Proceedings of the Twenty-Eighth International Florida Artificial Intelligence Research Society Conference

Discovering Emotions in the Wild: An Inductive Method to Identify Fine-Grained Emotion Categories in Tweets

Jasy Liew Suet Yan

School of Information Studies, Syracuse University, Syracuse, USA
jliewsue@syr.edu

Abstract

This paper describes a method to expose a set of categories that are representative of the emotions expressed on Twitter inductively from data. The method can be used to expand the range of emotions that automatic classifiers can detect through the identification of fine-grained emotion categories human annotators are capable of detecting in tweets. The inter-annotator reliability statistics for 18 annotators using different granularity of the emotion classification schemes are compared. An initial set of emotion categories representative of the range of emotions expressed in tweets is derived. Using this method, researchers can make more informed decisions regarding the level of granularity and representativeness of emotion labels that automatic emotion classifiers should be able to detect in text.

Introduction

Twitter[1], a microblogging site with more than 100 million monthly active users, is particularly rich with emotive content as it is a common place for users to publicly share how they feel about various events, entities and topics on a global scale. Emotive content on Twitter can be harnessed to gain insights regarding users' perceptions, behaviors, and the social interactions between individuals and populations of interest in a non-invasive manner. Growing interest in analyzing emotions on Twitter is evidenced by studies of how emotions expressed on microblogs affect stock market trends (Bollen, Mao, & Zeng, 2011), relate to fluctuations in social and economic indicators (Bollen, Pepe, & Mao, 2011), serve as a measure for the population's level of happiness (Dodds & Danforth, 2010), and provide situational awareness for both the authorities and the public in the event of disasters (Vo & Collier, 2013). With 500 million tweets being sent a day, automatic emotion detectors are needed to help augment our ability to analyze and understand emotive content and are especially useful in helping businesses learn about consumer emotional reactions toward various products, services, and events through large-scale analysis of online user-generated content.

Many current automatic emotion detection systems utilize knowledge-based methods (i.e., using emotion lexicons or ontologies) and statistical methods to detect emotions in text. An important starting point to build machine learning classifiers that can recognize the emotions represented in tweets is the selection of a set of suitable emotion categories. Prior work has mainly focused on distilling emotion knowledge from massive amount of unstructured data into a small set of basic emotions (*happiness*, *sadness*, *fear*, *anger*, *disgust*, and *surprise*) (Ekman, 1971) mainly because these emotion categories are assumed to be universal emotions according to the emotion theories in psychology (Mohammad, 2012; Wang, Chen, Thirunarayan, & Sheth, 2012). The current emphasis on the basic emotions poses limitations on the development of automatic emotion detectors that can truly capture the richness of actual human emotional experiences on Twitter.

First, it is unclear if these basic emotions are representative of the range of emotions humans express on Twitter. Mohammad & Kiritchenko (2014) have found a few hundred emotion words being expressed explicitly using hashtags (notably used to indicate topics) on Twitter, and indicated that identifying finer-grained emotions is useful for personality detection. This clearly shows that the basic emotions framework only give us a small glimpse of emotions being expressed in tweets, and is a poor fit or is too crude to adequately capture the rage of emotions expressed in tweets. While the basic emotions framework offers a starting point to study emotions expressed in text, it is crucial to note that the basic emotions only represent emotions that are important in dealing with "fundamental life tasks" (Ekman, 1999). Originally, the basic emotions framework was derived from facial expressions and physiology, and is not grounded on language theories. Human use of language

[1] Twitter: https://about.twitter.com/company

to express and describe emotions is not limited to basic emotions as illustrated in the examples below:

Example 1: *"My prayers go to family of Amb. Stevens & others affected by this tragedy. We must not allow the enemy to take another. http://t.co/X8xTzeE4" (Sympathy)*

Example 2: *"Thanks 2 my dad 4 going out on road today 4 southern MN gotv events. He is 84&strong as ever. He doesn't tweet but maybe when he turns 85." (Gratitude)*

Second, many emotions that are not part of the basic set are ignored or worse, force-fit into one of the basic emotion categories. Example 1 is a case of "sympathy" as the writer is expressing his or her condolences to people affected by a tragedy. Since "sympathy" is not one of the six basic emotions, Example 1 is most likely classified as the basic emotion "sadness". Example 2 is an expression of "gratitude" that is most likely classified as the basic emotion "happiness". The coarse fit between the emotions expressed in text and the basic set of categories makes it more difficult for automatic emotion detectors to extract pertinent linguistic patterns for each emotion category because of the considerable amount of fuzziness and noise introduced into the corpus.

