

EmoDex

An emotion detection tool composed of established techniques

Oxana Zhurakovskaya, Louis Steinkamp, Karsten Michael Tymann, and Carsten Gips

FH Bielefeld University of Applied Sciences, Minden, Germany
oxana.zhurakovskaya@fh-bielefeld.de, louis.steinkamp@fh-bielefeld.de,
ktymann@fh-bielefeld.de, carsten.gips@fh-bielefeld.de
<https://www.fh-bielefeld.de>

Abstract. In this work we created an emotion analysis tool consisting of established models and techniques: Ekman's and Plutchik's emotion models, WordEmbedding (GloVe), VADER sentiment analysis, emoji features and a RandomForest classifier. Additionally we composed a corpus based on existing corpora and with the help of distant supervision. As a result, our approach achieves an accuracy increase of up to 10% compared to other emotion analysis tools (ParallelDots and Twitter Emotion Recognition), while at the same time offering a broader set of emotion classes. In addition, adding a sentiment feature increased the accuracy by about 2%. We make the conclusion that a combination of features from multiple sources such as GloVe and VADER offer a good basis for a RandomForest classifier while only training on a very small set of texts (less than 70k sentences).

Keywords: Emotion detection · Random Forest · Word Embedding · GloVe · VADER · Emoji labelling · Sentiment analysis · Emotions · Ekman · Plutchik · Distant supervision · CrowdFlower · Feature selection

1 Introduction

Emotions are an important part of human communication. They influence the semantic meaning of sentences and can therefore convey additional information. While having a face to face conversation one is able to derive the emotional meaning of the partner's message not only by the actual spoken sentences but also by incorporating other factors such as facial expressions, gestures and voice. When reading texts, these natural factors are lost. Communications in textual form can therefore be misinterpreted, especially by machines. But being able to identify the emotion of an online text can be beneficial for multiple applications. With modern natural language processing (NLP) it is possible to process text messages to detect which emotions are expressed.

Copyright © 2020 by the paper's authors. Use permitted under Creative Commons License Attribution 4.0 International (CC BY 4.0).

Interpreting the emotions of a text can be done by a rule based algorithm or machine learning which requires a lot of training data. When relying on supervised learning the data needs to be labeled with emotion categories. To bootstrap the machine learning model, there exists a variety of already labeled corpora. There are also methods such as distant supervision (see section 2.2) to automatically annotate texts with emotions. Different models can be used for emotion labels, like Ekman's [5] and Plutchik's [13]. The more emotion categories, the more complex the classification can become. Thus emotion classification can be often harder than just deriving a sentence's sentiment.

Common approaches in using the text as classification features involve training on the word embeddings of the texts. Some also include separately crawled corpora for emojis which can serve as additional features. Others might as well include separately trained features such as hashtags or features by other classifiers such as sentiment analysis tools.

In this work, which was part of a student project at Bielefeld University of Applied Sciences, we created a free to use emotion analysis webservice, called EmoDex¹. We focus on existing tools and models such as Plutchik's emotion model [13], GloVe [12] for word embedding and existing corpora to train a Random-Forest-Classifer with scikit-learn². Additionally we add as separate features emoji categories as well as a sentiment rating provided by VADER [8]. We will describe all steps from collecting and preprocessing the texts, to testing out whether a separately calculated sentiment score or using distant supervision can improve the classification. EmoDex will be compared to two models, one being ParallelDots API³ while the other being TwitterEmotionRecognition [4].

2 Background

2.1 Emotion Models

To classify text with emotions first we should decide which are the basic emotions that can be identified. Two often used models are Ekman and Plutchik. Ekman's model highlights 6 basic emotions: sadness, happiness, anger, fear, disgust, surprise. The emotions selected in this model are discrete and based on facial expressions as well as neurobiological processes independent of cultural differences. Whereas Plutchik's multidimensional model of emotions is based on the psychoevolutionary theory of emotions. This model identifies 8 basic emotions: joy, trust, fear, surprise, sadness, disgust, anger, anticipation (see Fig. 1). Each have additional intensity levels. [5, 13, 16]

