MedCrave
Step into the World of Research

MOJ Applied Bionics and Biomechanics

Mini review

Open Access

CrossMark

Detecting emotional ambiguity in text

Abstract

An approach for determining emotional ambiguity in text data is described in this paper. The prediction confidences output from a text classifier are used to measure amount of ambiguity found in target entries. This measure can be used as a filtering mechanism to identify entries that require human feedback. This feedback loop can be implemented in a workflow which retrains a classifier model including newly disambiguated entries and resulting in a boost to classifier accuracy. This emotion ambiguity measure can be utilized to discover concrete emotional content in text data as well as reveal topics which do not have a concrete emotional consensus.

Keywords: emotion, ambiguity, deep learning, natural language processing

Volume 4 Issue 3 - 2020

Jeffrey Jenkins
The Catholic University of America, Dept. of Electrical Engineering and Computer Science

Correspondence: Jeffrey Jenkins, Dept. of Electrical Engineering and Computer Science, The Catholic University of America, USA, Email jeff.jenkins.1986@gmail.com

Received: August 29, 2019 | **Published:** May 25, 2020

Introduction

Text analysis has gained interest in the research community recently, due in part to the accessibility of analysis tools as well as the abundance of open source data. Additionally, ideologues using social media as a platform to spread fear throughout the world have prompted a need to further understand this complex subject from different angles. Psychologist Robert Plutchik created a wheel of emotions to illustrate different emotions and how they might be related. His circumplex model makes connections between the idea of an emotion circle and a color wheel. Like colors, primary emotions can be expressed at different intensities and can mix with one another to form other emotions.[1] Natural Language Processing (NLP) strategies have begun to address the need for defining emotional features, as well as identifying emotional ambiguity. Crowd sourcing has been employed to understand the common emotions in the English language.[2,3] Machine Learning (ML) contests have been held to produce algorithms that can recognize emotions in text samples.[4] In addition to capturing the emotion of text data in a single word, metrics such as polarity have been established to capture sentiment intensity.[5] The use of ML classifiers such as Convolutional Neural Networks has been proposed to analyze imagery from a smartphone camera for signs of emotional distress in order to provide a decision aid to law enforcement.[6] This paper describes a methodology to measure ambiguity in the classification of text data. This measure can be used as a means of sorting data based on a classifier's 'certainty' and then present ambiguous entries on a crowd sourcing platform for consensus driven labeling. Emotion consensus can be met when a percentage of users have labelled the same sentence with the same emotion, causing that entry to become 'finalized' and added to a queue for retraining. Ultimately, this forms a feedback loop between the classifier and a human analyst with the goal of retraining the classifier with disambiguated entries to improve overall classification accuracy on a benchmarking dataset.

Materials and methods

An approach to utilize NLP techniques and state-of-the-art text classification strategies for the purpose of detecting emotional ambiguity in text is described here. The derivation of ambiguity is suitable for any vector of prediction confidences output by a classifier. The ML libraries keras, tensor flow, and scikit-learn were dependencies in the python code which were used to train and perform classification using Deep Bidirectional Transformers (BERT)[7] (all code is available upon request). First, we define an emotional landscape that is a subset of Plutchik's emotion wheel Figure 1, consisting of the primary emotions found in the middle ring - anger, fear, joy, sadness, as well as the addition of neutral. We trained BERT with default hyper parameters on a composite dataset of 10k text entries pulled from numerous open source Twitter datasets,[8] with labels for ~2k sentences per emotion. We can visualize the manifestation of ambiguity with these 5 emotions on the output of the BERT classifier (Figure 2).

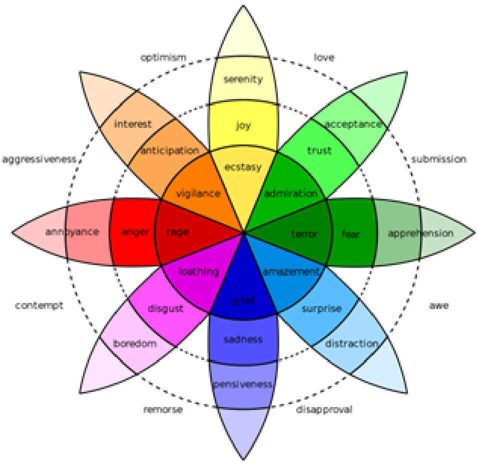

Figure 1 Plutchik's wheel of emotions. A subset of emotions (joy, sadness, fear, anger and neutral were used to classify the training data.

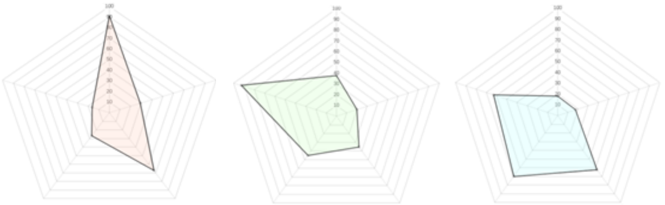

Figure 2 Radar plot with prediction confidences for text entries of increasing ambiguity. Starting from the top of the radar plot and going clockwise, the emotions for each axis are anger, fear, neutral, sadness, and joy. a) low ambiguity - "It makes me angry to drive in traffic", b) medium ambiguity - "I love the weather outside!", c) high ambiguity - "I love snow but hate sunshine".

