International Journal of Informatics, Information System and Computer Engineering

Aspect-Based Sentiment Analysis on Amazon Product Reviews

Muhammad Abubakar, Amir Shahzad, Husna Abbasi*

COMSATS University Islamabad Abbottabad Campus Pakistan, Pakistan.
*Corresponding Email: abubakarhameedch@gmail.com

ABSTRACTS

The focus of this paper was on Amazon product reviews. The goal of this is to study is two (NLP) for evaluating Amazon product review sentiment analysis. Customers can learn about a product's quality by reading reviews. Several product review characteristics, such as quality, time of evaluation, material in terms of product lifespan and excellent client feedback from the past, will have an impact on product rankings. Manual interventions are required to analyse these reviews, which are not only time consuming but also prone to errors. As a result, automatic models and procedures are required to effectively manage product reviews. (NLP) is the most practical method for training a neural network in this era of artificial intelligence. First, the Naive Bayes classifier was used to analyse the sentiment of consumer in this study. The (SVM) has categorized user sentiments into binary categories. The goal of the approach is to forecast some of the most important characteristics of an amazon-based product reviews, and then analyse Customer attitudes about these aspects. The suggested model is validated using a large-scale real-world dataset gathered specifically for this purpose. The dataset is made up of thousands of manually annotated product reviews gathered from amazon. After passing the input via the network model, (TF) and (IDF) pre-processing methods were used to evaluate the feature. The outcomes precision, recall and F1 score are very promising.

ARTICLE INFO

Article History:
Received 18 Dec 2021
Revised 20 Dec 2021
Accepted 25 Dec 2021
Available online 26 Dec 2021

Keywords:
Naïve bayes, Text Classification Algorithms , Natural Language Processing, Support Vector Machines, NLP, SVM

1. INTRODUCTION

Amazon is the largest online retailer in the world, as well as a significant cloud computing service provider (Rain, 2013). The company began as a book seller but has now evolved to include a wide range of consumer items and digital media, including the Kindle e-reader, Kindle Fire tablet, and Fire TV., a streaming media adaptor are among the company's own electronic devices. People nowadays prefer to trade things on an e-commerce website rather than at a physical store because of the time savings and convenience (Bhat et al., 2015). Before purchasing a product, it is usual practice to read the product review. The consumer's opinion of the product has been swayed either positive or negative by the reviews. Thousands of reviews were read, on the other hand, is an unnatural feat. In this era of ever-improving natural language processing algorithms, it takes time to wade through hundreds of comments to identify a product that uses a polarized review of a specific category to assess its popularity among consumers all around the world. This project aims to categorize customers' positive and negative product reviews, as well as construct a supervised learning model to polarize a wide range of reviews. According to an Amazon study from last year, 88% customers from internet trust reviews as much as a personal suggestion (Dey et al., 2020).

With a powerful remark, the credibility of an internet product with a high number of positive reviews is established. The absence of reviews, books, or any other thing on the internet creates a sense of distrust among potential customers. Pre-processing is used in this study to minimize the Multi-Domain Sentiment Dataset's dimensionality of the features applied. Following that, any frequent words above a certain threshold value are considered characteristics (Haque et al., 2018).

2. RELATED WORK

This section presents the results of the classic schema polarization analysis based on user reviews on the Amazon ecommerce website (Xiao et al., 2021). The criteria for compositional sentiment were set by Zhang et al. To find out how much textual sentiment there is. The system makes clear use of machine learning. In this work, film reviews were classified into binary classes using (SVM) and Naive Bayes classifiers (Joseph, 2020). The accuracy of the Naive Bayes model has been improved, while the SVM model has been extended. To summarize, there have been no studies comparing (SVM) with the Naive Bayes classifier. A comparison of two approaches (NLP) to analyze Amazon product evaluation sentiment is presented in this study (More et al., 2020). Comparative polarity analysis on Amazon product reviews using algorithms has also been carried out to evaluate the sentiment of Amazon product reviews (Karthikayini et al., 2017). In his research, Dadhich uses a rule-based hybrid to be able to create an automatic comment analyzer (Dadhich et al., 2022). Salmony also conducted a survey on amazon product reviews to assist in customer decision making (Salmony et al., 2021).