Automatic emotion detectors are built and evaluated using emotion corpora and other emotion-related language resources. With the basic emotions accepted as the state-of-the-art regardless of context, existing language resources are only annotated with the basic emotion categories at the finest level of granularity. For instance, Pak & Paroubek (2010) created a corpus with two emotion categories: positive and negative, while (Mohammad, 2012) and Wang et al. (2012) applied Ekman's six emotions in the construction of their corpora. As a result, automatic emotion detectors are only able to give us a limited picture of human emotional expression. Complex emotions, as well as variations within each basic emotion are "virgin territories" that have not yet been explored by researchers in this area. Efforts to increase the utility of automatic emotion detectors have to start with extending language resources to cover a richer set of emotion categories that both humans and computers can reliably detect in text.

Instead of picking a set of emotion categories from existing emotion theories in psychology, a systematic method to expose a set of categories that are more representative of the emotions expressed on Twitter inductively from data is developed and described in the paper. This method will allow us to expand the range of emotions automatic classifiers can detect through the identification of emotion categories human annotators can identify in tweets. This paper seeks to address two research questions:

R1: What emotions can humans detect in microblog text? Can human annotators reliably identify emotion categories at a level finer that the six basic emotions?

R2: What level of inter-annotator reliability can we expect from fine-grained emotion classification compared to the current state-of-the-art coarse-grained emotion classification?

Models of Emotion

A starting point to build automatic emotion detectors is to determine how emotions can be classified. The categorization of emotion is largely based on two common models of emotion adopted from emotion theories in psychology: 1) the categorical model, and 2) the dimensional model (Calvo & Mac Kim, 2012; Zachar & Ellis, 2012).

The dimensional model aims to account for all emotions in simpler and more general dimensions as opposed to discrete emotion categories. It holds that all emotional phenomena share the same fundamental structure, and can be identified from the composition of two or more independent dimensions (Zachar & Ellis, 2012). The two popular dimensions used in automatic emotion detection are valence (Alm, Roth, & Sproat, 2005; Kennedy et al., 2011; Strapparava & Mihalcea, 2008) and arousal (Aman & Szpakowicz, 2007; Neviarouskaya, Prendinger, & Ishizuka, 2007; Zhe & Boucouvalas, 2002). Valence indicates if an emotion is "positive", "negative" or "neutral". Arousal can be measured using labels representing varying intensities (e.g., low, moderate or high) or a numerical scale.

The dimensional measures allow researchers to capture more nuanced differences of emotions in text without the constraint of fitting all emotional phenomena into a limited set of categories. Nuanced variations within a dimension and combinations of different dimensions can then be used to determine discrete emotion categories. For instance, Cambria, Olsher, & Rajagopal (2014) employed a composition of 4 independent dimensions that can be used to determine the 24 emotion categories mapped in an hourglass model (Cambria, Livingstone, & Hussain, 2012). However, dimensional labels are less intuitive and more ambiguous to a lay person compared to categorical labels (Read, 2004). In addition, identifying the minimum number of dimensions to adequately define all emotions remains a difficult challenge.

In the categorical model, emotions are classified into discrete categories, and each category represents a distinct emotion (Cowie, 2009). Each emotion category is characterized by a set of emotion patterns or structures that sets it apart from other categories. An emotion label is used to represent each category (e.g., happy, sad, angry) but there are various lexical realizations for each emotion label. For instance, the emotion label "fear" is associated with different words used to describe someone feeling threatened (e.g., "terrified", "afraid", and "scared").

There are two advantages associated with categorical labels: 1) using intuitive labels makes it easier to understand the emotion associated with the label, and 2) researchers have the flexibility to use different dimensions or criteria to define each emotion category. It is crucial for researchers to draw clear distinctions between different emotion categories to avoid any confusion in the interpretation of the emotion labels. The common categorization models are Ekman's six basic emotions (Alm et al., 2005; Aman & Szpakowicz, 2007; Chaumartin, 2007; Liu, Lieberman, & Selker, 2003; Mohammad, 2012; Strapparava & Mihalcea, 2008) and Plutchik's eight prototypical emotions (Plutchik, 1962), an expansion of Ekman's basic emotion with the addition of *trust* and *anticipation* (Mohammad & Turney, 2013; Mohammad, Zhu, & Martin, 2014; Suttles & Ide, 2013). More recent work has also adopted the 22 emotion categories from the cognitive structure of emotions (Shaikh, Prendinger, & Ishizuka, 2009). However, the emotion categories are directly adopted from emotion theories in psychology without a systematic investigation of how well the categories fit the range of emotions expressed in text.

