

Sentimental Analysis on Amazon Product Review

Sadiul Arefin Rafi, Bedria Durdana, Muhtasim Fuad Ayon, Ashfaqur Rahman, SK Safoan, Mohammed Jufikar Ali Mahbub, Md Motahar Mahtab and Annajiat Alim Rasel

BRAC University, 66 Mohakhali, Dhaka 1212, Bangladesh
sadiul.arefin.rafi@g.bracu.ac.bd

Abstract. With time, most tasks have become digitized. The growth in the use of the internet along with smart devices brings all decision-making processes to the fingertips. Online shopping has become an easier alternative to in-person shopping now more than ever before. When customers are not able to rely on physical features to confirm a purchase, the next source of reliability comes from online product reviews. Such reviews not only help consumers but also help businesses make predictions about future trends. With the use of NLP, comments in bulk can be categorized with ease. This paper is based on an Amazon product review dataset focusing on gadgets. After tokenizing, in order to prepare the data for the machine learning model, reviews were transformed into numerical values. Ratings, content, and support votes are sectioned using a binary classifier to separate positive and negative reviews. The WordNet Lemmatizer from the NLTK library was used to make semantic associations between words. The NLTK POS tagger was used for nouns, verbs, and adjectives while treebank tags were mapped to WordNet parts of speech. In order to bring together distinct reviews with similar sentiments, the K-means clustering approach was used. The results showed that the logistic model had an accuracy of 0.892 and a TF-IDF score of 0.88 on gadgets, while the Naive Bayes model had comparable values of 0.887 and 0.884. In this paper, models that perform this type of sentiment analysis will be explored in order to prove its usefulness for upcoming predictions.

Keywords: NLP, sentiment analysis, POS tagger, tokenization, lemmatization, K-means clustering

1 Introduction

In this day and age, online shopping is becoming more and more popular due to the amount of convenience that comes along with it. Smart devices allow using the internet with ease for all generations. Throughout the pandemic, with people being stuck within the walls of their homes, the advantages of online shopping have come to the surface more than ever before. The availability of online services became the only way to survive for numerous businesses. However, e-commerce websites that were popular long before 2020 were able to see the jump in sales even more clearly.

Amazon reached annual revenue of more than 457 billion dollars at the end of 2021. This was almost a 60% increase since 2019. As a trusted portal that has gained an immense amount of exposure, it became a valued option for many online shoppers.

When it comes to not being able to see or touch products, customer reviews hold the most weight for ensuring a purchase. It makes it easier for consumers to compare between brands and prices as well as an overall source for reliability. Not only do such reviews make the shopping experience smoother for customers, but it also helps the seller understand the customer's needs. With this information, predictions and strategies for future sales can be made in a systematic manner.

One application for natural language processing is sentiment analysis, also known as opinion mining. With the help of computational linguistics and text analysis, opinion mining is able to determine the emotion behind different types of writing. From a generic angle, the text is classified into two types. A text can be either an opinion or a fact. "The bag is red", would be a fact. "The bag is not useful", would typically be more of an opinion. Customer reviews will mostly consist of opinions that can be categorized as good, bad, or neutral. This paper will be using a dataset of Amazon product reviews upon which sentiment analysis will determine the type of comments from various customers.

2. **Related Works**

A significant part of the research on product reviews, sentiment analysis, and opinion mining has just recently been finished. Elli, Maria, and Yi-Fan accumulated other people's judgment from surveys and dissected the outcomes to make a plan of action in their work.[1] They said that the instruments they displayed were sufficiently dependable to furnish them with high precision. Their judgment was made more appropriate thanks to the use of business analytics. They worked on recognizing emotions in reviews, gender-from names, and extracting bogus reviews, among other things. Python and R were the most often utilized programming languages. As classifiers, they mainly used support vector machine(SVM) and Multinomial Naive Bayesian (MNB). In the paper [2] the authors utilized various classifiers to determine and precision the recall value. The author of Paper [3] considered information from Amazon audit datasets to apply and used them to improve present work in the domain of NLP and sentiment analysis. In order to classify a review as favorable or negative, Naive Bayesian and the list of decisions that's supposed to be classified were taken advantage of. They have chosen books and reviewed the Kindle area on Amazon. In the paper [4] the author claimed of using NV, SVM, and maximum entropy. They displayed their output in the statistical chart and no accuracy

is being shown. The authors of the paper [5] fostered a model for anticipating item appraisals in light of rating text with the help of a sack of words. Unigrams and bigrams were used in the models that were tested. The most precise output was obtained using unigrams when compared to bigrams. Because of their higher variance, popular unigrams were an excellent predictor of ratings. Unigram results outperformed bigrams by 15.89%. The algorithm that the author of Paper [6] used was the naive Bayes classifier algorithm. It sadly failed to provide any adequate outcome. The Paper used [6] less complex algorithms to make it easier to understand. Because the system provides high accuracy on SVM, it cannot work efficiently on large-scale datasets. The authors employed the support vector machine (SVM), logistic regression, and decision trees methods. As an extra examination Tf-idf has been used in paper [6]. Using a bag of words, it can predict rating. However, there are only a few Classifiers used here. They used a linear regression model with root mean square error.

