Textual Cues for Online Depression in Community and Personal Settings

Thin Nguyen(✉), Svetha Venkatesh, and Dinh Phung

Deakin University, Geelong, Australia
{thin.nguyen,svetha.venkatesh,dinh.phung}@deakin.edu.au

Abstract. Depression is often associated with poor social skills. The Internet allows individuals who are depressed to connect with others via online communities, helping them to address the social skill deficit. While the difficulty of collecting data in traditional studies raises a bar for investigating the cues of depression, the user-generated media left by depression sufferers on social media enable us to learn more about depression signs. Previous studies examined the traces left in the posts of online depression communities in comparison with other online communities. This work further investigates if the content that members of the depression community contribute to the community blogs different from what they make in their own personal blogs? The answer to this question would help to improve the performance of online depression screening for different blogging settings. The content made in the two settings were compared in three textual features: affective information, topics, and language styles. Machine learning and statistical methods were used to discriminate the blog content. All three features were found to be significantly different between depression Community and Personal blogs. Noticeably, topic and language style features, either separately or jointly used, show strong indicative power in prediction of depression blogs in personal or community settings, illustrating the potential of using content-based multi-cues for early screening of online depression communities and individuals.

Keywords: Computer mediated communication · Weblog · Social media analysis · Mental health · Textual cues · Affective norms · Language styles · Topics

1 Introduction

In their lifetime, 12.1 % and 4.1 % of people have suicide ideation and attempts, respectively [15]. Psychiatry disorders were found in 90 % of suicide victims and the relative risk of suicide among those with the conditions was increased more than 9-fold [3]. Of the disorders, depression is among the major risk factor. Knowing what are the signs of depression would help detect early the sufferers and ultimately help alleviate the suicide crisis.

The signals could be non-verbal activities, such as “lack of eye contact”, “head drooped, looking at ground”, and “mouth turned down”, which were found

J. Li et al. (Eds.): ADMA 2016, LNAI 10086, pp. 19–34, 2016.
DOI: 10.1007/978-3-319-49586-6_2

as salient cues for depression [22]. Other markers of depression were vocal characteristics [11]. Social skills were also found as an indicator of depression [7]. Another aspect, language styles, such as the use of sadness or swearing words, was identified as depression cues for personal diaries and online blogs [19]. Most of these studies were conducted in small scale, such as for the study on the linguistic cues [19], the text was collected from fifty-seven participants.

On the other hand, the Internet and social media and networking systems built on it offers great sources to investigate causes and cues for depression in large scale. The new media has provided an excellent venue for individuals to express their ideas and like-minded people to share stories, including depression sufferers. This unintentionally leaves a huge data which are often difficult to be collected in traditional studies.

In this study, we collect data from *depression.livejournal.com*, the largest Live Journal community interested in "depression". Investigations into online communities have often been at the community setting, such as in social capital and mood patterns [13,18]. However, it is still questionable whether blogging by members of depression communities in personal blogs is different from community blogs. For example, *fillers* or *swearing* are expected to be used more in personal than in community pages, probably suggesting different weights for the cues when predicting depression in the two settings.

This paper examines if the content that members of the depression community contribute to the community blogs (*Community*) different from what they make in their own personal blogs (*Personal*)? In this study, the content is seen in three aspects: the topics discussed, the language styles expressed, and the affective information conveyed. These features were used as the base to detect online community [12], especially they have been found to be strong predictors of autism, and differentiate mental health communities from other online communities [14]. The features can be considered as potential text-based cues for depression.

We present an analysis of a large scale cohort of data authored by more than 4,000 members of the depression community. The content made by members of the depression community in their own personal blogs is also examined in comparison with what they make in the community blog. The way to collect data unobtrusively from online sources, such as online depression communities and individuals in this study, provides a valuable alternative approach to research in mental health, avoiding the issue of privacy as in traditional data collection for clinical studies.

The main contributions of this work are: (1) to introduce a relatively comprehensive view of the content made in depression personal and community blogs, regarding three aspects: sentiment information, language style, and topics of interest; (2) to propose an efficient approach to select features and do regression simultaneously, providing a set of powerful predictors of depression blogs in personal and community settings; and (3) to provide statistically reliable empirical evidence in a data-driven approach compared with small-scale questionnaire-based method in psychology. The result shows the potential of the new media in screening and monitoring of online depression blogging in both individual and community context. In addition, the same framework could be employed for at-risk individuals and communities. In a broad sense this work demonstrates the application of machine learning in medical practice and research.

The remainder of this paper is organized as follows. Section 2 describes the methodology. Section 3 presents the difference of blogging in personal and community settings. Section 4 shows the performance of the classification of Community vs. Personal blogs. Section 5 discusses the limitation of the work and Sect. 6 concludes the paper.

2 Method

2.1 Datasets

In this paper Live Journal data was chosen since it allows people to create or join communities of interest, along with their own personal pages. In particular, to examine blogging within online depression communities, data from the largest community in Live Journal interested in "depression" – depression.livejournal.com – was crawled. The community is described in its profile as "*a safe and open community for those experiencing depression*".

The community was founded in December 2000 and as of September 2015, it has more than 7,000 members and 40 thousand posts. This is the *Community* cohort in this study. We then construct a control dataset. The posts made in personal blogs of members of the depression community were also crawled. This is the *Personal* cohort. Only those members who have made posts in both cohorts were taken into the experiments, resulting in a corpus of 25,012 community posts and 104,033 personal posts made by 4,439 users. To create a balanced dataset, the same number of posts (the smaller number of posts, 25,012, which appeared in the Community category) for Personal category was randomly selected into the study, resulting in a corpus of 50,024 posts.

