Sentiment Classification of Social Media Interactions - Project Status Report

Abstract

A flexible polarity-sentiment and content-type classification machine learning algorithm can be used to help identify 1) patterns of personal satisfaction and overall well-being over time 2) at-risk friends who may be suffering from depression. A cross-platform sentence sentiment machine learning classifier was created using an ensemble method of tf-idf SVM linear kernel classifiers each focused on a different aspect of sentiment.

1. Application

Two final applications emerge from the underlying algorithm: 1) the tracking of ones personal mental health over the course of a day, and day-to-day based upon analysis of their correspondence with others 2) A social media plug-in based on the Facebook Messenger API that allows end users to be more conscientious of aware of the mental state of their friends and loved ones.

1.1. Motivation

According to psychologist Albert Mehrabian in his book Silent Messages, 93 percent of communication in a face-to-face interaction is actually non-verbal (i.e. body language, tone of voice, facial expressions, etc.)(Mehrabian, 1981). Whether this exact value is accurate or not, it stands that a great deal of information is lost in text based communication whether it be sms texting, tweeting, facebook messaging, email, or other forms of communication. What's more, many people struggle with accurately evaluating friend and partners emotional state - the so called emotional intelligence.

2. Survey of Related Work

Sentiment classification of text is hardly a novel concept. An early foray into the subject was undertaken by Bo Pang, Lillian Lee, and Shivakumar Vaithyanathan in 2002 using Naive Bayes, maximum entropy, and support vector machines. Pang et. al. were able to achieve 70-80 percent success on the classification of product reviews (Bo Pang & Vaithyanathan, 2002). The intention of this project is to improve on those concepts using newer text classification ideas such as TF-IDF to augment the process for higher accuracy, and apply the process to the more novel area of evaluating mental and emotional health and states of individuals.

3. Approach

3.1. Data

The data are collected from a variety of sources to enable a robust ensemble classifier. The sources come from the following institutions, studies, and resources: (Gao & Sebastiani), (John Blitzer, 2007), (Hu & Liu, 2004), (et. al, 2015), (Bing Liu & Cheng, 2005), and (Bo Pang & Vaithyanathan, 2002). The data span across multiple sources including SMS, Twitter, movie reviews, product reviews, and generic sentence classification into different types of content ranging from money issues to physical pain. The training data have three different types of classification: 1) discrete polarity classification (i.e. positive, neutral, or negative), 2) numeric, continuous polarity classification (i.e. a rating in [-5,5]), and 3) a content-type classification (i.e. a label from a set including money issue, physical pain, etc).

3.2. Data Processing and Feature Generation

3.3. Algorithm Steps

The algorithm works in two major phases for each testing entry: 1) identification of the sentiment of the entry and 2) identification of the most likely content-type classification of the entry. For part 1, polarity sentiment classification, the entry is compared to each of the training data sets to determine its relevance to each training set. For instance, for a Facebook Message from the Messenger API, the relevance to the SMS training set will be higher than the relevance to a movie review training set. The relevance is captured by a weighted coefficient. Then, the entry runs through a tf-idf SVM linear kernel for each training set. The output classification for each individual training set is then multiplied by the associated weighted relevance coefficient to determine the final polarity sentiment prediction. For part

University of Pennsylvania, CIS 419/519 Course Project.

Algorithm 1 Text Sentiment And Content Ensemble Classification

Input: data x, size m, numOfEnsembleClassifiers n
Initialize sentiments = array of n vectors each of size m.
Initialize weights = array of n 0s
for $i = 1$ **to** n **do**
  Compute relevance of x to the ith ensemble classifier and set weights[i] to this weighted relevance score
  Run x through the ith tf-idf SVM linear kernel and set the classification predictions into sentiments[i]
end for
Set $weightedSentimentsPredictions$ to the weighted average predictions based on $sentiments$ and $weights$
Initialize $contentPredictions$ to a vector of size m
Run predictions through the tf-idf SVM linear kernel for content predictions and set output into $contentPredictions$
Set $contentFrequency$ to a hashmap that describes the frequency of each content category over $contentPredictions$
Set $mostNegative$ to an array of the top 10 percent most negative entries.
return $weightedSentimentPredictions$, $contentPredictions$, $contentFrequency$, and $mostNegative$

2, content-type classification, each entry is simply run in a tf-idf SVM linear kernel based on the content-type training data from Qatar Computing Research Institute.

4. Results

4.1. Performance

As of the mid-project report, the current state of the product is a simple UI in which the user is prompted to enter a sentence and that sentence is then classified as either positive, neutral, or negative based on the aforementioned ensemble of tf-idf SVM linear kernels.

5. Conclusions

Currently, the vast majority of research and data collection for the project is complete. Additionally, the data has been partially cleaned, parsed, and categorized. Using subsets of the indented final datasets, a prototype sentiment classifier has been build that can take incoming user inputs and classify them as the user types. The next steps are to select the subcategories of emotion or mental health we wish to classify and expand on the framework to have a classifier for each metric of emotional health. Once this is complete, and the various classifiers may be combined using an ensemble method.

Acknowledgments

“None.”

References

Bing Liu, Minqing Hu and Cheng, Junsheng. Opinion observer: Analyzing and comparing opinions on the web. In *Proceedings of the 14th International World Wide Web conference (WWW-2005)*, pp. 168–177, Chiba, Japan, 2005. ACM. ISBN 1-58113-888-1. doi: http://doi.acm.org/10.1145/1014052.1014073.

Bo Pang, Lillian Lee and Vaithyanathan, Shivakumar. Thumbs up? sentiment classification using machine learning techniques, proceedings of emnlp 2002, 2002.

et. al, Kotzias. From group to individual labels using deep features, 2015.

Gao, Wei and Sebastiani, Fabrizio.

Hu, Minqing and Liu, Bing. Mining and summarizing customer reviews. In *KDD ’04: Proceedings of the tenth ACM SIGKDD international conference on Knowledge discovery and data mining*, pp. 168–177, New York, NY, USA, 2004. ACM. ISBN 1-58113-888-1. doi: http://doi.acm.org/10.1145/1014052.1014073.

John Blitzer, Mark Dredze, Fernando Pereira. Biographies, bollywood, boom-boxes and blenders: Domain adaptation for sentiment classification, 2007.

Mehrabian, A. *Silent messages: Implicit communication of emotions and attitudes*. Belmont, CA: Wadsworth (currently distributed by Albert Mehrabian, email: am@kaaj.com), 1981.