Another model type describes emotions in dimensions. The Valence-Arousal-Dominance (VAD) or also called Pleasure-Arousal-Dominance (PAD) model defines three axes which locate emotions in a space. First the pleasure or valence

¹ EmoDex <https://emodex.net/>

² <https://scikit-learn.org/stable/modules/generated/sklearn.ensemble.RandomForestClassifier.html>

³ ParallelDots Inc. <https://www.paralleldots.com/emotion-analysis>

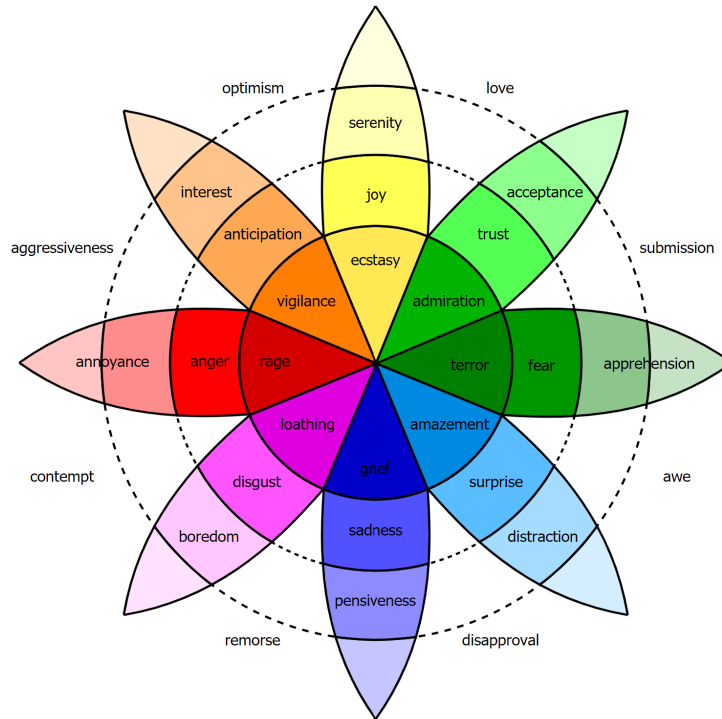


Fig. 1: Plutchik's wheel of emotion with the 8 inner basic emotions [7]

axis describes how pleasant a feeling is. Second the arousal axis shows how much a person feels “activated”. Being excited would for example be high in arousal whereas sadness or calmness have a low arousal value. From high arousal feelings an action can be more expected by the individual than from a person having a low arousal emotion. Thirdly the dominance scale shows how dominant or submissive the persons feeling is. Being angry would be a very dominant feeling while sadness would indicate a more submissive behaviour. [19, 14]

All three models are types of different emotions categorization. Ekman's model describes emotions as discrete emotions whereas the VAD/PAD model by A. Russell describes them dimensional. Plutchik's can be regarded as a hybrid model, where the 8 basic emotions can be extended by further emotions as dimensions. [19, 14, 16]

2.2 Distant supervision & expert labeling

To label the data for training purposes we used expert labeling in combination with distant supervision.

Expert labeling describes the process that the test data has been annotated by human experts. In the best case, a test data set is evaluated by several experts

so that the label of the data set is as accurate as possible. However, this type of labeling is very labor-intensive and subjective due to the human judgement [18].

The other type of labeling is the automatic creation of labeled datasets, called distant supervision. Especially via Twitter, the hashtag search can be used to filter for emotions. For example, if one searches with the hashtag #joy, tweets that are annotated with the hashtag will be returned. The assumption is that the user has used this hashtag to express his emotions in this tweet. Therefore the tweet will also be labeled with this emotion label in the data set.

The authors of the paper [18] compared the accuracy of expert annotation and remote supervision and created a test corpus of 400 tweets, which was annotated using both methods. The result is that the labels match 93.16% and are therefore suitable as a meaningful label for the dataset.

2.3 Corpora compilation

One of the main components of emotion recognition in texts is the corpus on which a ML algorithm can be trained. Since there are already several papers that have dealt with the topic of corpora creation, a collection of corpora which are free to use for research purposes will be used in this work.