Method

Annotation Scheme Development

In this exploratory method to discover a representative set of fine-grained emotions from tweets, 18 annotators were recruited annotate a sample of 300 tweets for annotation scheme development. The 300 tweets were sampled randomly from 96 official US Senators Twitter accounts[2] from November 2011 to November 2012. The sample, though too limited in size to make the findings generalizable, made annotation feasible for 18 annotators (see Table 1 for annotators' demographic information) and provided an initial set of emotion categories which can be tested and expanded using more data in the future. Annotators consisted of graduate students who were interested in undertaking the task as part of a class project (e.g., Natural Language Processing course) or to gain research experience in content analysis (e.g., independent study). Each annotator had to first attend a one hour training session to discuss the concept of emotion with the researcher. The annotators worked independently on the annotation task but met with the researcher in groups to discuss disagreements. Annotators annotated the 300 tweets in 3 stages within 3 weeks.

Tweets were segmented at the message-level. Since tweets were relatively short in length (i.e., 140 characters at maximum), annotation conducted at the message level provides annotators with sufficient context to identify the emotion expressed in text. Tweets were pre-processed beforehand to remove duplicates, retweets, and non-English text. Annotators were given instructions to annotate tweets based on two facets of emotions described in Table 2: polarity and emotion tag.

Demographic Aspect	**# of Annotators**
Gender	
- **Female**	11
- **Male**	7
Geographic region of origin	
- **USA**	3
- **China**	9
- **India**	3
- **Southeast Asia**	2
- **Middle East**	1

Table 1: Annotators' demographic information.

For emotion tag, annotators were asked to suggest the best-fitting emotion tag to describe the emotion expressed in each tweet (see Example 3). In cases where a tweet may contain multiple emotions, annotators are asked to first identify the primary emotion expressed in the tweet, and then include the other emotions observed (see Example 4).

Example 3: "*Alaska is so proud of our Spartans! The 4-25executed every mission in Afghanistan with honor & now, they're home http://t.co/r8pLpnud" (Positive: Pride)*

Example 4: *"Saw Argo yesterday, a movie about the 1979 Iranian Revolution. Chilling, sobering, and inspirational at the same time." (Positive: Inspiration, Negative: Fear)*

Facet	**Description**	**Codes**
Polarity	Expressing pleasure or displeasure towards events, objects or situations	Positive: Expressing pleasure (e.g., happiness) Negative: Expressing displeasure (e.g., anger) Neutral: Emotion that is neither positive nor negative (e.g., surprise) No Emotion
Emotion Tag	Emotion category that best describes the emotion expressed in a tweet	Open coding

Table 2: Emotion facets in the annotation scheme.

[2] US Senator Twitter usernames were obtained from http://www.tweetcongress.org/.

Emotion Category		
Positive	**Neutral**	**Negative**
Gratitude [1], Pleased [2], Happiness [3], Excitement [4], Pride [5], Inspiration [6], Admiration [7], Love [8], Hope [9], Confidence [10]	Surprise [11]: {Amazement [13], Shock [12]}, Curiosity [14], Exhaustion [15]	Disappointment [16], Displeased [17], Annoyance [18], Anger [19], Fear [20], Sadness [21], Regret [22], Guilt [23], Sympathy [24], Worry [25], Hate [26]

Table 3: Emotion categories discovered from tweets.

The inductive approach was used to construct the annotation scheme through observation of content (Potter & Levine-Donnerstein, 1999). Construction of the annotation scheme does not start from a theoretical framework. Instead, annotators begin by looking for themes in the data, and then move to empirical generalization. The annotation scheme is refined through an iterative coding process until a set of categories are finalized. This inductive approach, also known as grounded theory, was developed by Glaser & Strauss (1967) for the purpose of building theory that emerges from the data.

