Making Sales Strategies Based on the Existing Shopping Reviews

Jingcheng Shi1, *, Yan Yang2 and Shi Qiu1
1School of Information Science and Technology, Dalian Maritime University, Dalian, China
2School of Transportation Engineering, Dalian Maritime University, Dalian, China
*Corresponding author: shijingcheng@dlmu.edu.cn

Abstract. With the popularization of mobile Internet technology, people are increasingly inclined to choose to use the internet to purchase products. Consumers' evaluation of products after online shopping will directly affect the future sales of the products, so many online shopping platforms are committed to the research of intelligent evaluation and analysis models, so as to improve their sales strategies in a timely manner. We take the evaluation of three kinds of goods on Amazon as the research object. The platform allows purchasers to give star-rating and text-based review of products, and also lets them vote for the helpfulness of the former reviews. Therefore, we need to figure out the relation between these factors. We construct an innovative Maximum-Relation-Minimum-Redundancy Feature Selecting Model to analyze how the text-based measures and ratings-based measures affect the reviews, especially on some occasions at which the review has significant positive changes. We also discuss the connection between the specific quality descriptions of text-based reviews and the level of star-ratings. We use an innovative model, Bi-LSTM model to calculate the rate of different emotions contained in the review texts. Then we can find out a remarkably high correlation between certain descriptors and the star-ratings. This paper puts forward an intelligent and constructive scheme for e-commerce platform to adjust its own sales strategy based on customer's reviews on products.

Keywords: Bi-LSTM Model, Semantic Emotion Analysis, MRMR Model.

1. Introduction
Due to the rapid development of the Internet, information has become one of the determinants of a company’s efficiency [1, 2].

Customer reviews would help enterprises get real information about three products, thus enabling them to raise product quality to fit the customers’ needs. However, a large number of comments online make it difficult for an enterprise to extract useful information.

Amazon provides customers with the opportunity to rate and review purchases simultaneously. Customers can express their level of satisfaction with a product in the module called “Star Ratings”, and submit text-based messages in the module called “Reviews”. While other purchasers can submit ratings on these reviews for being helpful or not in the module called “Helpfulness Rating”.

Content from this work may be used under the terms of the Creative Commons Attribution 3.0 licence. Any further distribution of this work must maintain attribution to the author(s) and the title of the work, journal citation and DOI.
We used Amazon reviews of three products: a microwave oven, a baby pacifier and a hair dryer to suggest some sales strategies.

Since Amazon let the purchasers give a star rating and a text-based review on the products, and also let them remark on the helpfulness of the former reviews, we want to find out the relation among the factors [3].

Based on that, we designed a model to predict changes in product reputation based on star ratings and text-based reviews, as well as an algorithm to predict succeeded and failed products [4].

Finally, we analyzed the relation between certain descriptions of sentiments contained in the reviews and the level of star ratings [5].

2. Find Out Potential Successful Products

2.1. Text-based and rating-based Distribution Probability Analysis

To find out the potentially successful and failing products, we need to consider as many reviews as possible and combine them effectively.

Based on the Hierarchical Model and the Screening Model, we count and list the probability distribution of different star ratings and valid reviews. The result is as follows:

![Figure 1. Probability Distribution of star ratings and Valid Reviews.](image)

It can be seen in Fig 1, that the probability distributions of star ratings and reviews of the three products are similar. So we can use one factor to find out whether a product is a success or a failure. However, we can see from the figure that there are peaks and inflection points the number of positive reviews. That is to say, at some points, customers’ attitude toward the product changes greatly.

Therefore, we discuss the correlation between the former and the present star ratings and reviews at these points.

2.2. MRMR Feature Selection Model

To analyze how former comments impact on the rate of positive reviews at the particular points, we discuss the correlation between text-based measures and ratings-based measures and the change of positive-review-rate, namely, the correlation between the maximize feature and the categorical variables. However, when choosing the features, a single combination of positive features cannot improve the performance of a classifier [6]. Since the features of texts and star ratings are highly correlated, it would cause redundancy of feature variables. So we use the Maximum Relationship Minimum Redundancy Model to maximize the correlation between the features and the categorical variables and minimize the correlation between features.
2.2.1 Definition of Mutual Information. Taking time into regard, we set variables $x, y$ as the average level of text-based measures and ratings-based measures one week before the point. Their probability density functions are $p(x), p(y)$, and $p(x, y)$ respectively. Then the mutual information is

$$I(x, y) = \iiint p(x, y) \cdot \log \frac{p(x, y)}{p(x)p(y)} \, dx \, dy$$  \hspace{1cm} (1)$$

