

Evaluating Individual Subjective Well-being via Social Media

Yazhou Wang

School of Electronic and Computer Engineering Shenzhen Graduate School, Peking University

Shenzhen 518055, Guangdong, China

email: pkuwangyz@gmail.com

Keywords: subjective well-being, social media, semantic lexicon, feature extraction

Abstract. Subjective well-being (SWB) is now attracting more and more attention from policy makers. Traditional way of measuring SWB by self-reporting based on questionnaire has the limitation of small sample size and low survey frequency. With the rapid spread and use of social media, the online posts of users provide a new way to study SWB. But the existing work haven't gone far from word counting and ignore the components of SWB. In this paper we propose a new framework to evaluate individual SWB from social media. We leverage the language cues and introduce useful features to evaluate SWB. Moreover, we introduce a way to expand the SWB feature. To check the effectiveness of the framework, we crowdsource user tweets and their subjective well-being. Experiments on the dataset show that this way of evaluating SWB achieves a high accuracy in evaluating individual SWB.

Introduction

Subjective well-being (SWB) also known as happiness indicates how people evaluate their quality of life and the relative frequency of positive and negative affect. Quality of life, rather than material welfare is playing as a more important role at directing public policy. But traditional way of measuring SWB—survey research—is expensive and time-consuming, which makes it couldn't be conduct frequently, and survey usually covers a small percentage of the whole people.

With the rapid spread of social networking sites, such as Twitter, Facebook, Weibo, people usually share their emotional experiences and daily life through their updates [7]. This provides real-time information about users' daily life and offers a chance for studying SWB from online data. Facebook build Facebook Gross National Happiness (FGNH) in 2010 based on sentiment analysis to measure SWB. However, methods like FGNH use sentiment analysis just count the number of positive words and negative word, which may take many unrelated topics into account and they don't analysis the components of subjective well-being, thus making the result unreliable. For example, on the fourth of July, one talks about his job and another one talks about his pray for the nation, the first one may post: *"I'm so bored with this job. It's mind numbing."*, and the latter one: *"Happy Fourth of July everyone and one nation under God :-)"*. In this example, this two posts both have sentiment polarity, but they should not have the same contribution to the total SWB of one person. Apparently the first one should makes more contribution to one's subjective well-being because his job is part of his life and he is not satisfied with that. While the latter one express his joyful or his pray for America, but not express his life satisfaction. Most of the existing methods don't distinguish the different effect of posts.

In this paper, we propose a new framework for evaluating users' SWB through online posts. We utilize the research results in positive psychology on subjective well-being. And the framework takes different components of subjective well-being into account separately. To address the difference, we introduce new features related to different components of SWB. Moreover, we introduce one way to enrich the features which makes it could catch more signals related to SWB. At the end, we give a framework to evaluate user SWB from user posts. We demonstrate that our approach is more effective than existing methods in evaluating SWB on twitter dataset.

Our contributions are as follows: 1) We propose a new framework to measure and evaluate users' SWB from online posts at user-level. 2) We utilize the components of subjective well-being

from study results in positive psychology and map them to text as SWB features. We propose one way to expand the features by using semantic lexicon. 3) We evaluate our approach on tweets collected from crowdsourcing. Experimental results demonstrate that our framework is more effective than the existing methods of measuring SWB.

Related Work

Subjective well-being is an important field of positive psychology, and has been studied year. [8] reviewed the factors in producing SWB, and found the components of SWB. Now it is widely accepted that SWB is multidimensional: encompassing all aspects of human life [11]. When SWB is used to monitor progress and to inform policy, there are many ways in which SWB can be measured. [17] thinks that subjective well-being can be promoted from five dimensions: positive emotions, engagement, relationship, meaning and accomplishment. More generally, [12] gives some guidelines on measuring SWB, contains question construction, what to measure et al. Although these researches on measuring are based on questionnaire, they provide the components of SWB and which kind of information we should use to measure SWB. The features we used to evaluate SWB from texts are based on these works.