Refining Emotion Categories

To refine the emotion tags emerging from data, annotators were asked to perform a card sorting exercise in different teams to group emotion tags that are semantically similar into the same category once they have completed annotation for the 300 tweets. Annotators were divided into 5 teams, and performed the card sorting activity on the emotion tags used by the all members in their respective teams. Each team was instructed to follow the four-step procedures described below:

1) Group all the emotion tags into categories.
2) Decide a name for the emotion category. Collectively pick the most descriptive emotion tag or suggest a new name to represent each category.
3) Group all the emotion categories based on polarity: positive, negative and neutral.
4) Match emotion categories generated from other team's card sorting activity to the emotion categories proposed by your team.

At the end of each card sorting activity, the researcher merged, divided, and verified the final emotion categories to be included in the annotation scheme. Once the final emotion categories were identified, the original emotion tag labels generated from the open coding exercise were systematically replaced by the appropriate emotion category labels.

Results

The final 26 emotion categories that human annotators were able to identify from tweets are shown in Table 3. Two types of classification scheme emerged from the 5 card sorting activity sessions. Four teams proposed a flat classification scheme resulting in the 26 emotion categories as shown in Table 3 (labeled as Emotion-Category-26 in Table 4). One team came up with a hierarchical classification scheme (see Emotion-Category-17 in Table 4). In the hierarchical classification scheme, related emotion categories are further grouped together into higher-level parent categories.

Fleiss' kappa and Krippendorff's alpha were used as measures of inter-annotator reliability for multiple annotators as these statistics control for chance agreement whereas simple percent agreement does not (Neuendorf, 2002). Out of 300 tweets in the corpus, only 248 tweets tagged with a single emotion were used in the computation of the inter-annotator reliability statistics. Tweets tagged with multiple emotions were excluded. At the coarsest level of granularity concerning whether emotion was present or absent in a tweet, the three valence labels (positive, negative and neutral) were merged into a single category.

All 18 annotators were able to achieve an alpha of over 0.6 at the level of emotion being present or absent and polarity respectively even with minimal training in this exploratory annotation task. It is interesting to note that all 18 annotators were able to achieve a slightly higher alpha for polarity (4 categories) compared to presence or absence of emotion (2 categories).

From Table 4, it can be observed that although inter-annotator reliability statistics tend to be higher when classification scheme is coarser, human annotators still managed to achieve a moderate level of reliability ranging from 0.4 – 0.6 (Artstein & Poesio, 2008) when identifying finer-grained emotions in tweets. At the finest-grained level with 26 distinct emotion categories, annotators achieved an alpha of 0.443.

Further analysis on the annotations revealed that it was challenging for annotators to differentiate emotion categories that were closely related in meaning but varied in intensity. For instance, higher disagreements were observed in the use of the emotion categories "pleased", "happiness" and "excitement". When emotion categories displaying such characteristics were merged in a hierarchy proposed in one of the five card sorting exercise sessions (see Emotion-Category-17 in Table 4), an alpha of 0.529 was observed.

Level of Granularity	# of Cat	Categories	Percent Agreement	Fleiss' Kappa	Krippendorff's Alpha
Emotion present/absent	2	Emotion, None	81%	0.627	0.627
Polarity	4	Positive, Negative, Neutral, None	78%	0.630	0.630
Emotion-Category-8 (categories further merged by researcher)	9	Happiness: {1 – 8}, Hope: {9, 10}, Sadness: {21 – 24}, Fear: {20, 25}, Anger: {16 – 19, 26}, Surprise: {11 – 13}, Curiosity: 14, Exhaustion: 15, None	78%	0.595	0.595
Emotion-Category-6 (categories other than the 6 basic emotions recoded as "other")	8	Happiness: {2, 3, 4}, Sadness: {21, 22, 23}, Fear: 20, Anger: {16, 17, 18, 19}, Surprise: {11, 12, 13}, Hate, Other, None	73%	0.538	0.538
Emotion-Category-17	18	1, {2, 3, 4}, 5, 6, 7, 8, 9, 10, {11, 12, 13}, 14, 15, {16, 17, 18, 19}, 20, {21, 22, 23}, 24, 25, 26, None	72%	0.529	0.529
Emotion-Category-26	27	1 – 26, None	66%	0.442	0.443

Table 4: Inter-annotator reliability statistics corresponding to classification schemes at different levels of granularity.