2.2 Feature Sets

To characterised the posts made in community or personal blogs, three types of features are extracted: (1) Topics: what are discussed in the posts? (2) Language styles: how the posts are expressed? and (3) Expressed emotion: the affective information is conveyed in the content.

– *Affective information:* For the affective aspect, we used ANEW lexicon [2] to extract the sentiment conveyed in the content. Words in this lexicon are rated in term of valence, arousal, and dominance. The valence of words is on a scale of one, *very unpleasant*, to nine, *very pleasant*. The arousal is measured in the same scale, one for *least active* and nine for *most active*. The dominance is also in the same scale, ranging from *submissive* to *dominant*.

- *LIWC features:* We examined the proportions of words in psycho-linguistic categories as defined in the LIWC package [17]: linguistic, social, affective, cognitive, perceptual, biological, relativity, personal concerns and spoken.[1]
- *Topics:* to extract topics, latent Dirichlet allocation (LDA) [1] was used as a Bayesian probabilistic modeling framework. LDA extracts the probabilities $p\,(\text{vocabulary} \mid \text{topic})$ - that is, words in a topic, and then assigns a topic to each word in a document. For the inference part, we implemented Gibbs inference detailed in [10]. We set the number of topics to 50, run the Gibbs for 5000 samples and use the last Gibbs sample to interpret the results.

2.3 Statistical Testing

Statistical tests were conducted to examine the difference between Community and Personal blogs in the use of each feature. For each of 50 topics, 68 linguistic categories, and three ANEW sentiment scores, two following hypotheses were tested:

- H_1: $mean_{comm} = mean_{pers}$: the null hypothesis that, for the tested feature, the data made by the two examined populations, Community vs. Personal blogs, are samples from normal distributions with equal means. This test is conducted using the two-sample t-test, a parametric test.
- H_2: $median_{comm} = median_{pers}$: the null hypothesis that, for the tested feature, the data made by the two examined populations, Community vs. Personal blogs, are samples from continuous distributions with equal medians. This test is conducted using the Wilcoxon rank sum tests, a non-parametric test.

The relative difference in the use of a feature is determined as

$$diff = \frac{mean_{comm} - mean_{pers}}{(mean_{comm} + mean_{pers})/2} * 100 \tag{1}$$

If $diff > 0$, the feature is used more in Community than in Personal, and vice versa.

2.4 Classification

Lasso as the Classifier and Feature Selector. Denote by $\mathcal{B}$ a corpus of N posts made in community or personal blogs. A document $d \in \mathcal{B}$ is denoted as $\mathbf{x}^{(d)} = [\ldots, \boldsymbol{x}_i^{(d)}, \ldots]$, a vector of features. The feature sets experimented in this work are topics, extracted through topic modeling (LDA) and language styles (LIWC). When topics are the features, $\boldsymbol{x}_i^{(d)}$ is the probability of topic i in document d. If LIWC processes are the features, $\boldsymbol{x}_i^{(d)}$ represents the quantity of the process i in document d. Our experimental design examines the effect of

[1] http://www.liwc.net/descriptiontable1.php, retrieved Sept. 2015, cached: http://bit.ly/1PPbeSv.

these two feature sets in classifying a blog post into one of two target classes. Given a document $d \in \mathcal{B}$, we predict if the document belongs to a Community or Personal blog based on the textual features $\mathbf{x}^{(d)}$.

We are interested in not only which sets of features perform well in the classification but also which features in the sets are strongly predictive of depression. For this purpose, the least absolute shrinkage and selection operator (Lasso) [9], a regularized regression, is chosen. Lasso does logistic regression and selects features simultaneously, enabling an evaluation on both the prediction performance and the importance of each feature in the classification. Particularly, in prediction of community posts, Lasso assigns positive and negative weights to features associated with community and personal posts, respectively. To the features irrelevant to the prediction, Lasso assigns zero weight. Thus, by examining its weights, we can learn the importance of each feature in the prediction.

The regularization parameter (λ) in the regression model is chosen such that it is the largest number and the accuracy is still within one standard error of the optimum (1se rule). This way prevents over-fitting since not too many features are included in the model while the accuracy of classification is still assured.

For each run, we use five-fold cross-validation, that is, one held-out fold is used for testing and other four folds for training. Accuracy was used to evaluate the performance of the classifications.

Other Classifiers. For comparison with the classification performed by Lasso, classifiers from other paradigms were also included:

- Naive Bayes (NB): one of the probabilistic methods that construct the conditional probability distributions of underlying features given a class label. The classification on unseen cases is then done by comparing the class likelihood.
- Support vector machines (SVM): a non-probabilistic binary classifier that finds the separating plane between two classes with maximal margins.
- Logistic regression (LR): a non-regularized logistic regression model, as opposed to the regularized one as Lasso.

These classifiers will perform the binary classifications of *Community* versus *Personal* posts, using LIWC, topics, and a combination of them as the feature sets. The accuracy is used to compare with that of Lasso on the same classifications.

Table 1. The mean (±std) of sentiment scores for community and personal posts.

