Suicidal Tendencies: The Automatic Classification of Suicidal and Non-Suicidal Lyricists Using NLP

Matthew Mulholland
Montclair State University
Montclair, NJ, USA
mulhollandm2@mail.montclair.edu

Joanne Quinn
Montclair State University
Montclair, NJ, USA
quinnj11@mail.montclair.edu

Abstract

Can natural language processing be used to predict the likelihood that a musician will commit or has committed suicide? In order to explore this idea, we built a corpus of songs that includes a development set, a training set, and a test set, all consisting of different lyricists. Various vocabulary and syntactic features were then calculated in order to create a suicide/non-suicide song classifier. The features were input into the Weka machine learning suite and tested with an array of algorithms. We were able to achieve up to a 70.6% classification rate with the SimpleCart algorithm, a 12.8% increase over the majority-class baseline. Our findings suggest that syntactic and vocabulary features are useful indicators of the likelihood that a lyricist will commit or has committed suicide.

1 Introduction

Recently, research into the application of NLP to the detection of health illnesses has proved fruitful. For instance, in their study of the effects of dementia on writers, Le et al. (2011), guided by previous research, explored various hypotheses in the novels of three British writers. Their research found a decline in the type-token ratio of the novelists suffering from dementia. The use of passive constructions was also explored since it is generally believed that it represents a syntactic structure that is particularly complex and likely would be used less often by writers suffering from dementia. Indeed, they found that authors with dementia use less passives than their healthy peers. Their results indicate the potential for natural language processing in the language of mental illness. Similarly, much research into the application of NLP to depression and suicide prediction has been conducted in recent years (Pestian et al. 2012; Sohn et al. 2012). While one might not expect depression or a suicidal tendency to affect language in the same way as an illness such as dementia, it is reasonable to assume that there will be textual indications of these mental illnesses also. In Stirman and Pennebaker (2001), word use is treated as an indicator of the mental states of suicidal and non-suicidal poets. Stirman & Pennebaker developed the Linguistic Inquiry and Word Count (LIWC) program to analyze over 70 language dimensions, including: polarity, affect states, death, sexuality, tense, etc. Their research found a correlation between the likelihood of a poet committing suicide and his/her disengagement from society based on the suicidal poets' heavy use of first-person singular pronouns and decreased use of the first-person plural pronoun. Interestingly, they also noted that poets who had committed suicide generally used more sexual words and references to death than their non-suicidal counterparts. Additionally, latent semantic analysis has been used to detect depression in free texts (Neuman et al. 2010). By creating a semantic field from the words commonly associated with the concept of depression, researchers were able to accurately identify depressed people through their writing. Finally, other pertinent research involves the use of concrete nouns and the lack of abstract concepts in professionally-written poetry (Kao & Jurafsky, 2012). Based on the prior research, we anticipated that lyricists who have committed suicide would be more likely to use self-references than referring to others. We also hypothesized that certain features, like the use of sensual and morbid words, would be more prevalent in the works of suicidal lyricists. Additionally, we were interested in exploring the differences among other features such as TTR, the number of passives, the n-gram profiles of songs in relation the n-gram profiles of suicide/non-suicide lyricists, and the semantic fields of other target emotions and affect states of the two groups.

2 Methods

First, a corpus of songs by male suicide and non-suicide lyricists was constructed, which consists of a development set, a training set, and a test set. The development set includes five non-suicide artists (a total of 94 songs) and six suicide artists (a total of 74 songs). The training set is comprised of 533 songs, of which 253 were written by four lyricists who have not committed suicide. The remaining 280 songs were written by five lyricists who did commit suicide. Finally, the test set contains 63 songs written by five non-suicide artists and 46 songs written by 4 suicide artists. In our search for lyricists who committed suicide, we looked for lyricists who met the following prerequisite: the suicide had to be relatively unambiguous. Although we did not require that each song be solely written by the lyricist in question due to the fact that it is often murky who to attribute the lyrics to, we did make an attempt to exclude songs written entirely by bandmates or other musicians. Due to the lack of female lyricists who committed suicide, we were forced to consider exclusively male lyricists. 48% of the non-suicide corpus consists of songs written by one artist, Bob Dylan. Removal of 35% of those Dylan songs (and thus evening out the distribution of songs) did not significantly alter the classifier's results. The other non-suicide lyricists contributed between 32-45 songs each. The training set of suicide songs is slightly over-represented by Elliot Smith (about 33% of the songs belong to him). The four remaining suicide artists each contributed between 11% and 20% of all songs in this set. Table 1 displays the composition of the training and test sets. We additionally used a development set of five non-suicide lyricists (94 songs) and six suicide lyricists (74 songs), which was used to compute n-gram features. Each song considered was searched for in online lyrics databases and was cleaned by hand, the lines being joined into punctuated sentences or phrases so that a POS-tagger and lemmatizer could be used. The lyrics were then tokenized using the OpenNLP tokenizer and lemmatized. Features were computed using Python and with some help from the UAM corpus tool (O'Donnell 2008)(which uses the Stanford Parser), especially for grammatical analysis.