Another data transformation taking into account of only the 6 basic emotions commonly used in current automatic emotion detectors and recoding other emotions categories into "other" yielded an alpha of 0.538. This shows that human annotators are still able to achieve comparable inter-annotator reliability using a finer-grained classification scheme of 17 categories as opposed to a classification scheme that only takes into account of 6 basic emotions.

By merging the 17 emotion categories further into a coarser-grained classification scheme (see Emotion-Category-8 in Table 4), the alpha increased to 0.6. The 8 distinct emotion categories discovered from data differ slightly from Ekman and Plutchik's basic emotions but can be considered an expansion of the basic emotion model commonly used in current automatic emotion detectors. Notably, instances of "disgust" are absent from the corpus. Also, emotions such as "curiosity", "exhaustion" and "hope" have their own set of distinct characteristics that are linguistically identifiable in text.

In this exploratory stage aimed at discovering a suitable classification scheme for the automatic emotion detector, moderate agreement between annotators is acceptable. Once the granularity of the classification scheme is determined, annotators can be further trained to improve agreement.

Conclusion and Future Work

This paper describes a method to inductively derive fine-grained emotion categories that are representative of the range of emotions expressed in microblog text. The contribution of this paper is two-fold: 1) the introduction of an inductive method that can be used to discover the range of emotion expressed in text from various domains, and 2) the initial development of an emotion taxonomy that more accurately reflects the range of emotions expressed in text. Existing automatic emotion detectors usually adopt emotion labels directly from emotion theories in psychology that may not be well-suited to the rich range of emotions expressed in text. Using this method, researchers can compare the inter-annotator reliability achieved with different levels of granularity of the emotion classification scheme. Then, more informed decisions can be made regarding the level of granularity and representativeness of emotion labels that automatic emotion classifiers should be able to detect in text.

The emotion categories discovered from this small set of tweets are limited and need to be further tested on more tweets. As part of future work to make the classification scheme more generalizable, the emotion categories will be further tested and expanded by continuing annotation efforts on tweets generated by a broader range of Twitter users. Annotators will incrementally annotate more tweets to test the emotion categories until a point of saturation is reached, where new emotion categories stop emerging from data. Throughout this iterative process, new emotion categories can be added to the current set of emotion categories, and problematic emotion categories may be merged with other categories or removed from the current set. The researcher will continuously meet with the annotators to discuss disagreements until the expected inter-annotator agreement for the final set of emotion categories is achieved.

The annotated corpus generated from this process will be used to train and test machine learning models to automatically identify these emotion categories in microblog text, and will be made available to other researchers in the future.

Acknowledgments

I am grateful to Dr. Howard Turtle, Dr. Elizabeth Liddy and Dr. Nancy McCracken for their support, and the annotators for their contribution. I also want to thank the Ministry of Higher Education Malaysia and Universiti Sains Malaysia for making this research possible.

References

Alm, C. O., Roth, D., & Sproat, R. 2005. Emotions from text: Machine learning for text-based emotion prediction. In *Proceedings of the Conference on Human Language Technology and Empirical Methods in Natural Language Processing* (pp. 579–586). Stroudsburg, PA, USA.

Aman, S., & Szpakowicz, S. 2007. Identifying expressions of emotion in text. In *Text, Speech and Dialogue* (pp. 196–205).

Artstein, R., & Poesio, M. 2008. Inter-coder agreement for computational linguistics. *Computational Linguistics*, *34*(4), 555–596.

Bollen, J., Mao, H., & Zeng, X. 2011. Twitter mood predicts the stock market. *Journal of Computational Science*, *2*(1), 1–8.

Bollen, J., Pepe, A., & Mao, H. 2011. Modeling public mood and emotion: Twitter sentiment and socio-economic phenomena. In *Proceedings of the Fifth International AAAI Conference on Weblogs and Social Media* (pp. 450–453).

Calvo, R. A., & Mac Kim, S. 2012. Emotions in text: Dimensional and categorical models. *Computational Intelligence*, *29*(3), 527–543.

Cambria, E., Livingstone, A., & Hussain, A. 2012. The hourglass of emotions. *Cognitive Behavioural Systems* (pp. 144–157). Springer Berlin Heidelberg.

Cambria, E., Olsher, D., & Rajagopal, D. 2014. SenticNet 3: A common and common-sense knowledge base for cognition-driven sentiment analysis. In *28th AAAI Conference on Artificial Intelligence*.