Sentiment score	Community	Personal	h_1	h_2
Valence	5.29 ± 0.64	5.66 ± 0.62	1	1
Arousal	4.29 ± 0.34	4.23 ± 0.39	1	1
Dominance	5.33 ± 0.44	5.54 ± 0.42	1	1

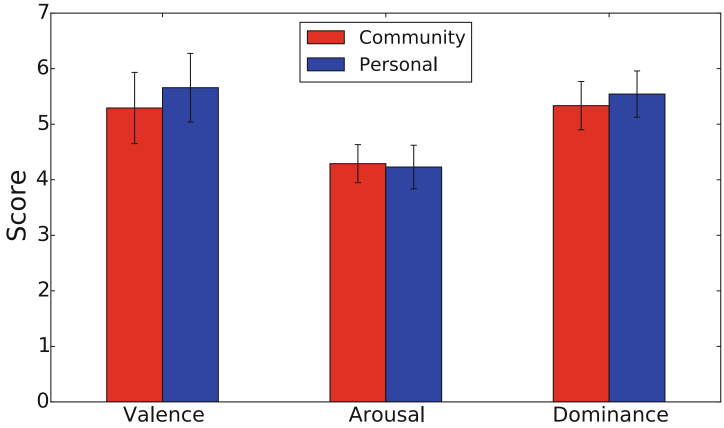


Fig. 1. The mean of sentiment scores for posts made in community and personal blogs.

3 Depression in Community and Personal Settings

In this section, the differences between *Community* and *Personal* blogs made by members of the depression community, with respect to the three feature sets, are discussed.

3.1 Affective Information

Table 1 shows the mean of sentiment scores for *Community* and *Personal* posts, accompanied with the result of t-test and rank sum test. If the result is one then the test rejects the null hypothesis of equal mean or median at the 5 % significance level. Otherwise the tests fail to reject the null hypothesis. As shown in the table, both tests rejected the null hypotheses.

Figure 1 shows that the mean valence of the community posts was lower than that of the personal posts. This means that the posts made in the community blogs have more unpleasant words and/or fewer pleasant words than those in the personal blogs.

Similarly, the average dominance score of community posts was lower than those of personal posts. It is likely that as a member of a depression community, people use more submissive and/or less dominant words, in comparison with what they write in their own pages.

On the other hand, the average arousal score of *Community* posts was slightly higher than that in *Personal* posts. This probably implies that members tend to use more high active words and/or less low active words in their posts.

3.2 Psycho-Linguistic Features

Table 2 shows the mean of LIWC linguistic features for the posts made in community and personal blogs, as well as the result of the statistical tests. For a

Table 2. The mean (±std) of LIWC features for posts made in community and personal blogs. Except for *wc* and *wps* are for the number of words for each post and number of words for each sentence in a post, respectively, the values for other features are the percentage of the LIWC categories in a post.

LIWC	Community	Personal	h_1	h_2	LIWC	Community	Personal	h_1	h_2
wc	225 ± 260	238 ± 456	1	1	anger	1.38 ± 1.9	1.09 ± 2.2	1	1
wps	15.2 ± 19.5	14.8 ± 24.3	0	1	sad	1.44 ± 1.8	0.60 ± 1.4	1	1
dic	90.88 ± 6.5	82.38 ± 14.3	1	1	cogmech	19.56 ± 5.3	15.57 ± 6.5	1	1
sixltr	12.98 ± 5.2	14.36 ± 7.2	1	1	insight	3.28 ± 2.2	2.12 ± 2	1	1
funct	60.86 ± 7.2	52.86 ± 13.1	1	1	cause	1.80 ± 1.6	1.39 ± 1.7	1	1
pronoun	20.86 ± 5.0	16.56 ± 6.8	1	1	discrep	2.28 ± 1.9	1.69 ± 2.0	1	1
ppron	14.29 ± 4.5	11.34 ± 5.6	1	1	tentat	3.43 ± 2.6	2.64 ± 2.5	1	1
i	11.13 ± 4.6	7.60 ± 5.0	1	1	certain	1.73 ± 1.6	1.36 ± 1.7	1	1
we	0.28 ± 0.8	0.49 ± 1.2	1	1	inhib	0.53 ± 1.0	0.57 ± 1.5	1	1
you	1.10 ± 2.2	1.65 ± 3.1	1	1	incl	4.32 ± 2.4	4.08 ± 2.9	1	1
shehe	1.20 ± 2.1	1.12 ± 2.1	1	1	excl	3.88 ± 2.4	2.84 ± 2.5	1	1
they	0.58 ± 1.1	0.49 ± 1.0	1	1	percept	2.34 ± 2.1	2.45 ± 2.8	1	1
ipron	6.57 ± 3.2	5.22 ± 3.4	1	1	see	0.54 ± 1.0	0.92 ± 1.7	1	1
article	4.10 ± 2.3	4.85 ± 3.2	1	1	hear	0.54 ± 1.0	0.62 ± 1.4	1	1
verb	18.21 ± 4.6	14.93 ± 6.0	1	1	feel	1.17 ± 1.4	0.74 ± 1.4	1	1
auxverb	10.94 ± 3.6	9.01 ± 4.5	1	1	bio	2.95 ± 2.7	2.86 ± 3.5	1	1
past	3.39 ± 2.7	3.25 ± 3.1	1	1	body	0.82 ± 1.3	0.97 ± 1.9	1	0
present	12.55 ± 4.6	9.56 ± 5.2	1	1	health	1.41 ± 1.8	0.76 ± 1.6	1	1
future	0.94 ± 1.1	0.99 ± 1.4	1	1	sexual	0.49 ± 1.1	0.70 ± 1.7	1	1
adverb	6.62 ± 3.0	5.43 ± 3.6	1	1	ingest	0.34 ± 1.0	0.57 ± 1.7	1	1
preps	11.63 ± 3.5	11.02 ± 4.4	1	1	relativ	12.83 ± 4.8	13.19 ± 6.2	1	1
conj	7.04 ± 2.8	5.94 ± 3.4	1	1	motion	1.64 ± 1.5	1.91 ± 2.2	1	1
negate	2.82 ± 2.1	1.91 ± 2.0	1	1	space	4.97 ± 2.7	5.16 ± 3.4	1	1
quant	2.92 ± 2.0	2.56 ± 2.2	1	1	time	5.84 ± 3.3	5.77 ± 4.2	0	1
number	0.65 ± 1.0	0.73 ± 1.4	1	1	work	1.17 ± 1.7	1.68 ± 2.7	1	1
swear	0.52 ± 1.2	0.57 ± 1.6	1	1	achieve	1.35 ± 1.4	1.32 ± 1.8	1	1
social	8.14 ± 5.2	8.04 ± 5.8	0	1	leisure	0.64 ± 1.2	1.45 ± 2.4	1	1
family	0.40 ± 0.9	0.33 ± 1.0	1	1	home	0.34 ± 0.7	0.49 ± 1.3	1	1
friend	0.36 ± 0.8	0.31 ± 1.0	1	1	money	0.29 ± 0.8	0.56 ± 1.5	1	1
humans	0.71 ± 1.1	0.72 ± 1.4	0	1	relig	0.22 ± 0.8	0.37 ± 1.3	1	1
affect	7.54 ± 3.8	6.57 ± 4.6	1	1	death	0.37 ± 1.0	0.23 ± 1.0	1	1
posemo	3.04 ± 2.5	3.82 ± 3.5	1	1	assent	0.25 ± 0.7	0.63 ± 1.5	1	1
negemo	4.39 ± 3.2	2.68 ± 3.2	1	1	nonfl	0.18 ± 0.5	0.23 ± 0.7	1	1
anx	0.67 ± 1.1	0.36 ± 0.9	1	1	filler	0.02 ± 0.2	0.02 ± 0.4	0	1

