Mapping Emotional Landscapes of Fiction Using Machine Learning Techniques

Nicholas Levitt[1]

[1]Department of Computer Science, College of Charleston, Charleston, SC

MOTIVATION AND INTRODUCTION

The perception of emotion in spoken language is largely based on context clues: facial expressions, body language, volume, etc. In written language however, none of these clues exist, there are just words on a page.

If a model can be trained to predict emotion experienced based off only written words, then theoretically that model could be applied to works of literary fiction to map the emotional landscape of a story.

This experiment set out to try to employ machine learning techniques to simulate the perception of emotion in written language, more specifically, the perception of joy, fear, anger, sadness, disgust, shame and guilt in works of literary fiction, and visualize the results.

METHODS

The International Survey On Emotion Antecedents And Reactions (ISEAR)[1] is a study done by Affective Sciences in which over 3000 persons were asked to report on situations in which they had experienced one of seven major emotions. The dataset is composed of 7666 individual observations. Each observation consists of the emotion experienced, a brief recount of the situation, and over 40 other variables such as sex, age and country of the reporter, intensity of the emotion felt and reaction to the situation.

This experiment focused only on the emotion experienced and the testimony of the situation. All testimonies are told in first-person and rarely span more than three sentences:

> "At the butcher's I saw an animal which had just been slaughtered; blood was dripping on the floor. "
> "New year's eve 1983/1984, I met my girlfriend. We stood on the steps outside her parents home and I kissed her for the first time."

All models were coded for in Python and trained and tested on the ISEAR dataset.

Naïve Bag-Of-Words

The first model tested was a bare-bones bag-of-words, trained on two-thirds of the data and tested on the remaining one-third. The algorithm partitions the ISEAR dataset by emotion and parses through each individual emotional dataset creating a dictionary of counts. The algorithm then parses through each individual sentence in the testing set and, using the dictionary of counts, tallies how many times each word appeared in each emotional dataset. The sentence as a whole is then predicted to be whichever emotion occurs most frequently:

> "When I got admission in M.Sc (I) Organic Chemistry, I was very happy."
> [Joy, Fear, Anger, Sadness, Disgust, Shame, Guilt] = [**154**, 49, 65, 38, 34, 36, 49] → Joy

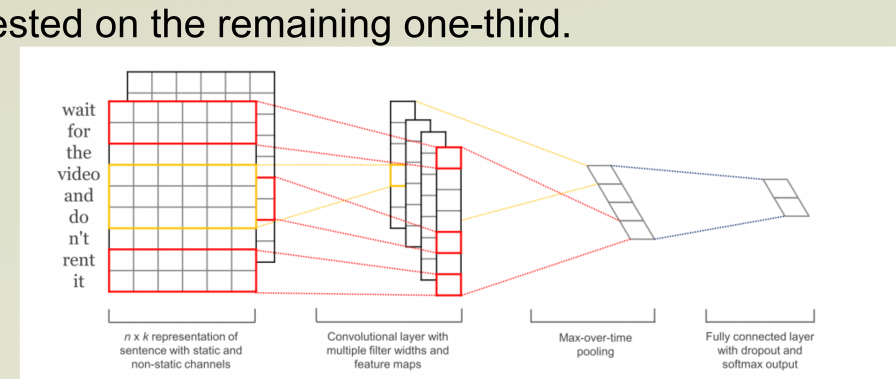
Support Vector Machine

The second model tested was a support vector machine (SVM). Using the scikit-learn module[2] a pipeline is built to tokenize the situations from the ISEAR dataset and build a dictionary of counts, transform the counts from occurrences to frequencies if applicable, and finally build a classifier using the training data.

The SVM is trained and tested using a 10-fold cross validation, with parameter tuning for determining if the model should be trained on every word individually or on pairings of two adjacent words, if the model should transform counts to frequencies, and if the model should include a penalty parameter of either .001 or .01.

Convoluted Neural Net

The third model tested was a convoluted neural net (CNN). The code is built upon an initial implementation of the TensorFlow module designed by Denny Britz and made available on Github[3], it is based loosely off of Yoon Kim's CNN from his paper *Convolutional Neural Networks for Sentence Classification*[4]. The model was trained on two-thirds of the data and tested on the remaining one-third.