Measuring SWB from online posts is related to sentiment analysis [7] [4]. Many studies has achieved a high accuracy by design effective features [13-14] in sentiment classification. But we cannot apply them to measuring SWB directly because it ignores the components of SWB. [15] use Affective Norms for English Words (ANEW) to evaluate happiness. Similarly, [16] use positive and negative words from OpinionFinder sentiment lexicon to measure happiness. And the SWB of one user is related to the number of positive terms and negative terms. These works all ignore the components of subjective well-being. We differ from these works in that we distinguish the most relevant components of SWB and other topics, so we get effective features to evaluate happiness.

The Proposed Framework

The framework of measuring and evaluating subjective well-being via social media shows in (Figure 1). In particular, we define the following three main components:

- 1) SWB feature definition: We define the features used in our framework.
- 2) Feature expansion: As the size of defined features is small and some words in features may have synonyms, we propose a novel approach to expand the features by using semantic lexicon. At the same time, the approach is a weight-based way, the features in expanded feature set have different weight based on the similarity to words in defined feature set.
- 3) Combined SWB classification: We introduce sentiment analysis to help to evaluated SWB. Combining the SWB features extracted from the posts and sentiment analysis result, we can get a vector describing one user. Then we train a model for classifying user into high or low happiness.

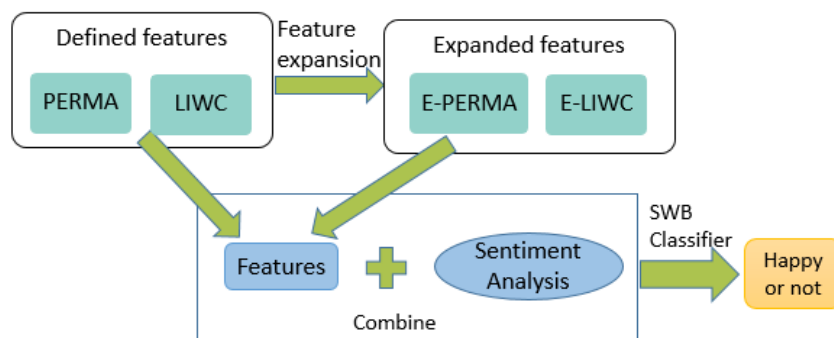


Fig. 1. Framework of evaluating subjective well-being

● SWB Features

Seligman [17] introduced the PERMA model with five core elements of well-being: 1) Positive emotions (P), positive emotion encompasses feelings like happiness, pleasure. This element is the most obvious connection to happiness. 2) Engagement (E), engagement means deep psychological

connection to some activity. 3) Relationships (R), relationship includes integration with society or one's social network. 4) Meaning (M), meaning refers to having a direction in life. 5) Accomplishment (A), making progress towards or achieving one's goals often leads to a sense of satisfaction. We also take the Seligman model as an important feature in our framework.

Besides the PERMA features, some works [9-10] show that SWB is also related to other variables such as health, income, religion and family life. To catch these variables, we take LIWC (Linguistic Inquiry and Word Count) as the second feature. LIWC counts words in psychologically meaningful categories. The 2001 LIWC contains 74 categories, 2,300 words and word stems. Each word or word stem may belong to more than one category.

So, we use PERMA and LIWC as the defined SWB features. The PERMA lexicon is from [2].

● Feature Expansion

A common way to analysis if one sentence is related to one category of PERMA or LIWC is to check whether the tokens of the sentence belong to the category. But the two lexicons only contain 3824 words. Besides, considering the words not in any of the categories but have a similar meaning with the words in one category, simply discarding the word may cause some important information missing. So we propose one way to solve this problem by knowledge graph.