Table 3. The mean (±std) of topic distribution for posts made in community and personal blogs.

Topic	Community	Personal	h_1	h_2	Topic	Community	Personal	h_1	h_2
T1	3.24 ± 3.3	1.56 ± 1.1	1	1	T26	2.14 ± 1.7	1.92 ± 1.7	1	1
T2	1.42 ± 0.8	1.93 ± 2.7	1	1	T27	2.41 ± 2.0	2.07 ± 1.7	1	1
T3	1.55 ± 0.8	1.81 ± 2.1	1	1	T28	1.75 ± 1.3	2.19 ± 2.5	1	1
T4	2.31 ± 2.3	1.76 ± 1.5	1	1	T29	1.55 ± 1.7	1.69 ± 2.2	1	1
T5	1.30 ± 0.8	1.67 ± 3.2	1	1	T30	1.73 ± 1.2	2.08 ± 2.7	1	1
T6	2.25 ± 1.7	1.86 ± 1.5	1	1	T31	1.81 ± 1.9	2.40 ± 4.0	1	1
T7	2.11 ± 2.1	2.24 ± 2.6	1	1	T32	1.85 ± 1.6	1.83 ± 1.6	0	1
T8	2.44 ± 2.1	1.75 ± 1.3	1	1	T33	2.09 ± 1.7	2.06 ± 1.9	0	1
T9	2.08 ± 1.9	2.03 ± 1.9	1	0	T34	1.60 ± 1.2	2.18 ± 2.7	1	1
T10	2.08 ± 1.8	1.90 ± 1.6	1	1	T35	2.28 ± 3.3	2.05 ± 2.0	1	1
T11	1.83 ± 1.6	2.08 ± 2.4	1	1	T36	2.26 ± 1.8	1.90 ± 1.5	1	1
T12	2.54 ± 2.2	1.79 ± 1.4	1	1	T37	1.75 ± 1.3	1.87 ± 1.9	1	1
T13	2.49 ± 2.2	1.85 ± 1.5	1	1	T38	2.13 ± 1.6	2.00 ± 1.9	1	1
T14	1.55 ± 0.9	2.06 ± 2.2	1	1	T39	1.90 ± 1.6	2.18 ± 2.3	1	1
T15	1.81 ± 1.4	2.30 ± 4.3	1	1	T40	2.45 ± 2.1	1.76 ± 1.3	1	1
T16	1.68 ± 1.2	1.89 ± 1.9	1	1	T41	1.86 ± 1.7	2.16 ± 2.3	1	1
T17	1.68 ± 1.4	2.31 ± 2.9	1	1	T42	1.86 ± 1.5	2.16 ± 2.2	1	1
T18	2.21 ± 1.6	2.01 ± 1.5	1	1	T43	1.79 ± 1.5	2.15 ± 2.4	1	1
T19	1.63 ± 1.2	2.29 ± 2.8	1	1	T44	1.72 ± 1.2	1.87 ± 2.1	1	1
T20	2.49 ± 2.7	2.05 ± 1.9	1	1	T45	1.63 ± 1.0	2.25 ± 3.6	1	1
T21	2.49 ± 2.0	1.74 ± 1.3	1	1	T46	1.52 ± 0.9	2.27 ± 3.2	1	1
T22	1.58 ± 1.6	2.08 ± 3.2	1	1	T47	2.09 ± 1.5	1.89 ± 1.4	1	1
T23	2.30 ± 2.1	1.90 ± 1.5	1	1	T48	1.61 ± 1.1	1.81 ± 1.9	1	1
T24	2.19 ± 2.3	2.10 ± 2.4	1	1	T49	2.84 ± 3.5	1.73 ± 1.7	1	1
T25	1.74 ± 1.2	2.37 ± 2.4	1	1	T50	2.41 ± 3.0	2.20 ± 2.6	1	1

majority of LIWC features, t-test and/or rank sum test rejected the null hypothesis of no difference between Community and Personal cohorts. There were only six of 68 LIWC features (wps, social, humans, body, time, and filler) whose the null hypothesis were failed to be rejected by either t-test or rank-test or both, at the 0.05 level.