3. METHODOLOGY

Amazon, as seen by the numerous evaluations accessible, is one of the most well-known e-commerce companies. The dataset was unlabeled, thus it needed to

be labelled before it could be used in a supervised learning model (Pandey, 2019). Only Amazon product feedback, specifically book feedback, was used for this study activity. To evaluating polarization, about 1, 47,000 book evaluations were analyzed. Data collecting was completed as the first step in the data labelling process. Manual labelling is impractical for a human to do because the dataset contains a high number of reviews. The term (TF) and (IDF), elimination of relevant nouns and frequent noun identifier methods were used to extract the dataset's features (Jagdale et al., 2019). TF-IDF: TF-IDF is a retrieval strategy that considers the frequency of a phrase (TF) as well as the (IDF). TF and IDF scores are assigned to each word or phrase. The TF and IDF product results of a term, on the other hand, refer to the TF-IDF weight of that term. As a result, the TF of a word represents its frequency, whereas the IDF is a metric for what percentage of the corpus is occupied by a term. The content will always be among the top search results if words have a high TF-IDF content weight, allowing anyone to avoid stop words while also effectively locating words with a higher search volume but a lower level of competition (Fang, 2015).

4. RESULTS AND DISCUSSION

The purpose of this part is to assess the experiment's performance. Evaluating metrics are important in determining classification efficiency, and assessing accuracy is the easiest way to do so. The system is assessed using three widely used statistical measures: The F-measure, which is generated from a confusion matrix, is derived from recall, precision, and the F-measure. The confusion matrix divided into four categories True Positive, True Negative, False Positive, and False Negative (See figures 1 and 2). True positive describes a situation in which the system accurately anticipates the positive class. False-positive highlights a situation in which the scheme predicts the positive class inaccurately. Tabulator form is used to show the (SVM) confusion matrix and the Naive Bayes Classifier A separate tabular format is used to display both the statistical measurement and the NPL (Table 1).

Table 1. SVM confusion matrix

Positive	3694
Neutral	158
Negative	90

In the train dataset, we have 3694 (~95.1%) sentiments labelled as positive, and 158 (~4%) sentiments labelled as Neutral and 90(~2.35%) sentiments as Negative. So, it is an imbalanced classification problem.

Naive Bayes

[[0 0 24]

[0 0 39]

[0 0 937]]

	Precision	Recall	f1-score	support
0	0.00	0.00	0.00	24
1	0.00	0.00	0.00	39
2	0.94	1.00	0.97	937

Micro avg	0.94	0.94	0.94	1000
Macro avg	0.31	0.33	0.32	1000

Weighted avg	0.88	0.94	0.91	1000

Accuracy: 93.7

Precision refers to the ratio of predicted positive cases to total positive instances indicated by the equation.

TF/IDF Vectorizer and logistic regression for under sampled data

[[10 6 8]
[15 7 17]
[314 195 428]]

	Precision	recall	f1-score	support
0	0.03	0.42	0.06	24
1	0.03	0.18	0.06	39
2	0.94	0.46	0.62	937
Micro avg	0.45	0.45	0.45	1000
Macro avg	0.34	0.35	0.24	1000
Weighted avg	0.89	0.45	0.58	1000