For every word in the whole dictionary and every category of the dictionary, we assign a weight to the word for the category, which indicates the probability that word belongs to the category. Apparently, the word in one category should have a weight of 1 for that category. Let $C = \{C_1, C_2, \dots, C_m\}$ donates the categories in PERMA and LIWC, and w_c^i donates the words in category C_i . For word w , the semantic similarity between w and words in C_i is based on the shortest path that

connects the words. Here, we define word similarity as $sim(w, w_c^{i,j}) = \frac{1}{dis(w, w_c^{i,j}) + 1}$, where $dis(w, w')$ is the shortest distance of word w and w' in semantic lexicon. So, when we can get the probability of word w belonging to category C_i from Equation (1), where k is the vocabulary size of category C_i .

$$P(C_i | w) = \begin{cases} 1 & w \in w_c^i \\ \frac{1}{k} \sum_{w_c^{i,j} \in w_c^i} sim(w, w_c^{i,j}) & w \notin w_c^i \end{cases} \quad (1)$$

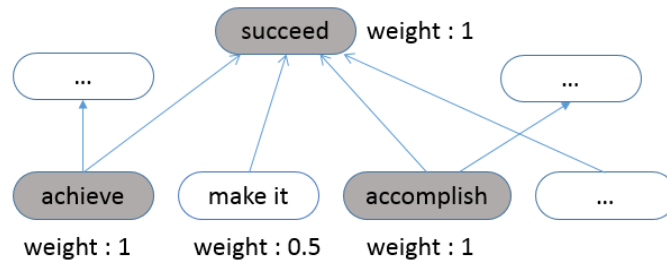


Fig. 2. An example of expanding lexicon

Figure 2 shows an example of expansion of features through WordNet. Gray oval frames represent the words in PERMA lexicon, while white oval frames represent words not in PERMA lexicon. In WordNet, “achieve” and “make it” and “accomplish” are all troponyms of “succeed”. Although word “make it” is not in PERMA lexicon, it can also be used a part of features through semantic lexicon. Here, words in gray oval frames are all part of positive accomplishment, their weights are all 1. Word “make it” should be part of positive accomplishment, and its weight is 0.5.

After the expansion, we can represent each word w by a vector $v = \langle w_1, w_2, \dots, w_m \rangle$ and w_i represent the probability of w belongs to category C_i . However, a vector with too many values, which indicates the word is related to many categories, is inconsistent with common sense. Therefore, we only retain the ‘1’s and top two non-‘1’ elements in the vector, and set other elements to zero.

● The Combined Framework

We combine SWB feature and sentiment analysis in our framework to achieve better result. The

task of sentiment analysis is to classify a tweet into three classes: Positive, Negative, and Neutral. We build a sentiment classifier by using a wide range of features. Before sentiment analysis, a tweet is tokenized and POS-tagged, the POS-tag information is used in both sentiment analysis and the SWB feature analysis step.

Sentiment analysis features that have been proven effective in previous works are used to train a SVM classifier. Specifically, we use the following features set: 1) n-grams, for $n \in \{1, 2, 3\}$. 2) Parts of speech, the number of each POS-tag. 3) Emoticons, the emoticons in a tweet, the polarity of emoticons serve as emoticons features. 4) Polarity dictionaries, we use three different polarity dictionary. They are NRC Emotion Lexicon created in [5], the MPQA Lexicon in [6], and the Bing Liu Lexicon. 5) Punctuation symbols, the number of single and non-single punctuation marks (e.g., ! or ? or !!!) are considered separately. 6) The number of hashtags.

After sentiment analysis, the posts of one user was classified to three classes. We run SWB feature analysis on the three classes separately. In each class, we regard the posts as a whole document. We follow the similar way used in [1] when analyzing a whole sentence (or user posts) using Equation (2).

$$usage(category, user) = \sum_{term} p(category | term) * \frac{freq(term, user)}{freq(*, user)} \quad (2)$$

Where $usage(category, user)$ is the total weight of one user in category $category$, $p(category | term)$ is the probability of $term$ belongs to category $category$, $freq(term, user)$ is frequency of the word for the given user, and $freq(*, user)$ is the total word count for the user.

After analysis on different class separately, we can get three vectors representing the posts of one user. We concatenate the vectors as the input for the SWB classifier. And the output of the classifier is whether the user is satisfied with life or not.