Figure 2 shows the relative difference in the use of LIWC features by Community and Personal cohorts. On the *affective* processes, it is observed that posts in the depression community blogs contained more *negative* (including *anxiety*, *sadness*, and *anger*) emotion and less *positive* words than those in the depression personal blogs. Community blogs also have more *affective* and *negate* words, compared with Personal blogs. In addition, they have more *pronoun* in the posts

Table 4. The prominent topics in favor of Community (number in red) and Personal (number in blue) blogs.

Topic	Word cloud	Topic	Word cloud
T1	depression anxiety therapist mental therapy disorder community diagnosed support issues	T46	movie movies watch watching character film season episode series scene star watched story big characters anime shows total princess evil
T49	doctor meds hospital medication pills anti appointment doctors weeks psychiatrist effects prozac medicine started zoloft psych anxiety pill	T19	hair black wear blue white red clothes wearing shirt dress pants looks green bought pink shoes
T21	cry stop crying anymore horrible upset wrong head tears cried sit hurts reason away wanting handle worse scream break feels	T45	favorite current color hair sex phone school movie crush food yep worst kissed gone date animal black opposite favourite eyes
T12	self thoughts feelings mind emotions emotional anger negative control sense esteem happiness physically emotionally fear normal lack feels	T17	sun water rain snow fire sky feet tree ground weather wind winter light trees warm sea place horse moon dark
T40	deal mood lately worse normal past problem worry anymore weeks control completely stress reason constantly fine gotten mind handle supposed	T25	birthday friday party saturday sunday monday thursday house wednesday tuesday wait coming sat hopefully dinner matt early decided spent plans

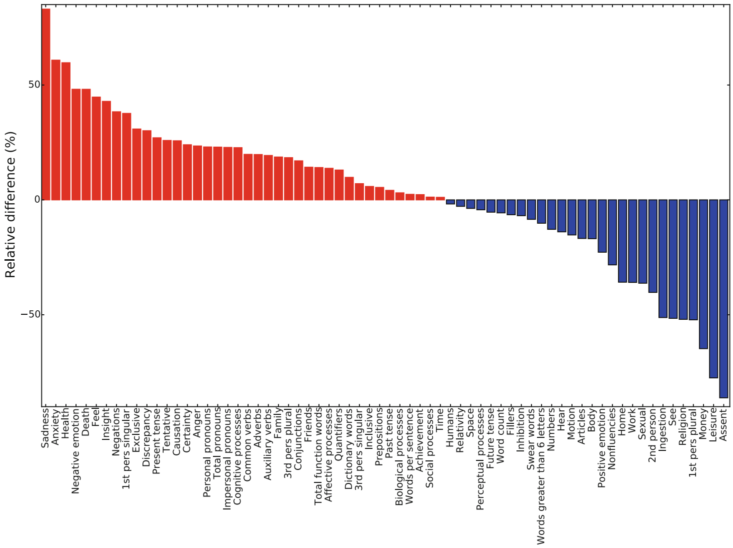


Fig. 2. The relative difference between Community and Personal blogs in the mean of LIWC features in the posts. The relative difference is in percent, as defined in Eq. 1. Features in red are in the preference of Community category, while those in blue are in the favor of Personal category. (Color figure online)

than do their Personal counterparts, except for *first personal plural* and *second personal*. This is partly in line with previous findings that depression people tend to use more pronouns [20].

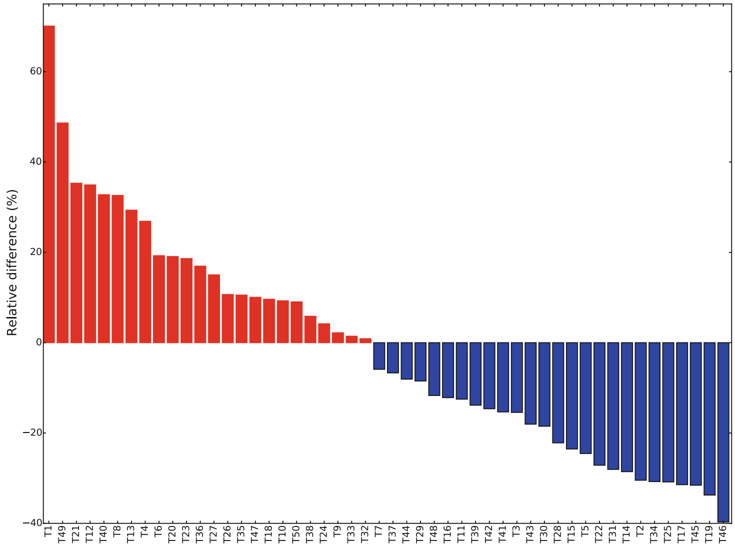


Fig. 3. The relative difference in the mean of topic proportion for the posts made in Community and Personal blogs. The relative difference is in percent, as defined in Eq. 1. Features in red are in the preference of Community category, while those in blue are in the favor of Personal category. (Color figure online)

On the other hand, on the *personal concerns*, except for *death*, people are likely to share all experience in all problems (*work*, *achievement*, *leisure*, *home*, *money*, and *religion*) more in personal than in community blogs.

On the *biological processes*, whilst *health* words were found more in community blogs, all the others (*body*, *sexual*, and *ingestion*) were used more in personal blogs. It shows a focus on the topic discussed within the community – health concerns.