Here we derive the formula for emotional ambiguity on the range from 0 to 1, where 1 represents maximal ambiguity. First, we define the ambiguity normalization factor as $A_n = 2(n-1) * \frac{1}{n}$, where n is the number of classes to be classified (5 in this case). The vector of class prediction confidences is defined as $\vec{p} = (p_1, p_2, \ldots, p_n)$, where p_i is the probability of the *ith* emotion label for a given text entry. The ambiguity for a given classification is given as $A = \left[\dfrac{A_n - \sum_1^n \left| \frac{1}{n} - p_i \right|}{A_n} \right] + 1$, which is a measure of similarity to a purely ambiguous entry where $p_1 = p_2 = \ldots = p_n$. We can then adopt semi-supervised learning to aid in classifier accuracy improvement with a re-training workflow. since a human in the loop for supervised learning is an expensive proposition, we can request clarification on ambiguous entries iteratively until the accuracy plateaus (Figure 3).

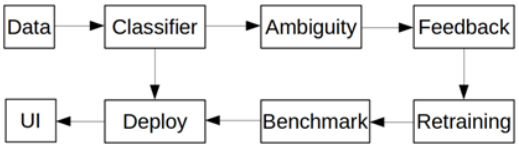

Figure 3 Workflow incorporating ambiguity metric into classifier retraining lifecycle from raw data to user interface for interactive applications.

We observe that a conditional probability with 2-point probability is defined as $P(X,Y) = P(X)P(Y) = P(Y)P(Y \mid X)$. Then, over time, $P(X \mid Z)$ as the negative example with Z as the condition of the positive X this condition can form a Markov chain $P(Y) = P(Z)P(Z \mid Y)$ etc, allowing the machine to gain experience, becoming 'older and wiser'. Some experimental studies will be performed in the near future to explore this technique with multiple modalities of data.

Preliminary results

We investigated the ambiguity contained in a 10,000 tweet subset from the collection of 3 million Twitter posts connected to the Internet Research Agency, a Russian "troll factory" and a defendant in an indictment filed by the Justice Department in February 2018, as part of special counsel Robert Mueller's Russia investigation.[9] The ambiguity histogram on Figure 4A depicts a bimodal distribution with maxima at both extremes – very ambiguous (right side) and not ambiguous (left side). Figure 4B. represents the ambiguity histogram after retraining, which shows a heavily one-sided distribution, with most entries now unambiguous. We validated the outcome against a benchmarking dataset containing hand-selected easy, medium, and hard tweets from the same collection. The graph below shows the accuracy for BERT before and after retraining with highly ambiguous labelled examples with representative sample of the 3 million tweets. Based on the benchmarking exercise (Figure 5) we can conclude that the BERT classifier does a better job on text entries of varying difficulty after retraining the model with ambiguous texts properly labelled. After re-running the dataset through the new model, the histogram in Figure 2A shows a more accurate representation of the ambiguity content in the dataset. This insight is useful for an analyst when we transformed into a tool that can perform a filtering operation on the data based on high ambiguity. If ambiguity is detected in an input sample, feedback from a human can be requested by the machine whereby a singular label can be assigned. Then, all models used after this point can be refined by retraining the classifier with the newly labeled training data. While the examples for this paper are typically single sentences, this approach can be extended to document level classification. Additional emotions can be added to more closely represent Plutchik's emotion wheel, where ambiguity between pairs of emotions can approach dyads, or emotion combinations.

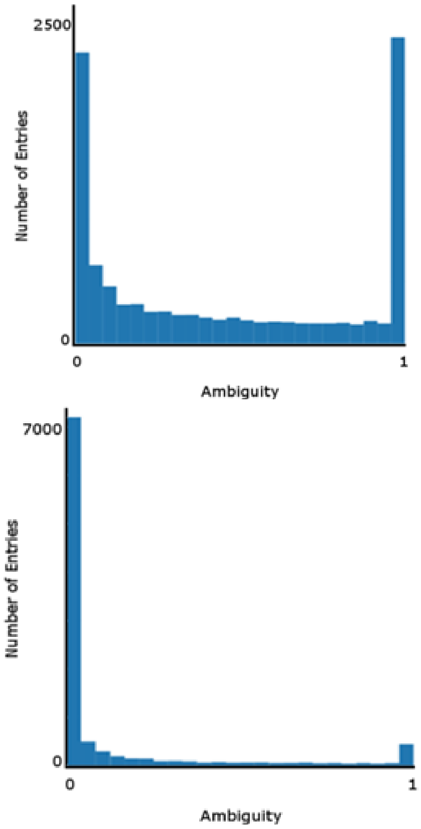

Figure 4 Ambiguity histograms; the x-axis is the ambiguity value and the y-axis is the number of entries. (A) Original BERT histogram indicates a bifurcation between highly ambiguous and unambiguous emotional content in samples. (B) BERT retrained histogram indicates that most entries are unambiguously a specific emotion.