The paper [3] has already examined various of such corpora in detail and analysed their suitability for classification. The corpus is based on 14 different corpora, which are labelled according to different emotion categories and come from different topics like news, blogs, weather or general. Furthermore their type of labeling process (distant supervision & expert labeling) is shown and they are differentiated in their granularity, like tweets, headlines or simple sentences. The result of their work included also a mapping of emotion classification, in order to merge all individual corpora in one data set. Only two of the corpora have Ekman and Plutchik as emotions model. As a result of this mapping, seven data sets use Ekman as a basis, while they extend the model by one or two additional emotions. Two corpora are according to the VA/PA and VAD/PAD model respectively and therefore find no observation in our further investigation. One model is only divided into happy and sad and is therefore comparable to a sentiment classification.

2.4 Detecting emotions with WordEmbedding

WordEmbedding describes a recent trend in ML and NLP where words are represented as vectors in a vector space. The embedded words are therefore in a relation to each other that can be measured as the vector distance. [9]

Word2Vec and GloVe, both being WordEmbedding techniques, are useful for emotion analysis since they represent the words with their semantic meaning in a vector space. Therefore the assumption is that the words in the same clusters offer a similar emotion. For an emotion analysis the sentence can therefore be split into the dimensional representation of the total of the words dimension vectors. With the so received dimension vector of the sentence, a machine learning classifier can be trained. [9, 1, 12]

2.5 Emoji Labelling

Emojis make a significant contribution to non-verbal communication in texts. [17] Through them, users are given another opportunity to express their emotions. [6] Normally displayed as icons, it is also possible to interpret emojis as unicodes which are defined by the Unicode Consortium⁴.

Emojis can therefore be used to abstract further information about the emotions in texts. The basis for this is an Emoji Emotion Mapping which assigns an emotion to selected emojis. The mapping makes it possible to classify a text with an emotion only on the basis of emojis in text.

The authors of the paper [6] have crowdsourced 202 emojis with an emotion label. In total 308 users submitted 15155 ratings. The result is an emoji emotion label mapping. As soon as an emoji was rated over 50% with this label, it was assigned to this emotion. Their work is used as a basis for our emoji classification. [6]

2.6 VADER

VADER is a sentiment analysis tool that is based on a crowd rated sentiment lexicon used in a rule based algorithm. The python tool rates sentences on a scale from -1 (negative) to 0 (neutral) to 1 (positive). The tool showed good results in the domain of social media. [8]

2.7 Benchmarking with other tools

For benchmark purposes and comparing our results to other approaches we picked two tools. We chose the ParallelDots API as well as the Twitter Emotion Recognition tool for our comparison.

ParallelDots is developing different NLP and AI products. They offer an API for their emotion recognition tool, which can detect emotions of 6 different categories: happy, sad, angry, fear, excited and indifferent. According to their blog, their model is based on Convolutional Neural Networks (CNNs).

The Twitter Emotion Recognition tool is able to predict emotions for English tweets. [4] It requires no preprocessing since it works on the words characters. It provides a trained Recurrent neural network (RNN) which can predict one of the following categories: Ekman's six basic emotions, Plutchik's eight basic emotions and Profile of Mood States six mood states.

3 Process

3.1 Corpus in detail

As described in chapter 2, the basis of the corpus in this work is a collection of different corpora consisting of tweets. For this purpose five corpora were used, which are described in more detail in the following:

⁴ Unicode Org <https://unicode.org/emoji/charts/full-emoji-list.html>

Name	Source	Emotion	Size	Labeling
Crowd-Flower	Unify Emotion Datasets [3]	Ekman + Love + NoEmotion	39.740	Crowdsourcing
Electoral-Tweets	Mohammad [11]	Plutchik	4.058	Crowdsourcing
EmoInt	Mohammad [2]	Ekman - Disgust - Surprise	7.097	Crowdsourcing
SSEC	Schuff et al. [15]	Plutchik	4.868	Expert Annotation
TEC	Mohammad [10]	Ekman	21.051	Distant Supervision

Table 1: Corpora

As shown in Table 1, three of the corpora are based on Ekman and two on Plutchik. CrowdFlower is extended by the emotion 'Love' and the label 'No Emotion'. EmoInt is shortened to four emotions in which 'disgust' and 'surprise' are removed. This results in a total of 10 labels with which the respective data records can be marked and a total number of 76.814 entries.