On the *spoken* categories, except *fillers* which was not found to be significantly different between the two cohorts, *assent* and *non-fluencies* were used more in personal than in community blogs. It appears that people could freely post in their own pages with informal and unprepared text.

Similarly, *swear* words were used more in personal than in community pages. It may be because posting to the communities is often gone through a moderation process. So, a post with inappropriate words could be rejected to be posted to the community pages.

3.3 Topical Representation

Table 3 shows the mean of the proportion of topics for community and personal blogs, as well as the result of the statistical tests.[2] For 47 of 50 topics, t-test

[2] All 50 topics learned from the corpus by LDA are placed at http://bit.ly/1KEgjpM.

Table 5. Performance, in terms of the predictive accuracy (percentage of correct predictions), of different classifiers on different feature sets in the binary classifications of Community versus Personal posts.

Feature	SVM	NB	LR	Lasso
LIWC	59.6	68	77.5	**78.9**
Topic	77.7	69.7	**77.9**	77.8
Joint features	59.2	74.6	80.3	**80.5**

and/or rank sum test rejected the null hypothesis of no difference between community and personal cohorts.

Figure 3 shows the difference in the distribution of topics for posts made by community and personal cohorts. The topics with the highest difference were presented in Table 4. It is obvious to see in the table that topic 1 with depression themes ("depression", "anxiety", "therapist") was used much more in community than in personal blogs. Likewise, topic 49 with "doctor", "medicine", "hospital" was found in the preference of community posts, with the second highest relative difference, in comparison with personal posts. This partly confirms the finding in the language section that community posts were more focused on health concerns than were personal posts.

However, it is not straightforward to interpret the meaning of the other three topics in the favor of the community cohort. Nevertheless, "emotional and submissive" sense could be observed across the three topics. They also consisted of many affective words, confirming that community posts contained more *affective* words than did personal posts.

On the other hand, personal cohort favored more on concrete concepts, from "movie" or fashion ("hair") to weather ("sun", "rain") or "birthday", than did Community cohort. Most of these topics contain words in the LIWC "*leisure*" feature, which was found to be a favor language category for personal posts.

4 Classification

4.1 Performance

The Lasso model [9] is used for the classification. Using the coefficients derived from the Lasso method, we implemented a pair-wise classifier of *Community* versus *Personal* posts, using three feature sets: LIWC, topics, and a combination of them. The accuracy of this classifier in different feature sets are shown in Table 5, accompanied with that of SVM, Naive Bayes, and Logistic Regression. Lasso outperformed other classifiers when LIWC and a combination of LIWC and topics were used as the features, and was second to LR (77.8 % and 77.9 %), when topics were the features.

Classifying based on the derived topics outperformed the LIWC linguistic feature analysis in three of the four classifiers. Noticeably, a fusion of the features gained the best performance in all the classifications, except for the case when SVM was the classifier. This shows the potential of using multi-cues for mental health prediction.

4.2 Linguistic Features as the Predictors

Figure 4a shows the model using language style cues as features to predict community versus personal posts. Several features found significantly different between community vs. personal cohorts (see Sect. 3.2) were chosen into the prediction model. For example, for the *affective* category, *negative*, *sadness*, and *anxiety* were assigned large positive coefficients, whilst *positive* emotion had negative coefficient. The positive predictors also included *negation*, which was found in the preference of community blogs, confirming the finding in Sect. 3.2.

Of *personal concerns* chosen into the model, only *death* was assigned positive coefficient, whilst the rest (*work*, *leisure*, *home*, *money*, and *religion*) had negative coefficients. This is consistent with the findings presented in Sect. 3.2.

Similarly, on the *biological processes*, *health* was a positive predictor while *body*, *sexual*, and *ingestion* were negative ones.

As expected, *assent* and *swear* words were also negative predictors in the model. Noticeably, *assent* had the largest negative coefficient, becoming the most predictive feature of personal posts.

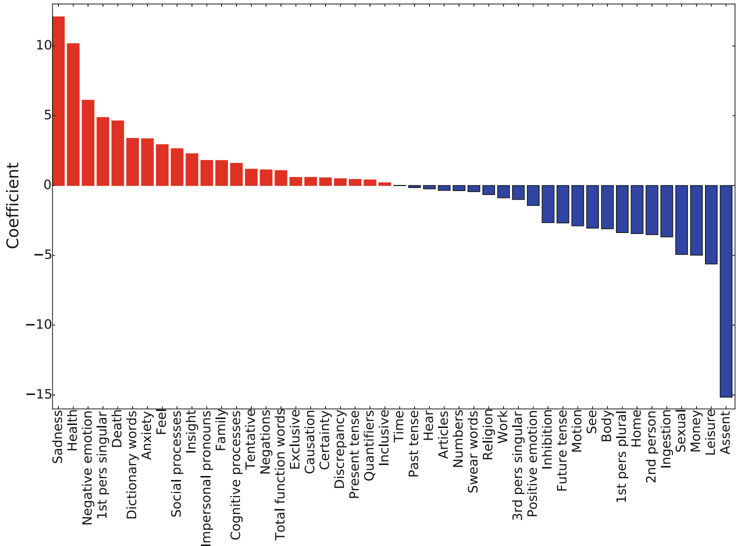
4.3 Topics as the Predictors

Figure 4b shows the classification model to predict community versus personal posts using topics as features. Table 6 shows the word cloud of topics whose the weights is largest in the model. It confirms that the top five topics in the preference of community blogs, which were "emotional and submissive" and focused on health concerns, as shown in Table 4, were among the top six positive predictors in the model.