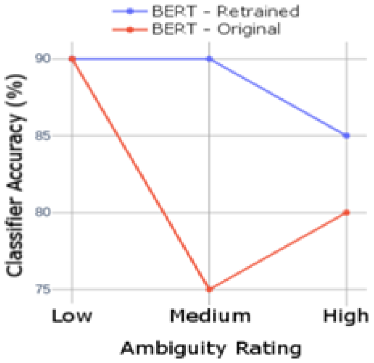

Figure 5 Classifier accuracy by ambiguity ranked difficulty. The blue (top) curve represents BERT after retraining based on ambiguity feedback, and the blue (bottom) curve represents original BERT performance. BERT performance was increased for both medium and highly ambiguous test examples.

Conclusion

The applications of emotional ambiguity detection and resolution are vast. Such a predictive mechanism could be used for discovering early onset of dementia and Alzheimer's for home alone seniors by detecting ambiguity in regular daily conversation. Teaching a machine to get to know it's user better through 6W (who, what, where, when, why, how) inference,[10] and developing an emotional-IQ over time based on fluctuations in 6W[11] can improve relevance in search engine recommender systems and personalized home-based assistants. Future work will include adding evaluating additional sources of ambiguity such as context, topic, and mental state. Psycho-physiological baselining and profiling can be enhanced by introducing emotional ambiguity detection within stimulus-response tasks in Electro Encephalo-Gram (EEG) brain-wave experiments.[12] Animal brains are made up of Biological Neural Networks (BNN), which have wired chemical signaling structures, such as calcium ion currents, via complex genetic and epigenetic structures. Animal brains also possess wireless chemical signaling mechanisms, i.e. hormones, that are clearly found to modulate fear, joy, anger, sadness, love, etc. Such wireless signaling mechanisms are missing from current computer architecture and we propose to represent an emotional BNN compound as a first step toward a more realistic Man-Machine Artificial Neural Network (ANN) interface.[13] Future work will include evaluating the ambiguity measure on different types of classification tasks such as image and audio processing.

Acknowledgments

Thanks to Harold Szu for helpful suggestions and Lin-Ching Chang for motivating discussions.

Conflicts of interest

The author declares no conflicts of interest.

Funding

None.

References

1. Robert Plutchik. Emotion: Theory, research, and experience: Theories of emotion, 1st ed. New York: Academic; 1980.
2. Saif M Mohammad, Peter D Turney. Emotions Evoked by Common Words and Phrases: Using Mechanical Turk to Create an Emotion Lexicon. Association For Computational Linguistics, Proceedings Of The NAACL HLT 2010 Workshop on Computational Approaches to Analysis and Generation of Emotion in Text. 2010. p. 26–34.
3. Saif M. Mohammad, Peter D Turney. Crowdsourcing a Word-Emotion Association Lexicon. Computational Intelligence. *Wiley*. 2012; 29(3):436–465.
4. Saif M Mohammad, Felipe Bravo Marquez. WASSA-2017 Shared Task On Emotion Intensity. Association For Computational Linguistics, Proceedings Of The 8Th Workshop On Computational Approaches To Subjectivity, Sentiment And Social Media Analysis, 2017. p. 34–49.
5. Saif M. Mohammad, and Felipe Bravo-Marquez. Emotion intensities in Tweets. In Proceedings of the 6th Joint Conference on Lexical and Computational Semantics. Vancouver, Canada: Association for Computational Linguistics; 2017. p. 65–77.
6. Soo-Young Lee, Harold Szu. Design of Smartphone to capture subtle emotional behavior. *MOJ App Bio Biomech*. 2017;11(1):1–6.
7. Jacob Devlin, Ming-Wei Chang, Kenton Le, et. Al. BERT: Pre-training of Deep Bidirectional Transformers for Language Understanding. arXiv: 1810.04805v2, 24 May, 2019.
8. Shay Palachy, awesome-twitter-data, GitHub repository, 2020.
9. Druhmil Mehta, russian-troll-tweets, GitHub repository, 2019.
10. Harold Szu, Jeffrey Jenkins, Charles Hsu, et. al. Digging for knowledge. Proc. SPIE 7343, ICA, Wavelets, Neural Networks, Biosystems, and Nanoengineering VII, 734304, 19 March. 2009.
11. Jeffrey Jenkins, Rutger van Bergem, Charles Sweet, et. Al. e-IQ and IQ knowledge mining for generalized LDA. Proc. SPIE 9496, ICA, Compressive Sampling, Large Data Analyses (LDA), Neural Networks, Biosystems, and Nanoengineering XIII, 94960C 20 May. 2015.
12. Jeffrey Jenkins, Charles Sweet, James Sweet, et al. Authentication, privacy, security can exploit brainwave by biomarker. Proc. SPIE 9118, ICA, Compressive Sampling, Wavelets, Neural Net, Biosystems, and Nanoengineering XII, 91180U 19 June. 2014.
13. Anthony Zador. A Critique of Pure Learning and What Artificial Neural Networks Can Learn from Animal Brains. Nature Communications. 2019;10(1).