Plutchik's emotions 'trust' and 'anticipation' were removed from the data set for this work, because their share is too small in comparison. What remains is a corpus consisting of 8 emotion labels and 72.762 data sets.

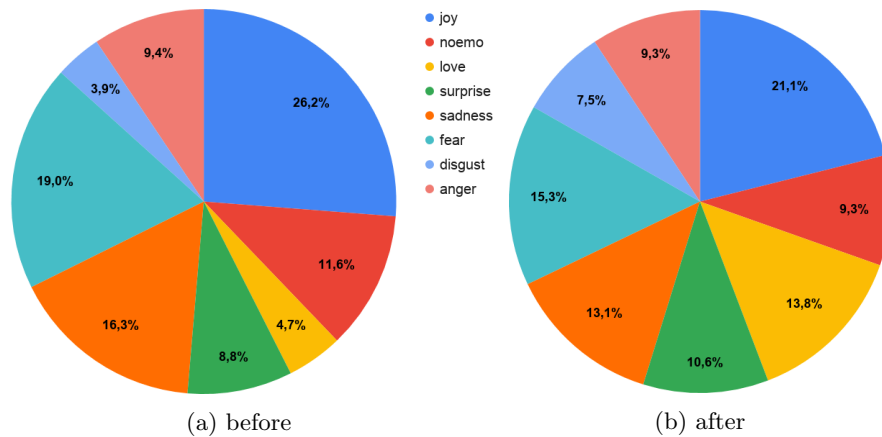


Fig. 2: Emotions distribution before and after distant supervision

Fig. 2a shows how the individual proportions of the respective emotion labels in the corpus are distributed. For eight emotions the average is 12.5 percent. While emotions like joy and fear are far above this average, emotions like disgust and love are far below.

If an ML algorithm is trained with this data set, emotions like disgust or love are hardly recognized because their share is too small. To compensate for this, the proportion of emotions in the corpus can be influenced. There are two possibilities for this. Either the shares of the dominant labels are reduced or the shares of the neglected labels are increased. Since reducing the data is not desirable, we have added more data to the data set using distant supervision.

Due to the fact that our selected corpora are all based on tweets, we decided to use Twitter as data source as well. To crawl Twitter we use the hashtag based search provided by the Twitter API. The hashtags we use for search are based on the National Research Council of Canada (NRC) Hashtag Lexicon for non-commercial purposes by Saif Mohammad [10] which provide multiple hashtags for six of our eight emotion labels. Every hashtag from the lexicon has a score which represents the strength of association between the hashtag and the emotion. We have chosen the highest rated hashtags and added our own tags, so that we arrived at ten hashtags per emotion. For the emotion label 'Love' we have only used our own hashtags, since the label is missing in the Hashtag Lexicon. Finally with this selection we have crawled further 19662 tweets for the emotion labels 'angry', 'disgust', 'surprise' and 'love', so that the corpus for this work comes to 92452 tweets in total.

A percentage distribution of the emotion labels in the final corpus can be seen in Fig. 2b. The percentages of the dominating labels have been reduced, while the percentages of the neglected labels have been increased.

3.2 Preprocessing

The prepared corpus should be preprocessed to remove irrelevant words and signs as well as emoticons and emojis. In order to take emojis into account we considered to add additional features, that represent the emotion category of emojis based on the idea of emoji labelling (see section 2.5). In contrast to the referenced approach, the emojis in this work are categorised by only one human and each emoji is assigned to only one category. There are eight categories in total, that are appropriate to the selected emotions features: joy, love, surprise, disgust, sadness, fear and neutral. Each text should have a count for each of these emojis categories. Thus in the first step emojis in each text should be counted and the count should be added to the feature vector. Fig. 3a shows an example of categorised emojis. All emojis are removed from the original text after counting.

In order to process the emoticons, the emoticons were replaced with a word representing them. Fig. 3b shows an example of emoticons and their descriptions. The replacement of emoticons with appropriate description is the second step in the preprocessing pipeline.