Kim, Y. (2014). Convolutional Neural Networks for Sentence Classification

RESULTS

ACCURACY		
BAG-OF-WORDS	SVM	CNN
0.470	0.580	0.511

Figure 1.

As can be seen from figure 1, the SVM outperformed the other two models with an accuracy rate of 0.580, roughly four times better than random chance (~.143). As a result, all further testing and visualization has been done using only the SVM model.

Some of the results are interesting from a semantics point-of-view. For example, as can be seen in Figure 2, joyful statements are misclassified more often as sadness than as any other emotion, and vice-versa. Shame and guilt are the most often misclassified emotions, both most commonly misclassified as anger. This may show an inter-relatedness between these three emotions in terms of how they are viewed from a first-person point of view; shame and guilt may be a sort of anger at the self.

JOY MISCLASSIFICATION RATE (0.261)					
FEAR	ANGER	SADNESS	DISGUST	SHAME	GUILT
0.203	0.217	0.377	0.101	0.043	0.058

FEAR MISCLASSIFICATION RATE (0.327)					
JOY	ANGER	SADNESS	DISGUST	SHAME	GUILT
0.281	0.229	0.198	0.146	0.063	0.083

ANGER MISCLASSIFICATION RATE (0.463)					
JOY	FEAR	SADNESS	DISGUST	SHAME	GUILT
0.183	0.206	0.238	0.183	0.103	0.087

SADNESS MISCLASSIFICATION RATE (0.477)					
JOY	FEAR	ANGER	DISGUST	SHAME	GUILT
0.280	0.243	0.229	0.075	0.056	0.047

DISGUST MISCLASSIFICATION RATE (0.463)					
JOY	FEAR	ANGER	SADNESS	SHAME	GUILT
0.181	0.153	0.386	0.130	0.107	0.042

SHAME MISCLASSIFICATION RATE (0.648)					
JOY	FEAR	ANGER	SADNESS	DISGUST	GUILT
0.190	0.151	0.310	0.120	0.120	0.108

GUILT MISCLASSIFICATION RATE (0.657)					
JOY	FEAR	ANGER	SADNGESS	SHAME	DISGUST
0.142	0.133	0.357	0.174	0.084	0.110

Figure 2.

VISUALIZATION

The final step of this experiment was to take the machine learning techniques and apply them to works of literary fiction. All novels and stories were taken from Project Gutenberg[5].

In order to visualize the emotional landscape, the SVM was first retrained on the entire ISEAR dataset in order to be as robust as possible. Fictional works were then partitioned into fifty sections and passed to the SVM. The output is a numerical representation of how strong the model perceives each of the seven emotions to be in each of the fifty sections, which can then be graphed. All graphing was done using the ggplot2 package for R. In order to make the visualizations more interactive, a R-Shiny application was built so a user can pick any of the available stories and choose any subset of the seven emotions to be visualized.

As can be seen in figure 3, on their own as data points these values don't appear to give any information, so a line graph joining them is built. The basic line graph of these values tends to be overly-acute and hard to follow though, so a loess function is applied to smooth out the plots.