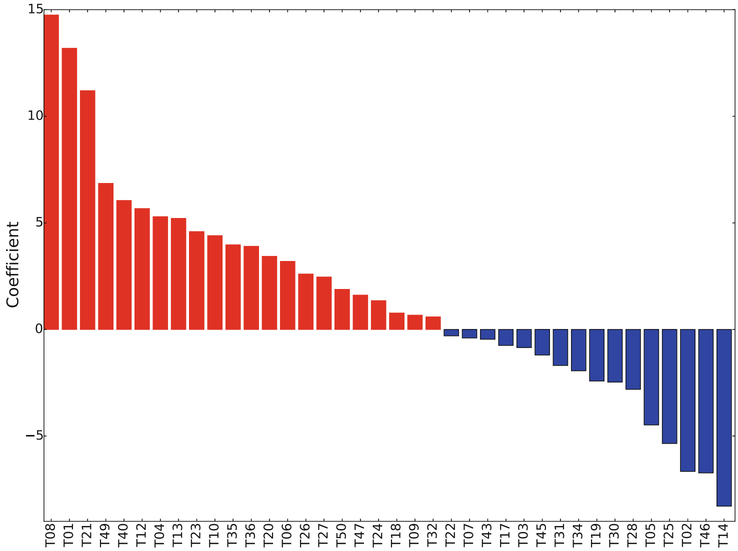
On the other hand, fashion and leisure topics, which were in the preference of personal blogs, were assigned negative coefficients. The topic on sex was also a negative predictor, which is in line with the finding on LIWC as the features that *sexual*, a *biological process*, was a negative predictor. Similarly, the topic of "eat", whose words belong to *ingestion*, another *biological process*, was also assigned a negative coefficient in the prediction model.

5 Limitation and Further Research

In the current work, the mental health status of the individuals in the online communities was not validated. Instead, labeling was "by affiliation" [5]. As such, it cannot be stated that all the individuals whose communications were analysed were experiencing depression. Future studies would benefit from attempting to validate this either by direct contact with the individuals or by analyzing the conversations for admission of diagnosis [4–6, 8]. If an admission is identified, other



(a) Lasso model coefficients for prediction of community (versus personal) posts using LIWC features as the predictors.



(b) Lasso model coefficients for prediction of community (versus personal) posts using topics as the predictors.

Fig. 4. Prediction models of community (versus personal) posts using topics and language styles as features. Features in red are positive predictors of community posts whilst the blues are the negatives. Individual coefficients that were not significant have been omitted, as have topics and LIWC features with no significant coefficients. (Color figure online)

Table 6. Topics with high weights in the prediction of community posts (versus personal ones). Positive weights are in red and negatives are in blue.

Topic	Word cloud	Topic	Word cloud
T8	end live kill suicide anymore death dead living point alive	T14	pictures watch picture amazing video yourwintr funny camera mike photos big pics youtube ryan watching sweet swandive random jon icon
T1	depression anxiety therapist mental therapy disorder community diagnosed support issues bipolar	T46	movie movies watch watching character film season episode series scene star watched story big characters anime shows total princess evil
T21	cry stop crying anymore horrible upset wrong head tears cried sit hurts reason away wanting handle worse scream break feels	T2	mark david finished james peter scott george michael robert bob jane king frank various richard read world ray alan mrs
T49	doctor meds hospital medication pills anti appointment doctors weeks psychiatrist effects prozac medicine started zoloft psych anxiety pill effexor medical	T25	birthday friday party saturday sunday monday thursday house wednesday tuesday wait coming sat hopefully dinner matt early decided spent plans
T40	deal mood lately worse normal past problem worry anymore weeks control completely stress reason constantly fine gotten mind handle supposed	T5	block writer oacute automatically tweets ntilde eacute shipped loudtwitter iacute asma aacute twitter
T12	self thoughts feelings mind emotions emotional anger negative control sense esteem happiness physically emotionally fear normal lack feels	T28	read write writing book reading story books written wrote paper art stories word letter words project poetry ideas interesting page
T4	cut pain cutting stop blood body hurt deep skin arm head scars arms	T30	sex woman women men gay sexual type female male test personality children straight marriage partner quizilla easily young body gender
T13	care hurt anymore trust matter deserve fault away lie real truth cares wrong feelings hurts somebody blame wanted inside cared	T19	hair black wear blue white red clothes wearing shirt dress pants looks green bought pink shoes
T23	boyfriend girl relationship months girlfriend girls past close broke date break dating meet hang boy friendship wanted single couple relationships	T34	music song band listening songs listen playing rock heard concert sound hear lyrics sing singing album radio guitar artist pop
T10	sleep bed wake sleeping woke asleep stay late fall slept hour energy half waking head early	T31	eat weight eating food fat dinner ate pounds lose sugar

statements made by that individual can be crawled for analysis. Alternatively, establishing a study and asking to recruit participants via these online communities would allow standardized diagnostic measurements of mental health to be administered. By having a more complete understanding of the mental health of the individuals in these communities, we are able to improve the profiling of their discussion and linguistic features so that a clearer match between online behavior and diagnosis can be made.

An example of this approach was conducted in [16], where depression scores of Twitter participants were collected and LIWC features were used as the predictors of the scores. Similarly, the depression ground-truth for Twitter users was collected individually in [21] (as opposed to "affiliation labeling" in this work), where tweeting behaviors were employed as the predictors.