Accuracy: 44.5

Characteristic of logistic regression of under sampled data

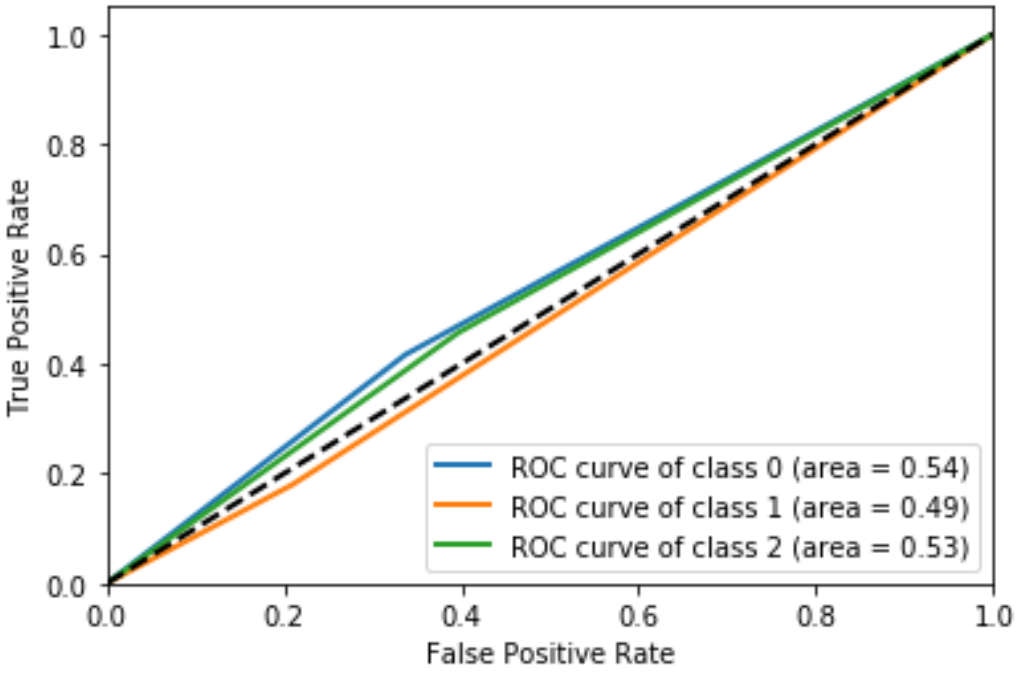

Figure 1. True and false positive rate under sampled data

TF/IDF and Logistic regression for over sampled data

[[13 3 8]
[10 10 19]
[214 171 552]]

	Precision	recall	f1-score	support
0	0.05	0.54	0.10	24
1	0.05	0.26	0.09	39
2	0.95	0.59	0.73	937
Micro avg	0.57	0.57	0.57	1000
Macro avg	0.35	0.46	0.31	1000
Weighted avg	0.90	0.57	0.69	1000

Accuracy: 57.49999999999999

Logistic Regression on over-sampled data is performing better than under-sampled data.

Characteristic of logistic regression of over sampled data

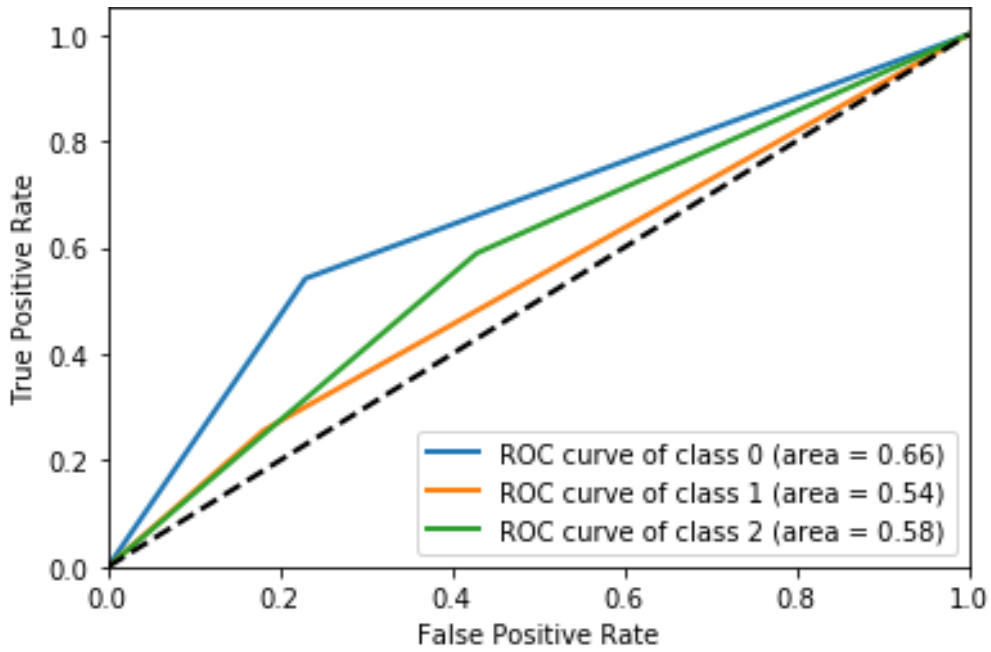

Figure 2. True and false over sampled data

Neural Network

[[9 2 13]

