SPH Method For Fluid Simulation

Summary

We have witnessed the rise of mass online marketplaces. For example Amazon, one of the biggest online platforms, is worth around $ 915 billion. Guided by the customer obsession principle, it provides an opportunity for the customers to rate the products from 1 to 5. Moreover, buyers can submit a text-based message, namely review, to express their feeling towards the products. The massive data of those ratings and reviews offer a wealth of information remained to be mined. Analysis of text-based messages or rating-based values has received wide attention, yet there is not a method severs as the combination of both, especially for the case of an online marketplace.

**PART I:** We have witnessed the rise of mass online marketplaces. For example Amazon, one of the biggest online platforms, is worth around $ 915 billion. Guided by the customer obsession principle, it provides an opportunity for the customers to rate the products from 1 to 5. Moreover, buyers can submit a text-based message, namely review, to express their feeling towards the products. The massive data of those ratings and reviews offer a wealth of information remained to be mined. Analysis of text-based messages or rating-based values has received wide attention, yet there is not a method severs as the combination of both, especially for the case of an online marketplace.

**PART II:** We have witnessed the rise of mass online marketplaces. For example Amazon, one of the biggest online platforms, is worth around $ 915 billion. Guided by the customer obsession principle, it provides an opportunity for the customers to rate the products from 1 to 5. Moreover, buyers can submit a text-based message, namely review, to express their feeling towards the products. The massive data of those ratings and reviews offer a wealth of information remained to be mined. Analysis of text-based messages or rating-based values has received wide attention, yet there is not a method severs as the combination of both, especially for the case of an online marketplace.

**PART III:** We have witnessed the rise of mass online marketplaces. For example Amazon, one of the biggest online platforms, is worth around $ 915 billion. Guided by the customer obsession principle, it provides an opportunity for the customers to rate the products from 1 to 5. Moreover, buyers can submit a text-based message, namely review, to express their feeling towards the products. The massive data of those ratings and reviews offer a wealth of information remained to be mined. Analysis of text-based messages or rating-based values has received wide attention, yet there is not a method severs as the combination of both, especially for the case of an online marketplace.

Our framework shows a strong accuracy, robustness. It can be easily implemented to other

data with our source codes.

**Keywords:** Text-Based Measure, Informative Text Selection, Reputation Quantification, Sales

Strategy Formation.

**Contents**

[1 Introduction 3](#_Toc56441471)

[2 Assumptions and Notations 4](#_Toc56441472)

[2.1 Assumptions 4](#_Toc56441473)

[2.2 Notations 4](#_Toc56441474)

[3 [todo] Model 4](#_Toc56441475)

[4 [todo] Model 6](#_Toc56441476)

[5 [todo] Model 7](#_Toc56441477)

[5.1 Simulation 7](#_Toc56441478)

[5.1.1 Detail 7](#_Toc56441479)

[6 Strengths and Weaknesses 8](#_Toc56441480)

[6.1 Strengths 8](#_Toc56441481)

[6.2 Weaknesses 8](#_Toc56441482)

[7 Conclusion 8](#_Toc56441483)

[References 10](#_Toc56441484)

[Appendices 11](#_Toc56441485)

[Appendix A 11](#_Toc56441486)

[Appendix B 11](#_Toc56441487)

# Introduction

Our society has witnessed the rise of many online marketplaces, with a total worldwide market value of 4.3 trillion dollars[1]. One salient feature of the online marketplace compared with traditional platforms is the massive review of texts and ratings. Among all of them, Amazon has received the most attention, as its greatest success[1]. Amazon also provides customers with chances to freely express their feeling and rate the products that they have purchased.

Previous work[2] indicates that customers will largely refer to the reviews and ratings before they buy the product on the platforms. Platforms can adjust their sales strategy by checking these comments. Hence, the ratings and the reviews both provide references to other potential buyers and massive data to analyze the demand of the customers, which can help to develop adaptive strategies. By making full use of these data, we can achieve a win-win situation for both the buyers and the platform.

One of the biggest challenges is the complexity and diversity of the texts of the reviews. In this paper, we propose a novel sentiment analysis model as the text-based measure to address this issue. In this paper, we develop a series of models as the combination of text-based, rating based, and time-based measures to pick out the most informative ratings and reviews to track. We also construct a novel evaluation framework to quantify the reputation of each product and predict potential success or failure. Then, we analyze the correlation between continuous same star ratings, word descriptors and the reputation of the products. We implement our model on the real data set generated from three different types of products, namely the pacifier, microwave, and the hair dryer.