As the third step all letters in the text are converted to lowercase. In order to reduce the word amount to be processed, some unnecessary words should be removed. Thus stopwords, URLs, usernames like "@name" and hashtags like "#tag" are removed. The negation words were however left in the text, because these influence the emotional meaning of the text. After the removal manipulation there can be left empty texts. Consequently the empty texts are removed from

Emoji	Category
😊	joy
😍	love
😮	surprise
😐	neutral
😒	disgust
😞	sadness
😨	fear
😡	anger

(a) Emoji category mapping

Emoticon	Description for replacement
&:	disgust
(%:	confused
((-:	smile
(-:<	gloatingly

(b) Emoticon replacement mapping

Fig. 3: Emoji and Emoticon mappings

the corpus. Additionally we add a sentiment classification score of the text to the feature vector. The idea was to enhance the classification of emotions by providing an additional feature, that allows to better distinguish negative and positive emotions like “joy” and “sadness”. Therefore we processed each text of the corpus with the VADER tool (see section 2.6) and added the result as a feature. In order to compare the influence of sentiment features on emotion classification we store one corpus with VADER preprocessing and one corpus without VADER classification.

3.3 Use of GloVe

After preprocessing, the texts can be converted into their vector representation. For this purpose a pre-trained word embedding model from GloVe is used in this paper. The model was trained on 2B tweets and contains a total vocabulary of 1.2 M words. For our work we used a word vector resolution of 100 dimensions. For each word in a tweet the vector was calculated from the model. Then the average of the sum of the words was used as a vector representation of the tweet and got appended to the corpus to serve as features. [12]

3.4 Random Forest Classifier

For classification we selected the Random Forest (RF) algorithm, which consists of multiple Decision Trees, each predicting the outcome class independently of each other. The class that gets the most “votes” is the result of the whole Random Forest. One of the advantages of this method is reducing the errors

of predictions that can occur when predicting only with a single individual tree. Another advantage is overfitting control.

For implementing the Random Forest we use the scikit-learn⁵ framework, that provides ready to use functions with configuration options. We use the Random Forest classifier with the default options, with 200 trees and the random state set to 42. We train the classifier on 75% (67.5k) of our corpus and tested it with the other 25% (22.5k).

4 Results

For the training and testing of the data we have divided our corpus in a ratio of 3 to 1. The test data was randomly selected from the entire corpus. We have trained four different models (with/without VADER, with 8/4 emotions) and use two types of analysis (with/without 20% threshold) to evaluate the results. We compared our results with the two tools explained in chapter 2.7: ParallelDots (PD) and Twitter Emotion Recognition Tool (TER). We have benchmarked both by predicting parts of our corpus, however some mappings for the emotions had to be made. The results are shown in Table 2.

The ParallelDots API only returned classification for the five emotions: happy, sad, angry, fear and excited. We mapped the emotion 'happy' to our emotion 'joy'. The emotion excited has been removed from the evaluation. We tested the API with the 4 corresponding emotions from our corpus. Thus we have a total of 26028 rated tweets and an accuracy of 40.84% (see Table 2 row 2).

TER uses Plutchik's emotion model and thus directly covers six of our eight emotions. We have removed the two additional emotions trust and anticipation as well as our data sets labeled love or noemo. For this tool we have had 8938 tweets predicted and have an accuracy of 31,02% (row 1).

Our tool was trained with 75% and tested with the other 25%. This resulted in a test data set of 22470 data sets which our model with VADER (row 4) rated correctly with 45,11%. Thereby all eight emotions were used. Without VADER as an additional feature (row 3), the result is 43,76% with eight emotions.

For both versions, with VADER as a feature and without, a model with only four emotions was trained and tested. These are joy, sadness, anger and fear and are in accordance with the four emotions, which are also used for our benchmarking for ParallelDots. For training and testing, the data sets with the remaining four emotions were removed from the corpus. The results are 53,35% accuracy with VADER (row 6) and 51,39% (row 5) without.

All results were evaluated by using the emotion label with the highest classification accuracy (row 1-6). Since the ratings of eight labels can be close to each other, we used a second rating strategy for this model. It was checked whether an emotion label was rated with more than 20%. If so, it was added to the result set. The prediction was then marked as true if the test record label was within the result set. As a result the classification accuracy increased (row 7, 8).