Experimental Setup

We employ crowdsourcing to collect Twitter user as ground truth data on evaluating SWB. We set up an Amazon’s Mechanical Turk (www.mturk.com) task which has previously used in [2] for sentence annotation, wherein workers were asked to take a survey about their satisfaction with life. The workers are also required to share their Twitter usernames with an agreement that their data could be downloaded and analyzed using a computer program, but will only be used in research purpose. And each worker was restricted to take the task only once.

As we want to evaluate users’ SWB from the posts, we only consider whether they are happy or not. So we designed the survey simply asking “All things considered, whether are you satisfied with your life as a whole in recent months?” following the question asked in World Values Survey (WVS) Wave 6 . And the workers just answer “yes” or “no”.

A total of 795 workers completed our task. We crawled the posts of these users by Twitter API. After filtering the users who only have a small number of tweets (<50), we got 515 users and their tweets. Among the 515 users, 297 are satisfied with their life in recent months. We only retain the tweets posted in the latest year, and get 34873 tweets.

Twitter sentiment analysis is part of measuring SWB, so the result of sentiment analysis should be reliable. We use public training tweets in SemEval-2015 to train our sentiment classifier.

We employ CMU ARK twitter POS tagger for POS tagging. And we use WordNet as the semantic lexicon to expand features. We use logistic regression with L2 regularization to classify users. We use 10-fold cross validation for evaluation.

Experiments

We evaluate the performance of our model with two experiments. In the first experiment, we compare our method with the baselines, to analyze the performance of our method. In the second experiment, we analyze the performance of different configurations in our method.

● Comparison with Compared Methods

We perform 10-fold cross-validation on the dataset to tune all these model and obtain their optimal parameters. The different methods are as follows: 1) SENA: This method used in [4], which uses two kinds of classifier to compute happiness. 2) PSWB: The method proposed in [2] uses n-grams and LDA topics from Facebook data and lexica as features. 3) PSWL: The method was used in [3] on Facebook datasets, they use many features like Big Five features, age of a user. The result of comparison with different methods are shown in Table 1.

Table 1. Performance on evaluating SWB of different methods

Method	Precision	Recall	Accuracy
<i>SENA</i>	0.632	0.593	0.627
<i>PSWB</i>	0.658	0.629	0.663
<i>PSWL</i>	0.594	0.614	0.581
<i>Our Method</i>	0.692	0.770	0.702

Compared with other methods, ours outperforms the three baselines in terms of precision, recall and mean accuracy. Among the methods, PSWL has the worst performance on our dataset. SENA and PSWB both use word count and sentiment analysis method and both have good performance. This shows that it is appropriate to use sentiment and word count method for evaluating SWB from social media. SENA only use one kind of lexicon, i.e. LIWC lexicon, which makes it can't catch much information about the components of subjective well-being. The PSWB model has the best performance among the baselines. This may because this method also use many information about the components of subjective well-being. The shortcoming of this model is that though it can catch the components of SWB, but it cannot effectively detect the sentiment polarity of the user towards the component. Our method effectively model the components of SWB and combine sentiment information with it, so it yields the best performance on the dataset.

● Comparison of Different Configurations

We present some experiment results to discuss the effects of different features we used in our model for evaluating subjective well-being.

The different settings we used are as follows: (1) S-PERMA: we use sentiment analysis, and only use PERMA lexicon, which means we remove the LIWC lexicon. (2) S-LIWC: we use sentiment only with LIWC lexicon and remove PERMA lexicon. (3) SL-no-expand: we use sentiment analysis and lexicon analysis with PERMA and LIWC lexicon, but we don't expand the lexicon with semantic lexicon. (4) Full: our full model.