# Assumptions and Notations

## Assumptions

To simplify our model and eliminate the complexity, we make the following main assumptions in this literature. All assumptions will be re-emphasized once they are used in the construction of our model. （假设部分）

**Assumption 1.** *The online marketplace operates stably. And there were no situations such as an outbreak of an epidemic which would seriously affect the production chain of online shopping.*

**Assumption 2.** *The ratings and reviews depict customers’ real experience and feeling about their purchased products. The sentiment in the review text reflects one’s feelings on the products.*

## Notations

In this work, we use the nomenclature in Table 1 in the model construction. Other none-frequent-used symbols will be introduced once they are used. （文中用到的变量的意思）

Table Notations used in this literature

|  |  |  |
| --- | --- | --- |
| Symbol | Definition | Type |
|  | review id | String |
|  | Star rate, subscript is its associated review id | Scalar |
|  | Review date, subscript is its associated review id | Date |
|  | Vector encoding of the star rating | Mapping |

# [todo] Model

图的风格尽量统一，建议完全参考python的plot默认配色（白蓝橙绿）

蓝：#1E76B3 橙：#FE7E0D 绿：#8FB132

图片和Figure绝不能跨页，因为非常不好看，当出现跨页时，务必调整一下前后文内容：

小图不大于整体文本框的60%。（自己估量）

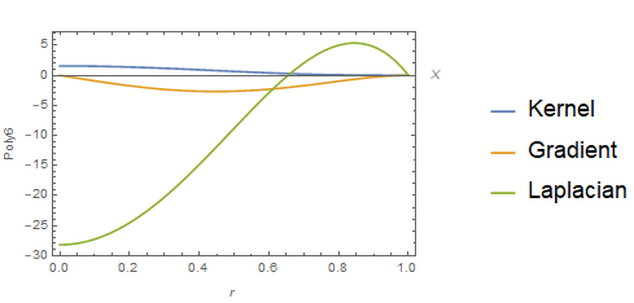


Figure : Demonstration of some representative seed words. Words annotated as "positive" are colored in light tone while the "negative" ones in a dark tone. The bigger the size, the higher word frequency. （这些是对图片的解释部分）

大图尽量不大于80%，以看清为第一标准（自己估量）

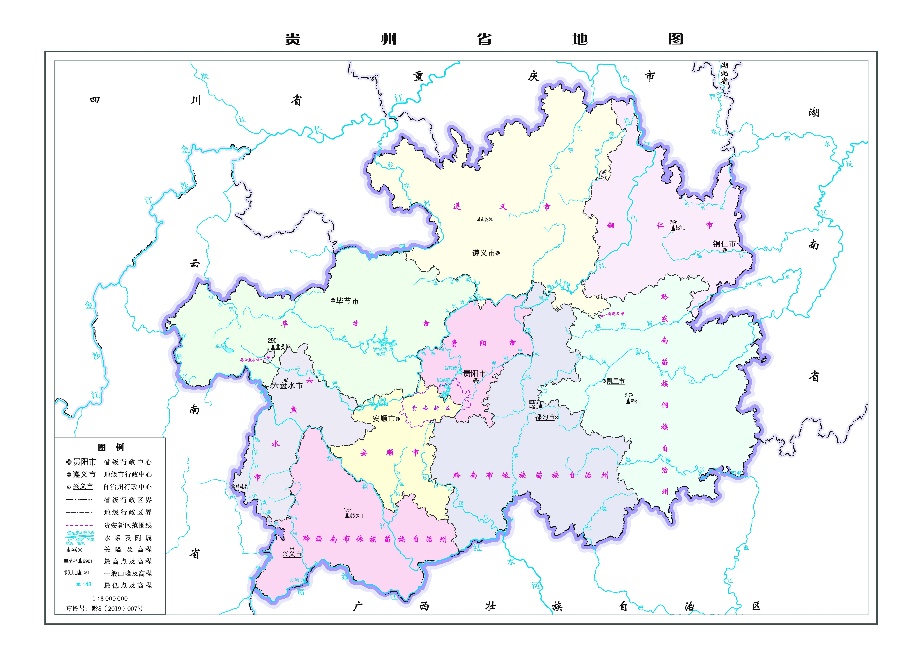


Figure : Demonstration of some representative seed words. Words annotated as "positive" are colored in light tone while the "negative" ones in a dark tone. The bigger the size, the higher word frequency. （这些是对图片的解释部分）