⁵ <https://scikit-learn.org/stable/modules/generated/sklearn.ensemble.RandomForestClassifier.html>

Row	Tool	Features	Tweets	True	False	Accuracy	Emotion Label
1	TER	Pretrained RNN	8938	2613	6325	31,02%	Anger, Disgust, Fear, Joy, Sadness, Surprise, Trust , Anticipation
2	PD	Pretrained CNN	26028	9526	16499	40,84%	Happy (Joy), Sad, Angry, Fear, Excited
3	EmoDex	RF, GloVe	22463	9829	12634	43,76%	Joy, Anger, Sad, Disgust, Fear, Surprise, Love, NoEmo
4	EmoDex	RF, GloVe, VADER	22470	10136	12334	45,11%	Joy, Anger, Sad, Disgust, Fear, Surprise, Love, NoEmo
5	EmoDex	RF, GloVe	12585	6467	6118	51,39%	Joy, Anger, Sad, Fear
6	EmoDex	RF, GloVe, VADER	12585	6714	5871	53,35%	Joy, Anger, Sad, Fear
7	EmoDex	RF, GloVe, 20% threshold	22463	13300	9163	59,21%	Joy, Anger, Sad, Fear
8	EmoDex	RF, GloVe, VADER, 20% threshold	22470	14227	8243	63,32%	Joy, Anger, Sad, Fear

Table 2: Benchmark results

In summary it can be said that the use of VADER as an additional feature brought a gain in accuracy of 1.35% for eight emotion labels and a gain of almost 2% for four emotion labels. The reduction from eight to four emotions brought a gain of about 8%, whereby it should be noted that the model with four emotions was trained and tested on a corpus that was almost 44% smaller.

5 Conclusion and future work

This works approach is mostly based on already established techniques in NLP. We have shown that combined techniques – labeled corpora, distant supervision, Glove, emoji categories, VADER, and random forest classifiers – complement each other to an efficient tool.

In this work the emojis were categorized by one human, thus the emoji labelling is subjective. In future works the emojis can be labeled using crowdsourcing methods or automatically with machine learning methods.

The emotions distribution on the used corpus was not even, thus the results can be affected by this. For future testing the corpus should be build with respect to emotion distribution. Compared to other papers, we have worked with a small dataset, thus our results should be tested with more data in a future analysis.

Compared to the other tools, this works approach has the highest classification score while also offering the most emotion categories in the described test environment. Moreover it is demonstrated that adding sentiment features by third party tools to the feature vector can increase the accuracy result. Additionally distant supervision proved itself to be useful for expanding the corpus.

References

- [1] Tomas Mikolov et al. *Learning Representations of Text using Neural Networks*. 2013. URL: <http://www.micc.unifi.it/downloads/readinggroup/TextRepresentationNeuralNetwork.pdf> (visited on 02/01/2020).
- [2] Alexandra Balahur, Saif M. Mohammad, and Erik van der Goot, eds. *Proceedings of the 8th Workshop on Computational Approaches to Subjectivity, Sentiment and Social Media Analysis*. Copenhagen, Denmark: Association for Computational Linguistics, Sept. 2017. DOI: 10.18653/v1/W17-52.
- [3] Laura Ana Maria Bostan and Roman Klinger. “An Analysis of Annotated Corpora for Emotion Classification in Text”. In: *Proceedings of the 27th International Conference on Computational Linguistics*. Santa Fe, New Mexico, USA: Association for Computational Linguistics, 2018, pp. 2104–2119. URL: <http://aclweb.org/anthology/C18-1179>.
- [4] N. Colnerić and J. Demsar. “Emotion Recognition on Twitter: Comparative Study and Training a Unison Model”. In: *IEEE Transactions on Affective Computing* (2018), pp. 1–1. DOI: 10.1109/TAFFC.2018.2807817.
- [5] Paul Ekman. “Are there basic emotions?” In: *Psychological Review* 99.3 (1992), pp. 550–553. DOI: 10.1037/0033-295X.99.3.550.
- [6] Abdallah El Ali, Torben Wallbaum, Merlin Wasmann, Wilko Heuten, and Susanne Boll. “Face2Emoji: Using Facial Emotional Expressions to Filter Emojis”. In: May 2017, pp. 1577–1584. DOI: 10.1145/3027063.3053086.
- [7] Wikipedia - The Free Encyclopedia. *Robert Plutchik*. URL: https://de.wikipedia.org/wiki/Robert_Plutchik (visited on 02/01/2020).
- [8] C. Hutto and Eric Gilbert. *VADER: A Parsimonious Rule-Based Model for Sentiment Analysis of Social Media Text*. 2014. URL: <https://www.aaai.org/ocs/index.php/ICWSM/ICWSM14/paper/view/8109>.
- [9] Tomas Mikolov, Ilya Sutskever, Kai Chen, Greg Corrado, and Jeffrey Dean. “Distributed Representations of Words and Phrases and Their Compositionality”. In: *Proceedings of the 26th International Conference on Neural Information Processing Systems - Volume 2*. NIPS’13. Lake Tahoe, Nevada: Curran Associates Inc., 2013, pp. 3111–3119.
- [10] Saif Mohammad. “#Emotional Tweets”. In: **SEM 2012: The First Joint Conference on Lexical and Computational Semantics – Volume 1: Proceedings of the main conference and the shared task, and Volume 2: Proceedings*

