HARK: Harshness-Aware Sentiment Analysis Framework for Product Review (Student Abstract)

Ting Zhou, Xun Wang, Yili Fang[*]

School of Computer and Information Engineering, Zhejiang Gongshang University,Hangzhou,China,310018

18020100012@pop.zjgsu.edu.cn,{xwang, fangyl}@zjgsu.edu.cn

Abstract

Sentiment analysis has been a helpful mechanism that targets to understand the market feedback on certain commodities by utilizing the user comments. In the process of providing comments, each user comment is generated based on his/her preference which is referred to as harshness. Existing methods mainly apply majority voting or its variants to directly infer the evaluation of products. Nevertheless, due to the ignorance of the harshness of users, these methods will lead to low-quality inference outcome of sentiment analysis, which is far from the result of the expert analysis report. To this end, we propose HARK, a harshness-aware product analysis framework. First, we employ a Bayesian-based model for sentiment analysis. Moreover, in order to infer the reliable sentiment concerning each product from all the comments, we present a probabilistic graphical model in which the harshness is incorporated. Extensive experimental evaluations have shown that the result of our method is more consistent with the expert evaluation than that of the state-of-the-art methods. And our method also outperforms the method which infers the final sentiment with the ground truth of comments but without involving the harshness of users.

Introduction

Sentiment analysis targets at the evaluation of the product quality with user comments or the analysis of the social event impact with public opinions. Recently, with the development of online commentary communities such as Douban, sentiment analysis has attracted widespread attention from researchers and industry. Nevertheless, the accuracy of evaluating products based on sentiment analysis is still far from the product analysis report of experts. Therefore, enhancing the accuracy of sentiment analysis is still an essential issue in product evaluation.

The product evaluation method aims at inferring the user's viewpoint of products based on the sentiment analysis of the text. However, the harshness of users often makes them provide biased comments. For a tolerant user, although the user is extremely dissatisfied with the product, harshness still makes her yield a neutral comment. This kind of harshness is very common in product evaluation of web sites. For example, as for the sentiment analysis dataset *IMDB* (Tang, Bing,
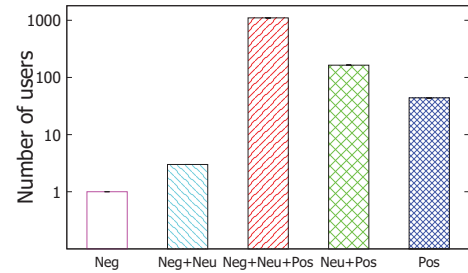
Figure 1: User distribution of dataset *IMDB*

and Liu 2015), it contains products and the corresponding comments from users. We analyzed the distribution of users on it. As shown in Fig 1, the analysis result shows that for any product, 44 users only give positive comments, 164 users only provide positive or neutral comments, 3 users only provide neutral or negative comments. We also observe that even there is a person who gives negative comments for all the products he/she has bought. From this point of view, the harshness plays a vital role in the process of evaluating the products. Taking into account that some comments are not objective due to users' preference, the work (Gong and Wang 2018) removes the comments of harsh users and tolerant users and conducts the viewpoint inference with the cut-off data. Obviously, it is difficult to obtain an accurate product evaluation due to the lack of useful data. Consequently, a problem of how to incorporate the harshness into the viewpoint inference with the aim of improving the accuracy of sentiment analysis rises. In this paper, we propose HARK, a harshness-aware product analysis framework. We first use a Bayesian-based model to analyze the comments. Then, we present a probabilistic graphic model to incorporate the harshness of users with text analysis into the inference of user comments and use the expectation-maximization (EM) algorithm to infer the evaluation of the product. Finally, we perform an extensive set of experiments to evaluate the effectiveness of our method.

Proposed Method

In our method, we analyze the user's evaluations about product attributes based on text and present a probabilistic graph-

ical model in which the harshness is incorporated to infer evaluations of product attributes.

In order to conduct aspect-based sentiment analysis, we first select aspect words. Next, the mean shift algorithm is used to cluster them in order to get product attributes such as the plot of the movie. Then, we divide comments according to product attributes. Followed by, combining document embedding pre-trained by TF-IDF model and SVM-based model, we can derive users' evaluations of product attributes. Let j^t be the attribute t of product j, $s_{ij}^t = \{s_{ij}^{tz}|z \in \{Negative, Neutral, Positive\}\}$ denotes the probability that user i give different evaluations to product attribute j^t. We construct a three-by-three confusion matrix $\mathbf{R}$ where $\mathbf{R}(x_s, x_y)$ represents the probability of the user comment is evaluated as x_s when the product attribute is evaluated as x_y. Specifically, $\mathbf{R}(x_s, x_y)$ is defined as:

$$\mathbf{R}(x_s, x_y) = \begin{cases} s_{ij}^{tx_s} & x_s = x_y \\ \dfrac{1 - s_{ij}^{tx_s}}{2} & x_s \neq x_y \end{cases}.$$