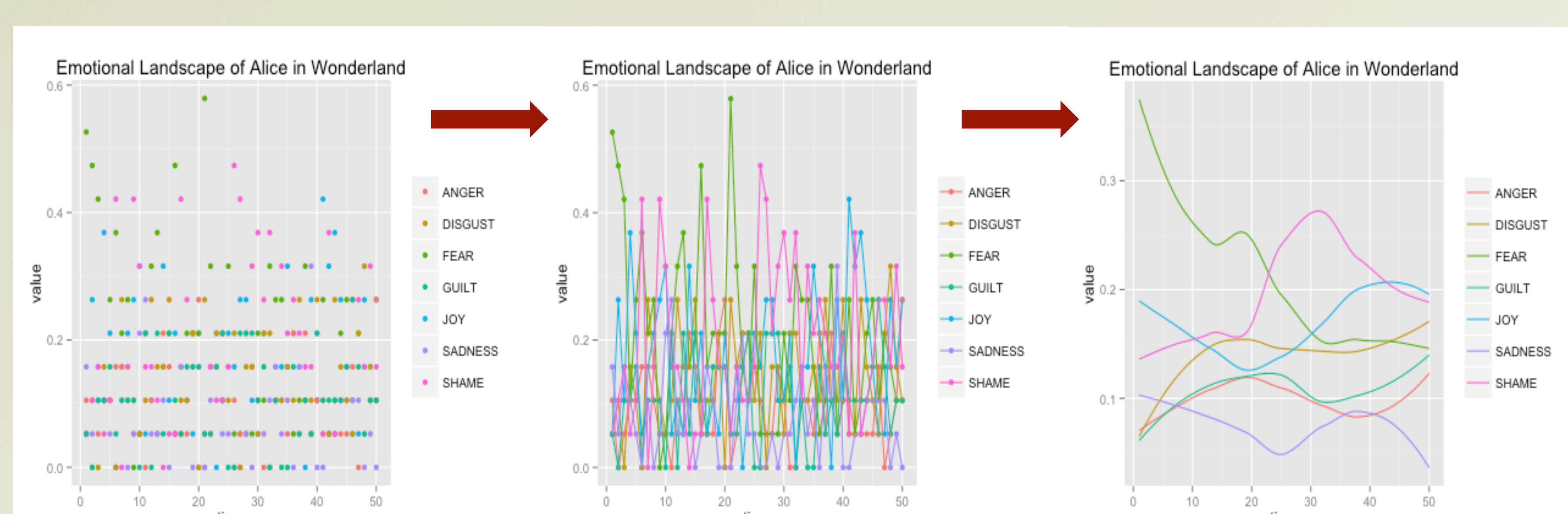


Figure 3.

With the visualizations, the writing styles of different authors can be viewed in a unique way. For example, it can be seen in Figure 4 that H.G. Wells instilled in both *War of The Worlds* and *Time Machine* an overwhelming amount of fear and a secondary serving of disgust, with the 5 other emotions maintaining rather low intensities through the stories.

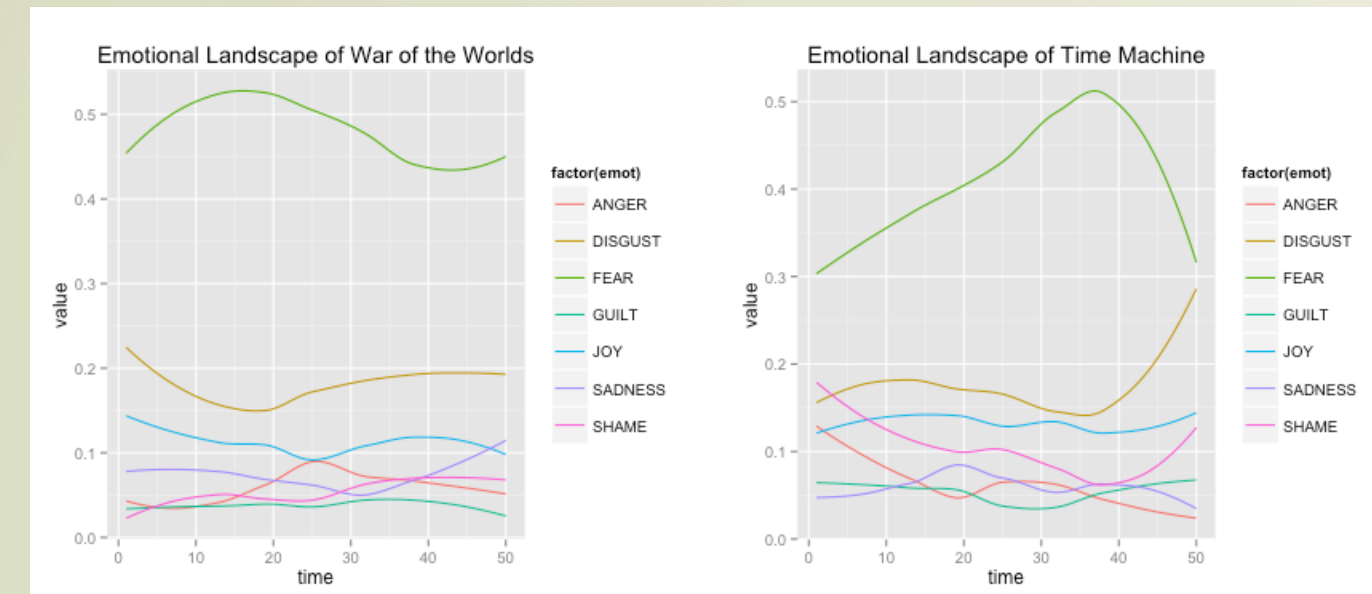


Figure 4.