2.3. Model Construction

The correlation between the feature subset $S$ and the categorical variable $c$ depends on the average mutual information value features $f_i$ and class $c$

$$D(S, c) = \frac{1}{|S|} \sum_{f_i \in S} I(f_i; c)$$  \hspace{1cm} (2)$$

All the redundancy of the features in the subset $S$ is the average value of all the mutual information between $f_i$ and $f_j$

$$R(S) = \frac{1}{|S|^2} \sum_{f_i, f_j \in S} I(f_i; f_j)$$ \hspace{1cm} (3)$$

For ontology, reviews and star ratings are discrete variables, so we compute with the discrete model. The maximum correlation is

$$\max D(S, c), D = \frac{1}{|S|} \sum_{x_i \in S} I(x_i; c)$$ \hspace{1cm} (4)$$

The minimum redundancy is

$$\min R(S), R = \frac{1}{|S|^2} \sum_{x_i, x_j \in S} I(x_i; x_j)$$ \hspace{1cm} (5)$$

Then we multiply the two values above

$$\max \Phi(D, R), \Phi = D/R$$  \hspace{1cm} (6)$$

Then we can analyze the relationship between the two types of measures and the trend of positive reviews.
2.4. Model Analysis

![Figure 2. The Relationship between MRMR Model Features and Learning Accuracy.](image)

Fig 2 shows that the learning accuracy of ML-MRMR Feature Selecting Model based on Similar Matrixes is higher than that of which without feature selecting. When we choose all the features, the accuracy turns out to be the same. The accuracy of the MRMR Feature Selecting Model reaches a peak when we choose 20-60% features and goes down if we add the features. That is to say that there is some redundancy in the features which harm the accuracy.

We compute the average precision of different products- *Ave. Prec* and *Ranking Loss* to discuss their effects on the positive reviews. The higher the *Ave. Prec* is, the stronger the correlation is; the smaller the *Ranking Loss* is, the stronger the match is. The concrete values are as follows:

|               | Hairdryer   | Microwave  | Pacifier   |
|---------------|-------------|------------|------------|
| Ave. Prec↑    |             |            |            |
| Text          | 0.8835±8.73E-4 | 0.8701±2.49E-4 | 0.8605±3.48E-4 |
| Ratings       | 0.6557±4.18E-4 | 0.5964±2.12E-4 | 0.5653±1.53E-4 |
| Ranking loss↓ |             |            |            |
| Text          | 0.0379±1.65E-5 | 0.0434±2.84E-5 | 0.0490±4.26E-5 |
| Ratings       | 0.1245±2.18E-5 | 0.1272±9.68E-5 | 0.1575±7.43E-5 |

We can see from Table 1 that, for all three products, text-based measures impact the number of positive reviews more. That is to say, customers tend to depend more on the text of the reviews. Then we use the MMRM model to calculate the rates of the two types of measures and calculate the differential of the curve fitting of the number of positive reviews. Then we standardize the data. The result is as follows:

|               | Hairdryer   | Microwave  | Baby pacifier |
|---------------|-------------|------------|---------------|
| Standard diff. | 2.1234±0.0255 | 4.0480±0.1030 | -3.1426±0.0465 |

We can learn from Table 2 that the positivity of hairdryers and the microwave ovens are growing up, while that of baby pacifiers is decreasing. Therefore, we regard the hairdryer and the microwave oven as potentially successful products and regard the baby pacifier as the potentially failing product.
3. Association Between Specific Quality Descriptors and Rating Levels

3.1. Long Short-term Memory Model
To analyze the association between the emotion contained in the review texts and the star ratings, we carried out semantic emotion analysis. Long Short-term Memory (LSTM) is widely used to model the information in the context. To connect the words into sentences, we can add up the meaning of all the words or calculate the average. However, these methods cannot mention the sequence of the words. LSTM learns the information needed to be memorized or to be forgotten by training system, which helps to catch long-term relations and improve the stability of the emotion identification.

LSTM model consists of input word $X_t$, cell state $C_t$, temporary cell state $\tilde{C}_t$, invisible state $h_t$, forgetting gate $f_t$, memory gate $i_t$, and output gate $o_t$. The computing progress is pass the information useful for subsequent moments by forgetting the information in the cell state and memorizing the new information and put out invisible state every moment. The memorizing, the forgetting, and the output are determined by the invisible state of the last moment and the forgetting gate, the memorizing gate, and the output gate computed from the current input. The steps of the algorithm are as follows:

**Step1:** computing the forgetting gate - derive and select the information to be forgotten from the last layer.

**Step2:** computing the memorizing gate - select information to be memorized from step1.

**Step3:** computing the cell state at the moment - process the invisible information of the last moment.

**Step4:** computing the output gate and the invisible state of the moment - process the invisible information of the last moment.