Since user comments are generated based on users' harshness, we model the harshness of each user i using the parameter $\mu_i \in (-\infty, \infty)$. Meanwhile, each product has a different evaluation difficulty. The difficulty of the evaluation of product attribute j^t is modeled by the parameter $1/\tau_j^t \in [0, \infty)$ where τ_j^t is constrained to be positive. We represent the probability of s_{ij}^t given y_j^t, μ_i, τ_j^t, and $\mathbf{R}$ as follows.

$$p(s_{ij}^t|y_j^t, \mu_i, \tau_j^t, \mathbf{R}) = \begin{cases} \dfrac{1}{1 + e^{-\mu_i \tau_j^t}} \mathbf{R}(s_{ij}^t, y_j^t) & s_{ij}^t = y_j^t \\ \dfrac{1 - \frac{1}{1 + e^{-\mu_i \tau_j^t}} \mathbf{R}(s_{ij}^t, y_j^t)}{T - 1} & s_{ij}^t \neq y_j^t \end{cases},$$

where y_j^t is the true evaluation of product attribute j^t and T is the total number of evaluation types that the user may give. Our goal is to find the posterior distribution of y_j^t and to select the evaluation y_j^t with the maximum a posterior estimation as the final evaluation of product attribute j^t. We use the EM algorithm to obtain the parameters of μ_i and τ_j^t.

Experimental Evaluation

We conduct experiments on the movie review dataset with user and product information obtained from *IMDB* (Tang, Bing, and Liu 2015). And we collect the expert evaluations as the ground truth. Besides our method, we also implemented four algorithms as the baseline methods, i.e., Statistical method with dataset label (ComTL) in which the ground truth of comments is available, attention-based LSTM with aspect embedding (ATAE) (Xue and Li 2018), CNN model based on aspect embedding (GCAE) (Xue and Li 2018), SVM with comment segmentation (SenSVM).

As shown in Fig 2, the methods (i.e., *SenSVM* and ours) which use comment segmentation outperform the methods (i.e., *ATAE* and *GCAE*) which use aspect embedding. This means the comment segmentation is more able to detect the product attributes expressed in each sentence. Then, our method has a 4% improvement of accuracy over *SenSVM*.
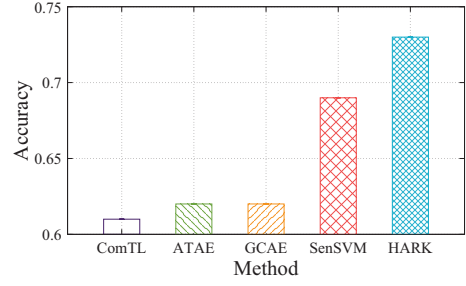

Figure 2: Accuracy of the methods

This is because incorporating users' harshness in the viewpoint inference can improve the quality of final outcome. Finally, we verify the consistency between the results of the five methods and expert evaluation by the Kappa coefficient and the Kendall coefficient. The results are presented in Table 1. As for these two coefficients we can observe that the result of our method is closer to expert evaluation than that of the other four methods.

Table 1: Kappa coefficient & Kendall coefficient

Methods	Kappa	Kendall
ComTL	25.00%	44.76%
ATAE	23.12%	31.12%
GCAE	27.83%	43.16%
SenSVM	45.17%	42.24%
HARK	**52.76%**	**50.66%**

Conclusions

In this paper, basing on the analysis of users' harshness, we propose HARK, a harshness-aware product analysis framework. To be specific, we first employ a Bayesian-based sentiment analysis method. Subsequently, considering users' harshness, we give a viewpoint inference method of the product based on a probabilistic graphical model. The experimental results show that the result of our method is more consistent with the result from the expert report than that of state-of-the-art methods. And, our method also outperforms *ComTL* in which the ground truth of comments is available.

Acknowledgements

References

Gong, L., and Wang, H. 2018. When sentiment analysis meets social network: A holistic user behavior modeling in opinionated data. In *SIGKDD*, 1455–1464.

Tang, D.; Bing, Q.; and Liu, T. 2015. Learning semantic representations of users and products for document level sentiment classification. In *IJCNLP*, 1014–1023.

Xue, W., and Li, T. 2018. Aspect based sentiment analysis with gated convolutional networks. In *ACL*, 2514–2523.