Table 2. Performance of different settings

Method	Precision	Recall	Accuracy
<i>S-PERMA</i>	0.645	0.541	0.621
<i>S-LIWC</i>	0.627	0.532	0.603
<i>SL-no-expand</i>	0.663	0.684	0.677
<i>Full</i>	0.692	0.770	0.702

As shown in Table 2, method that use only one lexicon perform worse due to the reason that it can't capture all useful information that may be components of SWB thus may contribute to measure SWB. From the results, we can also find that the PERMA lexicon is more effective than LIWC. This may because that the PERMA is directly associated with subjective well-being, while the LIWC lexicon contains too much other information. And using both the lexicon has a significant improvement to the mean accuracy. And when using WordNet to expand the lexicon, we can take more words into consideration and get more features, which can also help to improve the result.

Conclusion

Evaluating subjective well-being is important for studying the whole society and making public policy. Measuring SWB from social media is a real challenge, from both computational and linguistic perspective. In this paper, we present a novel method to evaluate SWB from social media posts by combining SWB features and sentiment analysis. We propose a SWB feature set which

takes the research results from psychology into account, and propose one way to expand the feature set. We compare our model against existing approaches on twitter data set. The experiments give promising results.

References

- [1] Sap, Maarten, et al. "Developing age and gender predictive lexica over social media." *Proceedings of the 2014 Conference on Empirical Methods in Natural Language Processing, EMNLP*. 2014.
- [2] Schwartz, H. ANDREW, et al. "Predicting individual well-being through the language of social media." *PSB*, 2016.
- [3] Collins, Susan, et al. "Are You Satisfied with Life?: Predicting Satisfaction with Life from Facebook." *Social Computing, Behavioral-Cultural Modeling, and Prediction*. Springer International Publishing, 2015. 24-33.
- [4] Quercia, Daniele, et al. "Tracking gross community happiness from tweets." *Proceedings of the ACM 2012 conference on Computer Supported Cooperative Work*. ACM, 2012.
- [5] Mohammad, Saif M., and Peter D. Turney. "Crowdsourcing a word–emotion association lexicon." *Computational Intelligence* 29.3 (2013): 436-465.
- [6] Wilson, Theresa, Janyce Wiebe, and Paul Hoffmann. "Recognizing contextual polarity in phrase-level sentiment analysis." *Proceedings of the conference on human language technology and empirical methods in natural language processing*. Association for Computational Linguistics, 2005.
- [7] Liu, Pan, et al. "Do Facebook Status Updates Reflect Subjective Well-Being?." *Cyberpsychology, Behavior, and Social Networking* 18.7 (2015): 373-379.
- [8] Diener, Ed. *Subjective well-being: The science of happiness and a proposal for a national index*. Vol. 55. No. 1. American Psychological Association, 2000.
- [9] Diener, Ed, and Eunkook M. Suh. *Culture and subjective well-being*. MIT press, 2000.
- [10] Graham, John R., and Micheal L. Shier. "The social work profession and subjective well-being: The impact of a profession on overall subjective well-being." *British Journal of Social Work* 40.5 (2010): 1553-1572.
- [11] Decancq, Koen, and Maria Ana Lugo. "Weights in multidimensional indices of wellbeing: An overview." *Econometric Reviews* 32.1 (2013): 7-34.
- [12] OECD M. *OECD Guidelines on Measuring Subjective Well-being*. 2013.
- [13] Kiritchenko, Svetlana, Xiaodan Zhu, and Saif M. Mohammad. "Sentiment analysis of short informal texts." *Journal of Artificial Intelligence Research* (2014): 723-762.
- [14] Kim, Yoon. "Convolutional neural networks for sentence classification." *arXiv preprint arXiv:1408.5882* (2014).
- [15] Dodds, Peter Sheridan, and Christopher M. Danforth. "Measuring the happiness of large-scale written expression: Songs, blogs, and presidents." *Journal of Happiness Studies* 11.4 (2010): 441-456.
- [16] Bollen, Johan, et al. "Happiness is assortative in online social net-works." *Artificial life* 17.3 (2011): 237-251.
- [17] Seligman, Martin. *Flourish: A new understanding of happiness, well-being-and how to achieve them*. Nicholas Brealey Pub., 2011.