- of the *Sixth International Workshop on Semantic Evaluation (SemEval 2012)*. Montréal, Canada: Association for Computational Linguistics, July 2012, pp. 246–255. URL: <http://www.aclweb.org/anthology/S12-1033>.
- [11] Saif Mohammad, Xiaodan Zhu, Svetlana Kiritchenko, and Joel Martin. “Sentiment, emotion, purpose, and style in electoral tweets”. In: *Information Processing & Management* 51 (Oct. 2014). DOI: 10.1016/j.ipm.2014.09.003.
- [12] Jeffrey Pennington, Richard Socher, and Christopher D. Manning. “GloVe: Global Vectors for Word Representation”. In: *Empirical Methods in Natural Language Processing (EMNLP)*. 2014, pp. 1532–1543. URL: <http://www.aclweb.org/anthology/D14-1162>.
- [13] Robert Plutchik. “A psychoevolutionary theory of emotions”. In: *Social Science Information* 21.4-5 (1982), pp. 529–553. DOI: 10.1177/053901882021004003.
- [14] James A Russell and Albert Mehrabian. “Evidence for a three-factor theory of emotions”. In: *Journal of Research in Personality* 11.3 (1977), pp. 273–294. DOI: [https://doi.org/10.1016/0092-6566\(77\)90037-X](https://doi.org/10.1016/0092-6566(77)90037-X).
- [15] Hendrik Schuff, Jeremy Barnes, Julian Mohme, Sebastian Padó, and Roman Klinger. “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*. Copenhagen, Denmark: Association for Computational Linguistics, Sept. 2017, pp. 13–23. DOI: 10.18653/v1/W17-5203.
- [16] Armin Seyeditabari, Narges Tabari, and Wlodek Zadrozny. “Emotion Detection in Text: a Review”. In: *CoRR* (2018). arXiv: 1806.00674.
- [17] Jessica L Tracy, Daniel Randles, and Conor M Steckler. “The nonverbal communication of emotions”. In: *Current Opinion in Behavioral Sciences* 3 (2015). Social behavior, pp. 25–30. DOI: <https://doi.org/10.1016/j.cobeha.2015.01.001>.
- [18] Wenbo Wang, Lu Chen, Krishnaprasad Thirunarayan, and Amit P. Sheth. “Harnessing Twitter “Big Data” for Automatic Emotion Identification”. In: *Proceedings of the 2012 ASE/IEEE International Conference on Social Computing and 2012 ASE/IEEE International Conference on Privacy, Security, Risk and Trust. SOCIALCOM-PASSAT '12*. USA: IEEE Computer Society, 2012, pp. 587–592. ISBN: 9780769548487. DOI: 10.1109/SocialCom-PASSAT.2012.119.
- [19] Shen Zhang, Zhiyong Wu, Helen Meng, and Lianhong Cai. “Facial Expression Synthesis Based on Emotion Dimensions for Affective Talking Avatar”. In: vol. 2010. June 2010, pp. 109–132. DOI: 10.1007/978-3-642-12604-8_6.