[0 12 27]

[2 8 927]]

	Precision	recall	f1-score	support
0	0.82	0.38	0.51	24
1	0.55	0.31	0.39	39
2	0.96	0.99	0.97	937
Micro avg	0.95	0.95	0.95	1000
Macro avg	0.77	0.56	0.63	1000
Weighted avg	0.94	0.95	0.94	1000

Using class-weights does not improve the performance.

3. CONCLUSION

In order to investigate the polarisation of Amazon product ratings, this study was able to compare SVM and Naive Bayes classifiers. Following the pre-processing step, almost 2250 features and over 6000 datasets were used to train the models. The SVM classifier in this system has a precision of 0.00 percent, a recall of 0.00 percent, f1 score 0.00 percent. The model yields SVM and Naive Bayes with 93.7 percent accuracy, respectively, which is confirmed to be superior to traditional approaches. With a higher accuracy rate, the (SVM) can polarise Amazon product feedback, according to the findings of experiments.

REFERENCES

Bhatt, A., Patel, A., Chheda, H., & Gawande, K. (2015). Amazon review classification and sentiment analysis. *International Journal of Computer Science and Information Technologies*, *6*(6), 5107-5110.

Dadhich, A., & Thankachan, B. (2022). Sentiment analysis of amazon product reviews using hybrid rule-based approach. In *Smart Systems: Innovations in Computing* (pp. 173-193). Springer, Singapore.

Dey, S., Wasif, S., Tonmoy, D. S., Sultana, S., Sarkar, J., & Dey, M. (2020, February). A comparative study of support vector machine and Naive Bayes classifier for sentiment analysis on Amazon product reviews. In *2020 International Conference on Contemporary Computing and Applications (IC3A)* (pp. 217-220). IEEE.

Fang, X., & Zhan, J. (2015). Sentiment analysis using product review data. *Journal of Big Data*, *2*(1), 1-14.

Haque, T. U., Saber, N. N., & Shah, F. M. (2018, May). Sentiment analysis on large scale Amazon product reviews. In 2018 IEEE international conference on innovative research and development (ICIRD) (pp. 1-6). IEEE.

Jagdale, R. S., Shirsat, V. S., & Deshmukh, S. N. (2019). Sentiment analysis on product reviews using machine learning techniques. In *Cognitive informatics and soft computing* (pp. 639-647). Springer, Singapore.

Joseph, R. P. S. (2020). *Amazon Reviews Sentiment Analysis: A Reinforcement Learning Approach* (Doctoral dissertation, MS Thesis, Griffith College Dublin, Ireland).

Karthikayini, T., & Srinath, N. K. (2017, December). Comparative polarity analysis on Amazon product reviews using existing machine learning algorithms. In *2017 2nd International Conference on Computational Systems and Information Technology for Sustainable Solution (CSITSS)* (pp. 1-6). IEEE.

More, G., Behara, H., & Suresha, A. M. (2020). Sentiment Analysis on Amazon Product Reviews with Stacked Neural Networks. *no. October*.

Pandey, P., & Soni, N. (2019, February). Sentiment analysis on customer feedback data: Amazon product reviews. In *2019 International Conference on Machine Learning, Big Data, Cloud and Parallel Computing (COMITCon)* (pp. 320-322). IEEE.

Rain, C. (2013). Sentiment analysis in amazon reviews using probabilistic machine learning. *Swarthmore College*.

Salmony, M. Y. A., & Faridi, A. R. (2021, April). Supervised Sentiment Analysis on Amazon Product Reviews: A survey. In *2021 2nd International Conference on Intelligent Engineering and Management (ICIEM)* (pp. 132-138). IEEE.

Xiao, Y., Qi, C., & Leng, H. (2021, March). Sentiment analysis of Amazon product reviews based on NLP. In *2021 4th International Conference on Advanced Electronic Materials, Computers and Software Engineering (AEMCSE)* (pp. 1218-1221). IEEE.