多图相关：

真多张图片：

每个图片中用**一个空格**分隔，尽量不大于80%，以看清为第一标准（自己估量）

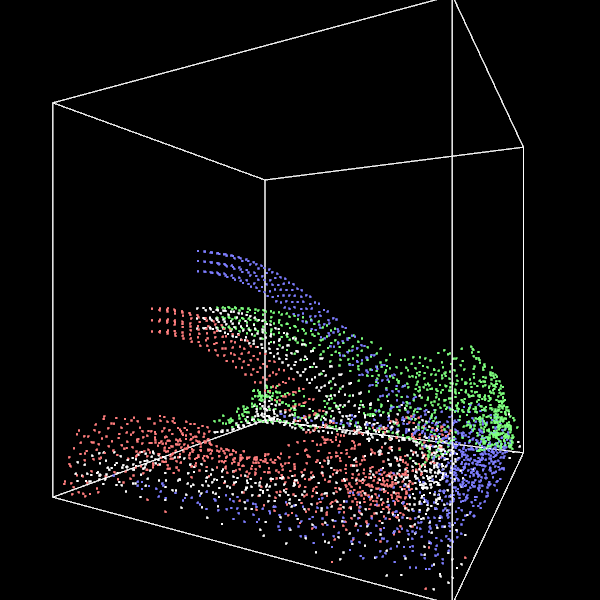
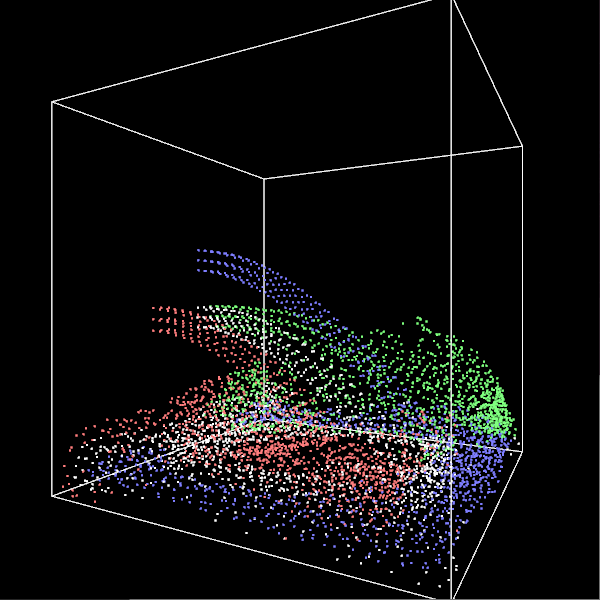
 

Figure : Demonstration of some representative seed words. Words annotated as "positive" are colored in light tone while the "negative" ones in a dark tone. The bigger the size, the higher word frequency. （这些是对图片的解释部分）

还有一种大图，这个就看内容是否看清，尽量不大于80%：

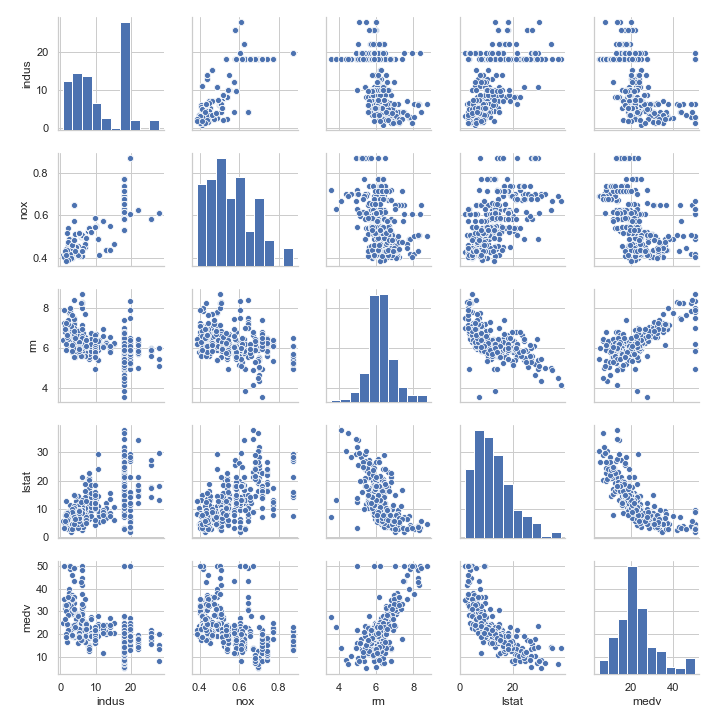


Figure : Demonstration of some representative seed words. Words annotated as "positive" are colored in light tone while the "negative" ones in a dark tone. The bigger the size, the higher word frequency. （这些是对图片的解释部分）