**Step5:** circulate the steps and get the final results - get a series $\{h_0, h_1, \ldots, h_{n-1}\}$ whose length is the same as the sentence.

The equation of the function is

\[
\begin{align*}
    f_t &= \sigma(W_f \cdot (h_{t-1}, x_t) + b_f) \\
    i_t &= \sigma(W_i \cdot (h_{t-1}, x_t) + b_i) \\
    \tilde{C}_t &= \tanh(W_c \cdot (h_{t-1}, x_t) + b_c) \\
    o_t &= \sigma(W_o \cdot (h_{t-1}, x_t) + b_o) \\
    C_t &= f_t \cdot C_{t-1} + i_t \cdot \tilde{C}_t \\
    h_t &= o_t \cdot \tanh(C_t)
\end{align*}
\]

3.2. Bi-directional Long Short-term Model
To improve the success rate of emotion identification, we do further improvement on the LSTM model. We construct an innovative model combing the forward LSTM and the backward LSTM, thus the reverse-ordered sentences could be identified. The map of the model is as follows:

![Figure 3. The Map of Bi-LSTM Model.](image-url)
3.3. Model Solving

We import the processed information into the Bi-LSTM model and use Pytorch to compute it, getting the rate of different emotions in the reviews of the three products [7, 8].

![Figure 4. Emotions in the Reviews of the Products.](image)

Compared with the NLP algorithm above, using the Bi-LSTM algorithm would help find better results in which there are tighter correlations between the reviews and the star ratings.

We divide the reviews into five levels according to the emotion contained in the text. Then we classify them into three types according to the correlation between the star ratings and the reviews. We calculate the difference between the emotions level and the star ratings. Then we calculate the absolute value. The rules are as follows:

a. If the value is 0, the two factors are “highly correlated”.

b. If the value is 1, the two factors are “generally correlated”.

c. If the value is larger than 1, the two factors are “not correlated”.

Here we use the data processed above. We regard the “highly correlated” and the “generally correlated” ones as “successfully correlated”. Then we can get the following results:

|                | Highly correlated | Generally correlated | Not correlated | Correlation rate   |
|----------------|-------------------|----------------------|---------------|--------------------|
| Hairdryer      | 6281              | 994                  | 575           | 92.6752%           |
| Microwave Oven | 638               | 465                  | 82            | 93.0802%           |
| Baby Pacifier  | 6281              | 1094                 | 475           | 93.9490%           |

4. Conclusions

For the evaluation and analysis of online shopping platforms, we discussed the star-rating and text-rating of three products on Amazon, and put forward an innovative analysis model based on sentiments analysis. Firstly, we analyzed the relation between different review indexes, and based on this, we proposed an innovative Maximum-Relation-Minimum-Redundancy Feature Selecting Model, which is used to analyze the impact of text-based evaluation and rate-based evaluation on reviews. We further discussed the connection between specific quality descriptions for text-based reviews and star rating levels. We comprehensively used the NLP model and the Bi-LSTM model to judge the attitudes of review text and determined the relation between review text and star ratings.

In general, the model has strong robustness and stability, and is suitable for e-commerce platforms that use multiple review indexes in products evaluation. Meanwhile, it can carry out semantic sentiment analysis of remarks, which has certain application and promotion value.

References

[1] Næss P, Peters S, Stefansdottir H, et al. Causality, not just correlation: Residential location, transport rationales and travel behavior across metropolitan contexts [J]. Journal of Transport
Geography, 2018, 69: 181-195.

[2] Wu P F. Motivation Crowding in Online Product Reviewing: A Qualitative Study of Amazon Reviewers [J]. Information & Management, 2019.

[3] Cruz R A, Lee H J. The Effects of Sentiment and Readability on Useful Votes for Customer Reviews with Count Type Review Usefulness Index [J]. 2016.

[4] Xialu Liu, Rong Chen. Threshold factor models for high-dimensional time series [J]. Journal of Econometrics, 2020.

[5] Ying Qin, Yingfei Zeng. Research of Clinical Named Entity Recognition Based on Bi-LSTM-CRF [J]. Journal of Shanghai Jiaotong University (Science), 2018, 23 (03): 392-397.

[6] M. Toğacar, B. Ergen, Z. Cömert. A Deep Feature Learning Model for Pneumonia Detection Applying a Combination of mRMR Feature Selection and Machine Learning Models [J]. IRBM, 2019.

[7] opinion-lexicon-English attained from http://www.cs.uic.edu/~liub/FBS/sentiment-analysis.html

[8] Qinglong An, Zhengrui Tao, Xingwei Xu, Mohamed El Mansori, Ming Chen. A data-driven model for milling tool remaining useful life prediction with convolutional and stacked LSTM network [J]. Measurement, 2020, 154.