Review Recommendation with Graphical Model and EM Algorithm

Richong Zhang
School of Information Technology and Engineering
University of Ottawa
800 King Edward Avenue
Ottawa, ON, Canada
rzhan025@site.uottawa.ca

Thomas Tran
School of Information Technology and Engineering
University of Ottawa
800 King Edward Avenue
Ottawa, ON, Canada
ttran@site.uottawa.ca

ABSTRACT

Automatically assessing the quality and helpfulness of consumer reviews is more and more desirable with the evolutionary development of online review systems. Existing helpfulness assessment methodologies make use of the positive vote fraction as a benchmark and heuristically find a "best guess" to estimate the helpfulness of review documents. This benchmarking methodology ignores the voter population size and treats the the same positive vote fraction as the same helpfulness value. We propose a review recommendation approach that make use of the probability density of the review helpfulness as the benchmark and exploit graphical model and Expectation Maximization (EM) algorithm for the inference of review helpfulness. The experimental results demonstrate that the proposed approach is superior to existing approaches.

Categories and Subject Descriptors

H.3.3 [**Information Search and Retrieval**]: Information Filtering; H.4.m [**Information Systems**]: Miscellaneous

General Terms

Algorithms, Measurement, Theory

Keywords

Helpfulness, Recommendation, Online Review

1. INTRODUCTION

Due to the growth of internet business, more and more web sites are providing services by requesting users to leave reviews after finishing a transaction. "Online product reviews provided by consumers who previously purchased products have become a major information source for consumers and marketers regarding product quality" [2]. There are, however, a number of challenges for consumers to make their best use of online reviews. Reviewers write reviews and rate products by using a number of stars. Such a star scale rating, together with the fact that reviews are usually unstructured and often mix the reviewers' feelings and opinions, make it difficult for consumers to get the real semantics of reviews.

WWW 2010, April 26–30, 2010, Raleigh, North Carolina, USA.
ACM 978-1-60558-799-8/10/04.

Search engines are good tools to assist in looking for information, but the result set of a query by search engine is huge. Also, an online community like Epinion.com usually receives more than 1000 reviews submitted by different users for a specific product. These justify why it is essentially important to develop systems that can recommend helpful reviews to consumers effectively. We notice that most of the review aggregation web sites provide helpfulness voting function for consumers to rate reviews. However, this progress takes time far before a really helpful review to be discovered, and the latest published review will always be the least voted one. Existing review helpfulness assessing approaches commonly take the positive vote fraction as the benchmark and apply some general machine learning tools to infer the helpfulness of the online review documents. We argue that using positive vote fraction as the benchmark is a significant limitation of these approaches because they treat the the same positive vote fraction as the same helpfulness no matter how many voters voted.

In this paper, we propose a novel probabilistic formulation to model the review helpfulness distribution and to recommend review documents with higher probability of being a helpful review to consumers. Under this probabilistic formulation, the helpfulness of a review document is given a precise mathematical meaning. We also develop a probabilistic graphical model and an Expectation Maximization algorithm for helpfulness inference. The experimental results obtained by this model show that our approach gives stat-of-the-art effectiveness.

2. PROPOSED PROBABILISTIC MODEL OF REVIEW HELPFULNESS

Given a collection of voters' opinion on a review document, we can use Bayes rule to infer the distribution of the helpfulness value. Formally, we use i to index the review documents set D, I to denote the index the set of all review documents. For each $i \in I$, let $\Gamma(i)$ index the set of all voters on the i^{th} review document and let $V_{\Gamma(i)}$ denote the collection of all votes on the i^{th} document. We also define $V_{\Gamma(i)} = \{V^j_{\Gamma(i)} : j \in \Gamma(i)\}$. Denote by $\Gamma := \cup_{i \in I}\Gamma(i)$ the set of all voters that have voted on at least one document in D_I. The helpfulness value α_i of a review document D_i then can be defined as the probability of a random chosen voter j, where $j \in \Gamma(i)$, voting positively on the review document D_i. Equivalently, the helpfulness of review D_i can be characterized by the conditional probability distribution

$P(V^j_{\Gamma(i)}|D_i)$, where D_i is a random review from D_I and $V^j_{\Gamma i}$ is a random voter's $\{0,1\}$-valued opinion from $V_{\Gamma(i)}$.