Chaumartin, F. R. 2007. UPAR7: A knowledge-based system for headline sentiment tagging. In *Proceedings of the 4th International Workshop on Semantic Evaluations* (pp. 422–425). Prague.

Cowie, R. 2009. Perceiving emotion: Towards a realistic understanding of the task. *Philosophical Transactions of the Royal Society of London B: Biological Sciences*, *364*(1535), 3515–3525.

Dodds, P. S., & Danforth, C. M. 2010. Measuring the happiness of large-scale written expression: Songs, blogs, and Presidents. *Journal of Happiness Studies*, *11*(4), 441–456.

Ekman, P. 1971. Universals and cultural differences in facial expressions of emotion. *Nebraska Symposium on Motivation*, *19*, 207–283.

Ekman, P. 1999. Basic emotions. In *Handbook of Cognition and Emotion* (pp. 45–60). John Wiley & Sons, Ltd.

Glaser, B. G., & Strauss, A. L. 1967. *The discovery of grounded theory: Strategies for qualitative research*. Chicago: Aldine.

Kennedy, A., Kazantseva, A., Mohammad, S., Copeck, T., Inkpen, D., & Szpakowicz, S. 2011. Getting emotional about news. In *Proceedings of the Text Analysis Conference*. Gaithersburg, MD.

Liu, H., Lieberman, H., & Selker, T. 2003. A model of textual affect sensing using real-world knowledge. In *Proceedings of the 8th International Conference on Intelligent User Interfaces* (pp. 125–132). New York, USA.

Mohammad, S. M. 2012. #Emotional tweets. In *Proceedings of the First Joint Conference on Lexical and Computational Semantics* (pp. 246–255). Montreal, QC.

Mohammad, S. M., & Kiritchenko, S. 2014. Using hashtags to capture fine emotion categories from tweets. *Computational Intelligence*.

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

Mohammad, S. M., Zhu, X., & Martin, J. 2014. Semantic role labeling of emotions in tweets. In *Proceedings of the ACL 2014 Workshop on Computational Approaches to Subjectivity, Sentiment, and Social Media (WASSA)* (pp. 32–41). Baltimore, MD.

Neuendorf, K. A. 2002. *The content analysis guidebook*. SAGE.

Neviarouskaya, A., Prendinger, H., & Ishizuka, M. 2007. Textual affect sensing for sociable and expressive online communication. In *Affective Computing and Intelligent Interaction* (pp. 218–229). Springer Berlin Heidelberg.

Pak, A., & Paroubek, P. 2010. Twitter as a corpus for sentiment analysis and opinion mining. In *Seventh International Conference on Language Resources and Evaluation (LREC)*.

Plutchik, R. 1962. *The Emotions: Facts, theories, and a new model*. New York: Random House.

Potter, W. J., & Levine-Donnerstein, D. 1999. Rethinking validity and reliability in content analysis. *Journal of Applied Communication Research*, *27*(3), 258–284.

Read, J. 2004. *Recognizing affect in text using pointwise-mutual information* (M. Sc. Dissertation). University of Sussex.

Shaikh, M. A. M., Prendinger, H., & Ishizuka, M. 2009. A linguistic interpretation of the OCC emotion model for affect sensing from text. In *Affective Information Processing* (pp. 45–73). Springer London.

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

Suttles, J., & Ide, N. 2013. Distant supervision for emotion classification with discrete binary values. *Computational Linguistics and Intelligent Text Processing* (pp. 121–136). Springer Berlin Heidelberg.

Vo, B.-K. H., & Collier, N. 2013. Twitter emotion analysis in earthquake situations. *International Journal of Computational Linguistics and Applications*, *4*(1), 159–173.

Wang, W., Chen, L., Thirunarayan, K., & Sheth, A. P. 2012. Harnessing Twitter "big data" for automatic emotion identification. In *2012 International Conference on Privacy, Security, Risk and Trust (PASSAT), and 2012 International Conference on Social Computing (SocialCom)* (pp. 587–592).

Zachar, P., & Ellis, R. D. (2012). *Categorical versus dimensional models of affect: A seminar on the theories of Panksepp and Russell* (Vol. 7). John Benjamins Publishing Company.

Zhe, X., & Boucouvalas, A. C. (2002). Text-to-emotion engine for real time internet communication. In *Proceedings of International Symposium on Communication Systems, Networks and Digital Signal Processing* (pp. 164–168).