OJPHI

ISDS 2015 Conference Abstracts

Identifying Depression-Related Tweets from Twitter for Public Health Monitoring

Danielle Mowery*[1], Hilary A. Smith[2], Tyler Cheney[2], Craig Bryan[2] and Michael Conway[1]

[1]Biomedical Informatics, University of Utah, Salt Lake City, UT, USA; [2]Psychology, University of Utah, Salt Lake City, UT, USA

Objective

We aim to develop an annotation scheme and corpus of depression-related tweets to serve as a test-bed for the development of natural language processing algorithms capable of automatically identifying depression-related symptoms from Twitter feeds.

Introduction

Major depressive disorder has a lifetime prevalence of 16.6% in the United States. Social media platforms – e.g. Twitter, Facebook, Reddit – are potential resources for better understanding and monitoring population-level mental health status over time. Based on DSM-5 [1] diagnostic criteria, our research aims to develop a natural language processing-based system for monitoring major depressive disorder at the population-level using public social media data.

Methods

In this pilot study, three annotators - two psychology undergraduates (A1, A2) and a postdoctoral biomedical informatics researcher (A3) - annotated 900 tweets using a linguistic annotation scheme based on DSM-5 depression criteria (e.g. *anhedonia* – "*I don't enjoy singing anymore*") [2]. We report agreement between annotator pairs computing F-score, a surrogate for kappa. Finally, we trained and tested three machine learning classifiers – *support vector machine, naïve bayes,* and *decision tree* – for predicting two depression-related classes: 1) whether a tweet represents **clinical evidence of depression** or **not** and 2) if the tweet is depression-related, whether it is classed as *low mood, fatigue or loss of energy,* or *problems with social environment.* We trained and tested each classifier using the Weka toolkit (v.3.6.8) with 10-fold cross validation using unigram features, and then reported classifier performance compared against 884 adjudicated tweets with sensitivity and positive predictive value.

Results

We observed high agreement between annotator pairs: A1/A2: 81%, A1/A3: 76%, and A2/A3: 78%. Of the 884 adjudicated tweets, the majority of tweets represented **no clinical evidence of depression** (n=635; 72%) then **clinical evidence of depression** (n=249; 28%). The skewed distribution of the most frequent 3 **clinical evidence of depression** symptoms/stressors ranged from *low mood* (n=114; 13%), *fatigue or loss of energy* (n=51; 6%), and *problems with social environment* (n=36; 4%). Overall, we observed comparable sensitivities and positive predictive values among classifiers for discerning whether a tweet represented **clinical evidence of depression** or **not (Table 1). No clinical evidence of depression** and *fatigue or loss of energy* can be more accurately identified than *low mood*, *problems with social environment*, and other types of **clinical evidence of depression** by all three classifiers. No one classifier performs with both superior sensitivity and positive predictive value for any one symptom/stressor, suggesting that different approaches may be necessary for reliable classification.

Conclusions

Automatically identifying depression-tweets with a moderate degree of accuracy is feasible. We are actively improving our training models leveraging a corpus of ~10,000 tweets, expanding symptoms/stressors being detected, and experimenting with richer features for symptom/stressor detection (e.g. gender, age).

Table 1. Classification performance using unigrams. SVM=Support Vector Machine; NB=Naïve Bayes; DT=Decision Tree; Sens=sensitivity; PPV=positive predictive value

	SVM		NB		DT	
	Sens	PPV	Sens	PPV	Sens	PPV
clinical evidence of depression or not	75	74	75	75	75	73
no clinical evidence of depression	84	81	83	83	90	78
clinical evidence of depression	51	56	56	56	37	60
low mood	33	41	37	36	23	50
fatigue or loss of energy	88	73	84	84	90	69
problems with social environment	6	17	14	8	0	0

Keywords

social media; mental health; Twitter; public health; natural language processing

Acknowledgments

This work was funded by a grant from the National Library of Medicine (R00LM011393) and was granted an exemption from review by the University of Utah Institutional Review Board (IRB 00076188).

References

1. American Psychiatric Association. Diagnostic and Statistical Manual of Mental Disorders, Fifth Edition (DSM-5). American Psychiatric Publishing. 2013.
2. Mowery DL, Bryan C, Conway M. Toward Developing an Annotation Scheme for Depressive Disorder Symptoms: A Preliminary Study using Twitter Data. In: Mitchell M, Coppersmith G, Hollingshead K, editors. Proceeding of 2nd Workshop on Computational Linguistics and Clinical Psychology - From Linguistic Signal to Clinical Reality; 2015 June 5; Denver, CO. Association for Computational Linguistics; c2015. p. 89-98

***Danielle Mowery**
E-mail: danielle.mowery@utah.edu

ISDS Annual Conference Proceedings 2015. This is an Open Access article distributed under the terms of the Creative Commons Attribution-Noncommercial 3.0 Unported License (http://creativecommons.org/licenses/by-nc/3.0/), permitting all non-commercial use, distribution, and reproduction in any medium, provided the original work is properly cited.