Figure 1 gives the graphical representation of the review helpfulness model. F_i denotes the feature of a review document D_i. α_i denotes the helpfulness of D_i. $V^j_{\Gamma(i)}$ denotes a vote submitted by the readers to D_i. A review document is a collection of N features denoted by $F_i = \{f_1, f_2, \ldots, f_N\}$, where f_n is the n^{th} feature in the collection. From the graphical model specified in Figure 1, it can be easily find that:

$$P(\alpha_i, V_{\Gamma(i)}|F_i) = P(V_{\Gamma(i)}|\alpha_i)P(\alpha_i|F_i)$$

and

$$P(V_{\Gamma(i)}|F_i) = \int_{\alpha_i} P(V_{\Gamma(i)}|\alpha_i)P(\alpha_i|F_i)d_{\alpha_i}.$$

We assume that the distribution of review helpfulness is a Gaussian distribution: $P(\alpha_i|F_i) \sim \mathcal{N}\left(\alpha_i; F_i A^T, \sigma^2\right)$. Collectively, we denote $\theta = \{A, \sigma^2\}$. Consider the conditional joint distribution of votes $V_{\Gamma(i)}$ and helpfulness α_I, we have:

$$P(V_\Gamma, \alpha_I|F_I, \theta) = \prod_{i\in I} P(\alpha_i|F_i, \theta)P(V_{\Gamma(i)}|\alpha_i)$$

Maximizing $P(V_\Gamma|F_I, \theta)$ then leads to the following optimization problem:

$$\hat{\theta} = \operatorname*{argmax}_{\theta} \sum_{i\in I} \log \int_{\alpha_i} P(\alpha_i|F_i, \theta) \prod_{j\in\Gamma(i)} P(V^j_{\Gamma(i)}|\alpha_i)d\alpha_i$$

In this paper, we apply EM algorithm to estimate the $\hat{\theta}$ which maximize $P(V_\Gamma|F_I, \theta)$. In order to find the $\hat{\theta}$ to maximize the log-likelihood of $P(V_\Gamma|F_I, \theta)$, we define $q(\alpha_I)$ as an arbitrary distribution over the hidden variable α_I where E-step:

$$q^{t+1}(\alpha_I) = \prod_{i\in I} P(\alpha_i|F_i, V_{\Gamma(i)}, \theta^t)$$

and M-step:

$$\theta^{t+1} = \operatorname*{argmax}_{\theta} \int_{\alpha_I} q^{t+1}(\alpha_I) \log P(V_\Gamma, \alpha_I|\theta, F_I)d\alpha_I.$$

The algorithm alternates between E-step and M-step. The local optimality can be guaranteed by the convergence property of EM algorithm in [1]. Upon the convergence, we obtain the value of parameter θ. The next issue is how to use this model to predict the helpfulness value of online reviews. Based on Bayesian formula, we can rewrite this problem as, let V_i denote a random vote on D_i, then

$$P(V_i = 1|F_i) = \int_{\alpha_i} P(V_i = 1, \alpha_i|F_i)d\alpha_i$$

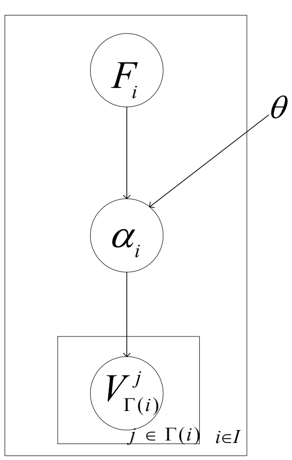

Figure 1: Detailed Graphical Model Representation

3. EXPERIMENTAL RESULTS

Table 1: Ranking Correlation of our probabilistic model, SVM Regression, ANN and Linear Regression (10-fold cross-validation).

	Ranking Correlation
Probabilistic Model	0.56778
SVM Regression	0.53874
ANN	0.53484
Linear Regression	0.51369

Our experiments focus on the product category of LCD HDTV. We crawled 501 LCD HDTV reviews from Amazon.com which have been evaluated by at least 10 consumers as helpful or not helpful. To validate the ranking performance of our model, we compare our probabilistic model with Support Vector Regression (SVR), Artificial Neural Network (ANN), and Linear Regression. We choose $\sigma = 0.2$ in our empirical study to estimate the helpfulness distribution. This can assure that most of the Gaussian distribution stays between (0,1). Table 1 shows the experimental results of our probabilistic model and SVR algorithm and Linear Regression algorithm. The results indicate that our model consistently outperforms SVR, ANN and Linear Regression in terms of Spearman's ranking correlation coefficient.

4. CONCLUSION

We have investigated the disadvantages of the conventional helpfulness definition and proposed a probabilistic approach to analyze the helpfulness distribution of online product review. To the best of our knowledge, this is the first model which is built to assess the helpfulness distribution of product reviews and does not rely on the simple positive vote fraction. We have shown that probabilistic model is a more generative model for the review documents helpfulness assessment than conventional models. Finally, we have executed an empirical study on the reviews of Amazon.com. The experimental results show that our model can effectively predict the review's helpfulness. By applying this model, helpful online reviews can be recommended to the possible consumers to assist them with their purchase decision making.

5. REFERENCES

[1] A. P. Dempster, N. M. Laird, and D. B. Rubin. Maximum likelihood from incomplete data via the em algorithm. *Journal of the Royal Statistical Society. Series B (Methodological)*, 39(1):1–38, 1977.

[2] N. Hu, L. Liu and J. J. Zhang. Do online reviews affect product sales? the role of reviewer characteristics and temporal effects. *Information Technology and Management*, 9(3):201–214, 2008.