6 Conclusion

We have investigated online depression blogs in community and personal settings. Affect information, language styles, and topics were extracted from the posts made in depression community blogs, as well as those made by members of the depression community in their own personal blogs. Statistical and machine learning techniques were used to discriminate the content in the two settings. A majority of features in all aspects were found to be significantly different between community and personal blogging by members of the community. Also, language styles and topics were found to have strong indicative powers in prediction of depression in personal and community settings. High performance in the classifications illustrates the effectiveness of the content-based cues as the discriminators. The result shows the potential of the new media in early screening and monitoring of online depression individuals and communities.

References

1. Blei, D.M., Ng, A.Y., Jordan, M.I.: Latent Dirichlet allocation. J. Mach. Learn. Res. **3**, 993–1022 (2003)
2. Bradley, M.M., Lang, P.J.: Affective norms for English words (ANEW): instruction manual and affective ratings. Technical report, University of Florida (1999)
3. Bridge, J.A., Goldstein, T.R., Brent, D.A.: Adolescent suicide and suicidal behavior. J. Child Psychol. Psychiatry **47**(3–4), 372–394 (2006)
4. Coppersmith, G., Dredze, M., Harman, C.: Quantifying mental health signals in Twitter. In: Proceedings of the Workshop on Computational Linguistics and Clinical Psychology: From Linguistic Signal to Clinical Reality, pp. 51–60 (2014)
5. Coppersmith, G., Dredze, M., Harman, C., Hollingshead, K.: From ADHD to SAD: analyzing the language of mental health on Twitter through self-reported diagnoses. In: Proceedings of the Workshop on Computational Linguistics and Clinical Psychology: From Linguistic Signal to Clinical Reality, pp. 1–10 (2015)
6. Coppersmith, G., Harman, C., Dredze, M.: Measuring post traumatic stress disorder in Twitter. In: International AAAI Conference on Weblogs and Social Media, pp. 579–582 (2014)
7. Cruwys, T., Haslam, S.A., Dingle, G.A., Haslam, C., Jetten, J.: Depression and social identity: an integrative review. Pers. Soc. Psychol. Rev. **18**(3), 215–238 (2014)
8. De Choudhury, M., Counts, S., Horvitz, E.: Major life changes and behavioral markers in social media: Case of childbirth. In: Proceedings of the 2013 Conference on Computer Supported Cooperative Work, pp. 1431–1442 (2013)
9. Friedman, J., Hastie, T., Tibshirani, R.: Regularization paths for generalized linear models via coordinate descent. J. Stat. Softw. **33**(1), 1–22 (2010)
10. Griffiths, T.L., Steyvers, M.: Finding scientific topics. Proc. Natl. Acad. Sci. **101**(90001), 5228–5235 (2004)

11. Mundt, J.C., Vogel, A.P., Feltner, D.E., Lenderking, W.R.: Vocal acoustic biomarkers of depression severity and treatment response. Biol. Psychiatry **72**(7), 580–587 (2012)
12. Nguyen, T., Phung, D., Adams, B., Venkatesh, S.: A sentiment-aware approach to community formation in social media. In: Proceedings of the International AAAI Conference on Weblogs and Social Media, pp. 527–530 (2012)
13. Nguyen, T.: Mood patterns and affective lexicon access in weblogs. In: Proceedings of the ACL Student Research Workshop, pp. 43–48 (2010)
14. Nguyen, T., Phung, D., Venkatesh, S.: Analysis of psycholinguistic processes and topics in online autism communities. In: Proceedings of the IEEE International Conference on Multimedia and Expo, pp. 1–6 (2013)
15. Nock, M.K., Green, J.G., Hwang, I., McLaughlin, K.A., Sampson, N.A., Zaslavsky, A.M., Kessler, R.C.: Prevalence, correlates, and treatment of lifetime suicidal behavior among adolescents: Results from the national comorbidity survey replication adolescent supplement. JAMA Psychiatry **70**(3), 300–310 (2013)
16. Park, M., Cha, C., Cha, M.: Depressive moods of users portrayed in Twitter. In: Proceedings of the ACM SIGKDD Workshop on Healthcare Informatics, pp. 1–8 (2012)
17. Pennebaker, J.W., Francis, M.E., Booth, R.J.: Linguistic Inquiry and Word Count (LIWC) [Computer software]. LIWC Inc. (2007)
18. Phung, D., Gupta, S., Nguyen, T., Venkatesh, S.: Connectivity, online social capital, and mood: a Bayesian nonparametric analysis. IEEE Trans. Multimedia **15**(6), 1316–1325 (2013)
19. Rodriguez, A.J., Holleran, S.E., Mehl, M.R.: Reading between the lines: the lay assessment of subclinical depression from written self-descriptions. J. Pers. **78**(2), 575–598 (2010)
20. Tausczik, Y.R., Pennebaker, J.W.: The psychological meaning of words: LIWC and computerized text analysis methods. J. Lang. Soc. Psychol. **29**(1), 24–54 (2010)
21. Tsugawa, S., Kikuchi, Y., Kishino, F., Nakajima, K., Itoh, Y., Ohsaki, H.: Recognizing depression from Twitter activity. In: Proceedings of the ACM Conference on Human Factors in Computing Systems, pp. 3187–3196 (2015)
22. Waxer, P.H.: Nonverbal cues for depth of depression: set versus no set. J. Consult. Clin. Psychol. **44**(3), 493 (1976)

http://www.springer.com/978-3-319-49585-9

Advanced Data Mining and Applications
12th International Conference, ADMA 2016, Gold Coast, QLD, Australia, December 12-15, 2016, Proceedings
Li, J.; Li, X.; Wang, S.; Li, J.; Sheng, Q.Z. (Eds.)
2016, XVI, 817 p. 280 illus., Softcover
ISBN: 978-3-319-49585-9