Artist	Class	Set	Songs
Bob Dylan	non-suicide	training	123
Bob Marley	non-suicide	training	42
Mike Ness	non-suicide	training	43
Trent Reznor	non-suicide	training	45
Elliott Smith	suicide	training	99
Ian Curtis	suicide	training	40
Kurt Cobain	suicide	training	53
Pete Ham	suicide	training	34
Phil Ochs	suicide	training	54
Ben Folds	non-suicide	test	11
Chris Bell	non-suicide	test	12
John Lennon	non-suicide	test	20
Neil Young	non-suicide	test	10
Paul Simon	non-suicide	test	10
Doug Hopkins	suicide	test	4
Peter Bellamy	suicide	test	18
Richard Manuel	suicide	test	10
Tom Evans	suicide	test	14

Table 1: Training and Test Set Artist List

3 Features

In order to create the suicide/non-suicide lyrics classifier, similar features to those used in the previous research were explored in conjunction with a set of original features. In all, there were 87 features.

3.1 Vocabulary Features

Used in many text classification systems to give some indication of the vocabulary size are features involving the number of types and tokens. While a few of our features were based on types and tokens, they, along with length of song, are mostly used to normalize many of the features in the following two sections. Below are the features we explored:

- typ: type/token ratio (TTR) and type/time ratio
- tok: token/time ratio

3.2 Syntactic Features

As in Le et al. (2011) and Stirman & Pennebaker (2001), we expected to find differences in the use of the passive constuction and in the proportions of the first-person pronouns to the rest of the pronouns. We expected a greater use of passive constructions in the lyrics of the suicide lyricists in comparison to the non-suicide lyricists since it might signify a greater sense of disengagement from the external world. Additionally, we hypothesized a higher proportion of first-person pronouns to other pronouns in the suicidal lyrics, since the common perception about suicide cases and depressive people in general is that they are more self-centered or that they are less concerned with others.

In addition to the exploration of the first-person pronouns and passive constructions, we also looked at the differences in the usage of mental-state verbs, including the use of verbs like *think* and *feel*, coupled with the first-person singular pronoun. Our expectation here was that the suicide writers might use such constructions more often due, perhaps, to a preoccupation with the thoughts and feelings and an inclination to think and feel more often than act.

These features were all computed using the UAM corpus tool, which makes use of the Stanford parser. The UAM tool allows one to create autocoding rules and it presents annotation statistics and counts in a somewhat convenient way on the text level. The syntactic features were all normalized by type, token, and time. In cases where the counts in the denominator frequently contained zeroes and thus would result in undefined values, we simply added 0.01 to the denominator.

3.3 Semantic Class Features

Our expectations about the content of the suicide lyrics in comparison to the non-suicide lyrics was that they would deal with more negative, depressing subjects than positive ones. We also hypothesized that, as in Pennebaker & Stirman (2001), we would see a difference in the use of sexual words in the suicide lyrics (specifically, a heightened rate of sexual word usage and death related word usage). To measure the rates of negativity, neutrality, and positivity, we used the MPQA prior polarity word lexicon (Weibe et al. 2005). We also compiled our own dictionaries of sensual words, action verbs, concrete nouns, and death-, love-, depression, and drug-related words using a (simplified) collection method based on that used by Neuman et al. (2010) in their research on depression detection. We then calculated the number of occurrences of each feature within the text. As with the syntactic features, these measurements were also normalized by type, token, and time. Again, there were some ratios whose denominators contained zeroes and we dealt with these cases in the same manner as above, adding 0.01 to the denominator. Additionally, we used the AFINN (Nielsen, 2011) word-valence resource to calculate the total and average polar value of each song.

3.4 N-gram Features

A Python script was written to build unique n-gram (for n = 1 to 6) profiles for the classes in the development set. (This set was used exclusively for this purpose and was not needed for the other syntactic and semantic features.) The n-gram profiles of the songs in the training and test sets were compared to the suicide and non-suicide n-gram profiles constructed from the development set to calculate features such as the percentage of bigrams in a song that overlap with either of the two bigram profiles from the development set.

a	b	<– classified as
29	17	a = suicide
15	48	b = non-suicide

Table 2: Confusion Matrix.

4 Results

After computing the features in the previous sections, we input all features into Weka and ran a number of different ML algorithms to create a classifier. As a baseline for comparison, we used the majority-class prediction rate of 57.7%. The classifier was trained on the training set and then tested on the test set of different artists' songs. The most successful algorithm was the SimpleCart algorithm, which correctly classified the songs as either suicide or non-suicide 70.6% of the time and which achieved a 0.39 Cohen's kappa value. The correct classification rate represents an increase of 12.9% over the baseline. The SimpleCart algorithm achieved a precision of 0.71, recall of 0.71, an F-measure of 0.71, and an ROC Area value of 0.70. In Table 4 above, we report the confusion matrix for the test set. As we computed such a great number of features, our expectation was not that they would all be useful. All algorithms were run on the subset of features chosen by Weka. The subset consisted of a combination of vocabulary, syntactic, and semantic features in addition to the n-gram features.

