-7, 2018*.
- Yoon Kim. 2014. Convolutional neural networks for sentence classification. In *Proceedings of the 2014 Conference on Empirical Methods in Natural Language Processing (EMNLP)*, page 1746?1751.
- Dan Klein and Christopher D. Manning. 2003. Accurate unlexicalized parsing. In *Proceedings of the 41st Annual Meeting on Association for Computational Linguistics - Volume 1, ACL '03*, pages 423–430, Stroudsburg, PA, USA. Association for Computational Linguistics.
- Upendra Kumar, Vishal Singh Rana, Srinivas Pykl, and Amitava Das. 2017. “A pessimist sees the difficulty in every opportunity; an optimist sees the opportunity in every difficulty” - understanding the psychosociological influences to it. In *Proceedings of the 14th International Conference on Natural Language Processing*.
- Bing Liu. 2012. Sentiment analysis and opinion mining. *Synthesis lectures on human language technologies*, 5(1):1–167.
- Bo Pang and Lillian Lee. 2005. Seeing stars: Exploiting class relationships for sentiment categorization with respect to rating scales. In *Proceedings of the 43rd Annual Meeting on Association for Computational Linguistics, ACL '05*, pages 115–124, Stroudsburg, PA, USA. Association for Computational Linguistics.
- Bo Pang and Lillian Lee. 2008. Opinion mining and sentiment analysis. *Journal of Foundations and Trends in Information Retrieval*, 2(1-2):1–135.
- Fabian Pedregosa, Gaël Varoquaux, Alexandre Gramfort, Vincent Michel, Bertrand Thirion, Olivier Grisel, Mathieu Blondel, Peter Prettenhofer, Ron Weiss, Vincent Dubourg, Jake Vanderplas, Alexandre Passos, David Cournapeau, Matthieu Brucher, Matthieu Perrot, and Édouard Duchesnay. 2011. Scikit-learn: Machine learning in python. *The Journal of Machine Learning Research*, 12:2825–2830.
- Jeffrey Pennington, Richard Socher, and Christopher D. Manning. 2014. Glove: Global vectors for word representation. In *Empirical Methods in Natural Language Processing (EMNLP)*, pages 1532–1543.
- Christopher Peterson and Lisa M Bossio. 2001. Optimism and physical well-being. *Optimism and pessimism: Implications for theory, research, and practice*, pages 127–145.
- Baojun Qiu, Kang Zhao, Prasenjit Mitra, Dinghao Wu, Cornelia Caragea, John Yen, Greta E Greer, and

- Kenneth Portier. 2011. Get online support, feel better—sentiment analysis and dynamics in an online cancer survivor community. In *Privacy, Security, Risk and Trust (PASSAT) and 2011 IEEE Third International Conference on Social Computing (Social-Com), 2011 IEEE Third International Conference on*, pages 274–281. IEEE.
- Heather N Rasmussen, Michael F Scheier, and Joel B Greenhouse. 2009. Optimism and physical health: A meta-analytic review. *Annals of behavioral medicine*, 37(3):239–256.
- Xianzhi Ruan, Steven Wilson, and Rada Mihalcea. 2016. Finding optimists and pessimists on twitter. In *Proceedings of the 54th Annual Meeting of the Association for Computational Linguistics (Volume 2: Short Papers)*, pages 320–325, Berlin, Germany. Association for Computational Linguistics.
- Michael F Scheier and Charles S Carver. 1992. Effects of optimism on psychological and physical well-being: Theoretical overview and empirical update. *Cognitive therapy and research*, 16(2):201–228.
- Michael F Scheier, Charles S Carver, and Michael W Bridges. 2001. Optimism, pessimism, and psychological well-being. *Optimism and pessimism: Implications for theory, research, and practice*, 1:189–216.
- Richard Socher, Alex Perelygin, Jean Wu, Jason Chuang, Christopher D. Manning, Andrew Ng, and Christopher Potts. 2013. Recursive deep models for semantic compositionality over a sentiment treebank. In *Proceedings of the 2013 Conference on Empirical Methods in Natural Language Processing*, pages 1631–1642, Seattle, Washington, USA. Association for Computational Linguistics.
- Lu Yan and Yong Tan. 2014. Feeling blue? go online: an empirical study of social support among patients. *Information Systems Research*, 25(4):690–709.