# [todo] Model

公式相关：

公式统一用表格模拟，因为这样可以实现编号，就如下图所见：

|  |  |
| --- | --- |
|  |  |

隐藏框线就可以得到一个好看的公式了：

|  |  |
| --- | --- |
|  |  |

当你不需要带编号的公式时，删除后面框内的编号就可以了，就如下式子所示：

|  |  |
| --- | --- |
|  |  |

还有种公式就是句内公式，这个由于句间距的限制，不能用复杂的式子填充，尽量填充如。

公式对齐，语法是：

\eqarray<空格> (你的公式)

e.g. \eqarray<空格> (A &= B+C @ B+C+1 &= D @ D&=2)

下图是一个对齐加号的效果：

|  |  |
| --- | --- |
|  |  |

**Where** 对公式部分的解释。

写完后若公式没转换，按shift + enter生成即可（别忘了删掉生成后多出来的那个小东西）

# [todo] Model

## Simulation

### Detail

我将表格竖直线隐藏了.jpg

参数：上线2.25磅，底线1.5磅，中间1磅，内容尽量用居中，除非不好看。

推荐Excel画好后粘入word。

全选表格后，拖动上面游标中最后那个拖动条，可整体水平缩放表格。

下面是大中小表格大小的参照：

Table small table

|  |  |  |
| --- | --- | --- |
| Symbol | Definition | Type |
|  | review id | String |
|  | Star rate | Scalar |
|  | Review date | Date |
|  | Vector encoding | Mapping |

Table middle table

|  |  |  |
| --- | --- | --- |
| Symbol | Definition | Type |
|  | review id | String |
|  | Star rate | Scalar |
|  | Review date | Date |
|  | Vector encoding | Mapping |

Table big table

|  |  |  |
| --- | --- | --- |
| Symbol | Definition | Type |
|  | review id | String |
|  | Star rate, subscript is its associated review id | Scalar |
|  | Review date, subscript is its associated review id | Date |
|  | Vector encoding of the star rating | Mapping |

# Strengths and Weaknesses

## Strengths

1. **Novelty.** To the best of our knowledge, we are the first to propose a CE-VADER hybrid model for review text-based sentiment evaluations on an online marketplace.
2. **Accuracy.** The maximum RMSE of the 50 order AR model is 0.031 on the validation time domain. The text-based measures correlate well with the rating-based measures.
3. Generalization. Our proposed framework can freely be implemented to any data set e.g., reviews and star-ratings of any products from any online platforms.
4. Robustness. Our model shows great robustness to most of the parameters.

## Weaknesses

1. **Time consuming manual annotations.** Annotating the seed words generated from reviews for CE-VADER model manually is time-consuming.
2. **CE-VADER model can not identify different forms of the same word with special variation rules.** CE-VADER model cannot identify the past form of verbs, plural of nouns and comparative forms of adjectives with special variation e.g., is (v.s. was), children (v.s. child) and better (v.s. good).

# Conclusion

To crack the secret of Amazon’s ratings and reviews, we proposed a series of novel models to address the sub-issues from selecting the most informative reviews to identifying reviews’ quality descriptors. The proposed model achieves high accuracy and robustness.

1. Information Evaluation Model can combine the text-based measure with the rating-based measure, where we propose a novel CE-VADER hybrid model for the sentiment analysis as the text-based measure. We can rank how informative each review and the rating is with the proposed model. The informative rate correlates to the helpful votes. To be more specific, the more help votes there the more five-star ratings they own and the longer review bodies are, the more likely they are evaluated as informative reviews. However, PLEASE REMEMBER that moderately-rated reviews with high information entropy, which have both positive and negative comments, marked by words like “however” and “but” also have great reference value.
2. We employ the Difference Equation Model to construct a “reputation rate” to quantify the reputation of three products, namely baby pacifier, microwave and hair dryer. Baby pacifier has a positive reputation, hair dryer has a weak negative one, and the microwave has the worst reputation. With a modified AR algorithm, we predicted the future reputation tendency of these three products.
3. In analyzing the distribution of star ratings and specific words, we identified special review descriptors by employing a continuous extreme rating and a set of special words. Continuous extreme ratings can obviously affect the total sale volume and special words’ appearance can judge the rating of reviews with high probability.

# References

1. Paleontology; Findings on Paleontology Detailed by Investigators at National University of La Pampa (Recognition of Fossil Nebkha Deposits: Clues From Neoichnology and Sedimentology). 2020, :414-.
2. M. Cristina Cardonatto, Ricardo Nestor Melchor. RECOGNITION OF FOSSIL NEBKHA DEPOSITS: CLUES FROM NEOICHNOLOGY AND SEDIMENTOLOGY. 2020, 35(7):277-291.

# Appendices

## Appendix A

## Appendix B