5 Discussion

Regarding the first-person sing. + plural and first-person sing. features, the results did not appear to match our expectations, which were that the suicidal lyricists would use more first-person pronouns. As for the first-person to second-/third-person and first-person sing. to first-person plural pronouns features, there is no indication that the suicide lyricists used more first-person singular pronouns in relation to first-person plural pronouns. While the results for the passive/time feature seemed to align with our expectations about the suicidal writers, the passive/typ and passive/tok features either did not show a significant difference or were contrary to our expectations (though they were still used in Weka). The suicide lyricists did appear to use more mental-verb syntactic constructions and, though both categories seemed to be more negative than positive, the suicide songs seemed to be even more negative in relation to positive terms. In regard to the use of sensual words, it appears that the non-suicidal lyricists actually used sensual words more often than the suicidal writers, maybe not significantly in relation to tokens, but more so in relation to types. With respect to types, the suicidal writers appeared to use less action verbs, but with respect to tokens the difference between the groups is not significantly different. Furthermore, in relation to tokens, the suicidal writers seemed to use less concrete nouns. These last two findings might indicate a greater reliance on static and non-descriptive writing. However, the suicide lyricists did appear to have a significantly higher TTR, which might indicate a more expansive vocabulary, which would seem to be at odds with the last statement. In terms of death- and drug-themed token count features, it did not appear that the suicide lyricists definitely used more drug- and death-related terms; however, with regard to the love-related token count feature, it seems that the suicide lyricists did write about love less often.

6 Further Research

Besides expanding the corpus to include many more non-suicide lyricists and (at the very least) to include more songs from each of the suicide artists, it would perhaps be fruitful to extend the analysis to other types of features and new lexicons since it has been demonstrated that this task could be solved using NLP.

Acknowledgments

We would like to take this space to acknowledge again the use of the UAM corpus tool, which proved valuable to our analysis, the MPQA prior polarity lexicon, the AFINN word valence resource, the Weka machine learning software, and the many online contributors of lyrics and word-lists. We would also like to thank Michael Flor for letting us use a lemmatizer that he built for his own use.

This material is based in part upon work supported by the National Science Foundation under Grant Numbers 0916280 and 1048406. Any opinions, findings, and conclusions or recommendations expressed in this material are those of the authors and do not necessarily reflect the views of the National Science Foundation.

References

Justine Kao and Dan Jurafsky 2012. A Computational Analysis of Style, Affect, and Imagery in Contemporary Poetry. *NAACL-HLT* 2012, 8.

Xuan Le, Ian Lancashire, Graeme Hirst, and Regina Jokel. 2011. Longitudinal detection of dementia through lexical and syntactic changes in writing: A case study of three British novelists. *Literary and linguistic computing*, 26(4):435-61.

Y. Neuman, Y. Cohen, G. Kedma, and O. Nave. 2010. Using Web-Intelligence for Excavating the Emerging Meaning of Target-Concepts. *2010 IEEE/WIC/ACM International Conference on Web Intelligence and Intelligent Agent Technology*, Toronto, Aug-Sept 2010. IEEE Computer Society.

F. Nielsen. 2011. A new ANEW: Evaluation of a word list for sentiment analysis in microblogs. *Proceedings of the ESWC2011 Workshop on Making Sense of Microposts: Big things come in small pack- ages,* number 718 in CEUR Workshop Proceedings, Heraklion 2011.

Mick O'Donnell 2008. emonstration of the UAM CorpusTool for text and image annotation. *2010 Proceedings of the 46th Annual Meeting of the Association for Computational Linguistics on Human Language Technologies Demo Session - HLT* Association for Computational Linguistics, 2008.

J. P. Pestian, P. Matykiewicz, and M. Linn-Gust. 2012. Whats in a Note: Construction of a Suicide Note Corpus. *Biomedical Informatics Insights*, 2012:5 1-6.

M. F. Porter 1980. An algorithm for suffix stripping. *Program*, 14(3):130137.

S. Sohn, M. Torii, D. Li, K. Wagholikar, S. Wu, and H. Liu. 2012. A Hybrid Approach to Sentiment Sentence Classification in Suicide Notes. *Biomedical Informatics Insights*, 2012:5 (Suppl. 1) 43 50.

S. W. Stirman and J. W. Pennebaker. 2001. Word Use in the Poetry of Suicidal and Nonsuicidal Poets. *Psychosomatic Medicine* 2001, 63:517-522.

Janyce Wiebe, Theresa Wilson, and Claire Cardie. 2005. Annotating expressions of opinions and emotions in language. *Language Resources and Evaluation* , 39(2-3):165210.

Ian H. Witten and Eibe Frank. 1999. Practical Machine Learning Tools and Techniques with Java Implementations. Morgan Kauffman (1999).