As would be expected, stories that are widely considered to be more complex, such as *Anna Karenina* by Leo Tolstoy or *Alice in Wonderland* by Lewis Carroll, have more complex visualizations, with emotions overtaking one another for prominence in any given section of the story, as can be seen in figures 3 and 5.

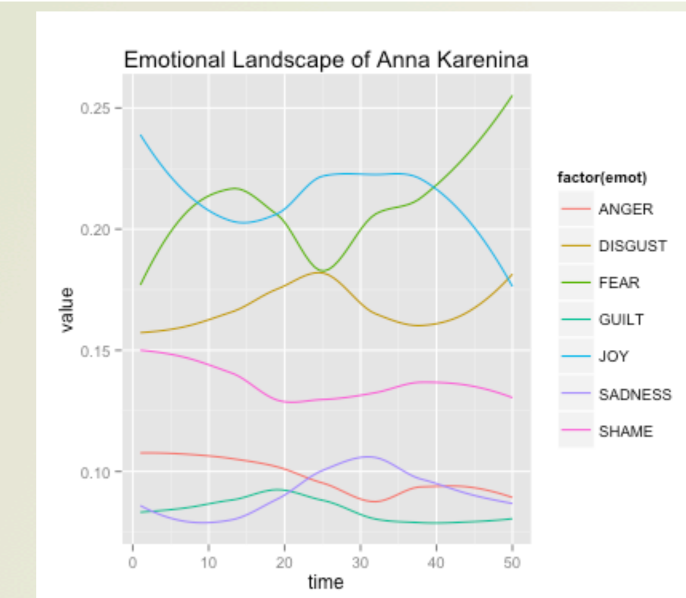


Figure 5.

DISCUSSION AND FUTURE WORK

This experiment gives an interesting insight into the abilities of different machine learning techniques when tasked with multi-class sentiment analysis. The convoluted neural net as a technique has been garnering a lot of attention and support recently for its ability to complete a myriad of tasks, but this experiment shows that there really is no free lunch.

There are many ways this experiment could be improved upon, some more intensive than others. Obviously it would be ideal to raise the accuracy of the models when testing on the ISEAR dataset. This could be done by employing more pre-processing of the text before passing it to the models. For example, part-of-speech (POS) tagging could add more insight, as could looking at the text as trigrams rather than as individual words or bigrams.

It would also be interesting to find a different dataset with labeled emotional text, preferably one that is not composed of only first-person testaments. As not all fiction is written in first person, it would be more appropriate to train a model on a diverse dataset.

Lastly, it is hard to validate the results of the visualizations because perception of emotion in literary fiction is an amorphous concept, it changes from reader to reader. Whereas one reader may view *Alice in Wonderland* as a lighthearted children's story, another may view it as a darker, more adult-centric novel. There is not much that can be done about this fact outside of going through a novel sentence by sentence and labeling their emotions in a supervised manner to be used in testing the machine learning algorithms.

REFERENCES

1 *International Survey On Emotion Antecedents And Reactions (ISEAR).* Affective Sciences, n.d. Web. Spring 2016.

2 Cournapeau, David, and Et Al. *Scikit-learn.* Computer software. *Scikit-learn.* Vers. 0.17.1. N.p., n.d. Web. Spring 2016.

3 Britz, Danny. *Cnn-text-classification-tf.* Computer software. N.p., n.d. Web. Spring 2016.

4 Kim, Yoon. "Convolutional Neural Networks for Sentence Classification." *EMNLP* (2014): n. pag. *Arxiv.* Web.

5 "Project Gutenberg." *Project Gutenberg.* N.p., n.d. Web. Spring 2016.

ACKNOWLEDGEMENTS

Thank you to Dr. Anderson and the rest of the College of Charleston Computer Science Department