3. Methodology

3.1 Data Preprocessing

The dataset was collected from Kaggle. It includes Amazon item reviews and information. Reviews (ratings, content, and support votes), item metadata (portraits, class data, value, brand, and image highlights), and linkages are all included in this collection. The following dataset contains mainly the reviews of gadgets such as Amazon kindle, Bluetooth speaker, TV, and Tablet. 'Reviewer ID', 'ASIN', 'Reviewer Name', 'Review title', 'Helpful', 'Summary', 'Rating', and 'Review time' are all properties in the files. It indicates that group ranks range from 1 to 5. We'll need to transform these ratings into two categories, 1 and 0 because we are going to use a binary classification model. Positive (1) ratings are those that are greater than or equal to three, while negative ratings (0) are those that are less than three. Because machine learning models work with numerical characteristics, we'll need to transform our review column into numerical values before we start creating our model. We will tokenize the text and conversions into tokens. Tokenization means the task of breaking down the text into smaller parts which can be words, characters, or subwords. Then we will remove the stop words. Things that are present in a phrase that is not required in any text mining area are stop words. We usually omit certain terms to improve the accuracy of the analysis. And Finally POS tagging. Parts of speech tag (POS) is a particular label applied to each token in a text corpus to denote the part of speech and, in certain cases, additional grammatical categories such as tense, number (plural/singular), case, and so on.

3.2 Model

The input words must be tagged with their senses in supervised algorithms for word sense disambiguation. The word's meaning is the label that has been attached to it. Words are not marked with their senses in unsupervised procedures, which must be inferred using other techniques. The context-sense set is used as the training data in supervised approaches such as naive Bayes (or any classifier). The process of lemmatization is the determination of two words that have the same root regardless of their surface difference. The NLTK library's Wordnet Lemmatizer was used to create structured semantic associations between words. For tagging adjectives, nouns, and verbs, the NLTK POS tagger was used immediately without any training, and treebank tags were mapped to WordNet part of speech names. For modeling, the final lemmatized textbase was divided into train and test datasets. To categorize reviews as positive or negative, logistic regression and naive Bayes classifiers were compared. Few product IDs were filtered out for the narratives, and feelings were applied to their text vectors. We utilized the K-means clustering approach to create narratives, grouping several reviews for an ASIN based on their feelings such that each cluster has distinct reviews with similar sentiments.

3.3 Results

On the bag of words, the logistic model had an accuracy of 0.892 and a TF-IDF score of 0.88 on gadgets, while the Naive Bayes model had comparable values of 0.887 and 0.884. The ratio of $tp / (tp + fp)$ is called precision, the meaning of tp and fp is true positive and false positive respectively. Precision gives the classifier the ability to avoid labeling a negative sample as a positive sample. The weighted harmonic mean of accuracy is the F-beta score where 0 is the lowest value and 1 is the highest value.

4 Conclusion

To conclude, reviews of products can vary over different factors and places. In the research conducted, we can see the utilization of models like Multinomial Naive Bayes and Support Vector Machine both of which provided an almost precise result. As e-commerce sites are determined to keep their sensitive data intact, it is quite a hassle to gather essential datasets in this regard. Nonetheless, the research has been regulated by popular electronic devices and their user reviews. Algorithms like Active Learning, POS Tagging, and TF-IDF have been proved quite convenient in retrieving product reviews. Also, processes like PCA are being developed to redeem data

comfortably. Therefore, the concept of using distinct models and algorithms to analyze the reviews is constantly updating and will help greatly in future endeavors.

References

1. Elli, Maria Soledad, and Yi-Fan Wang. "Amazon Reviews, business analytics with sentiment analysis." 2016
2. Xu, Yun, Xinhui Wu, and Qinxia Wang. "Sentiment Analysis of Yelp,s Ratings Based on Text Reviews." (2015).
3. Rain, Callen. "Sentiment Analysis in Amazon Reviews Using Probabilistic Machine Learning."Swarthmore College (2013)
4. Bhatt, Aashutosh, et al. "Amazon Review Classification and Sentiment Analysis." International Journal of Computer Science and Information Technologies 6.6 (2015): 5107-5110.
5. Nasr, Mona Mohamed, Essam Mohamed Shaaban, and Ahmed Mostafa Hafez. "Building Sentiment analysis Model using Graphlab." IJSER, 2017
6. Text mining for yelp dataset challenge; Mingshan Wang; University of California San Diego, (2017)
7. Amazon revenue 2006-2021: AMZN. Macrotrends. (n.d.). Retrieved January 15, 2022, from <https://www.macrotrends.net/stocks/charts/AMZN/amazon/revenue>
8. Klebnikov, S. (2021, June 28). 5 big numbers that show Amazon's explosive growth during the coronavirus pandemic. Forbes. Retrieved January 15, 2022, from <https://www.forbes.com/sites/sergeiklebnikov/2020/07/23/5-big-numbers-that-show-amazons-explosive-growth-during-the-coronavirus-pandemic/?sh=47798fe04137>
9. LIU, B. I. N. G. (2020). *Sentiment analysis: Mining opinions, sentiments, and emotions*. CAMBRIDGE UNIV PRESS.
10. Muhammad Ihsan Zul, Feoni Yulia, Dini Nurmallasari, "Social Media Sentiment Analysis Using K-means and Naïve Bayes Algorithm", 2nd International Conference of Electrical Engineering and Informatics(Icon EEI 2018), October 2018