SAPR - Sentiment Analysis of Product Reviews

D. Ramana Kumar[1]
[1]*Assistant Professor, Department of Computer Science and Engineering, Anurag University*
Email- ramanacse@anurag.edu.in
Sowdari Akshay Kumar Goud[2]Yennam Laxmi Pranaya[3]Potta Bhanu Priyanka[4], IthagoniMounika[5]
[2,3,4,5]*Department of Computer Science and Engineering, Anurag University,*
Email- sowdariakshay@gmail.com,pranayayennam999@gmail.com, bhanupriyanka56@gmail.com, mounigoud338@gmail.com

Abstract - Sentiment analysis is defined as the process of predicting the underlying emotion of a sentence through Natural Language Processing (NLP). Itinvolves classification of a statement into three phases: "Positive","Neutral" or "Negative". Sentiment analysis is a mostly used method for finding and extracting the polarity of text sources. This paper focuses on examining the efficiency of supervised machine learning algorithm, Logistic Regression, for the classification of online reviews available on E-Commerce websites. The Main purpose of sentiment analysis is to automatethe process of collectingonline reviews of a product and analyzingthemto find the underlying sentiment about features. It involves computation of the sentiments of a large number of useful reviews andfiltering of irrelevant and unhelpful reviews. Also providingvisualization for our result summarization.

Keywords: Machine Learning, Polarity, Sentiment Analysis, Logistic Regression, Natural language Processing, Visualization.

I. Introduction

Everyday we come across different varieties of products in our lives and on the digital medium hundreds of product choices are available under a single category. It will be quite difficult for the customer to make a selection among them.To ease this problem we came up with a reviews system where customers who have already got that product leave a rating after using them and brief their experience by giving reviews. As we know ratings can be easily sorted and judged whether a product is good or bad, but when it comes to sentence reviews we need to go through every line to identify the underlying sentiment. In the era of artificial intelligence, things like that have gotten easier with the Natural Language Processing(NLP) technology. E-commerce websites such as Amazon,Flipkart,Ebay etc.,have a great influence in our everyday routine. In this modern era, these websites have made things sostraightforward that an individual can make purchases from anywhere around the globe. The best part of these websites isthat anyone can purchase things without stopping by at any store and it also providesusers with a wide range of options so users can select the finest among them.

Sentiment analysis attempts to determine which features of text are indicative of its context and build systems to take advantage of these features. The problem of classifying text as positive,neutral or negative is not the whole problem in and of itself, but it offers a simple enough premise to build upon further.

Amazon employs a 1-to-5 scale for all products, regardless of their category, and it becomes challenging to determine the advantages and disadvantages to different parts of a product.In this paper we attempted to categorize opinions about different parts of a product and present them independently to give readers more information that positive,neutral or negative sentiment.

II. Literature Survey

As a preliminary part of the work, an elaborative literature survey was carried out.

Callen Rain et Al. [1] proposed probabilistic machine learning approach to carry out the process of sentiment analysis on amazon reviews.Finding optimal parameter values for the two classifiers made it apparent how drastically they can change the accuracy of the algorithm. By having evidence that the number of features is so important, it also gives a sense of how important the composition of these vectors is.

Lijisha T et Al. [2] designed a system to deal with the basic issue of sentiment analysis which is sentiment polarity categorization. Online reviews of products from E-commerce website(flipkart) are selected for this model. These reviews enabled the system to analyze and generate the results which have been aligned based on different geographical areas. This system helped manufacturers to view the sentiments as a function of time and to estimate the improvisions required.

Lingling Zhao et Al. [3] explored the necessity of the unsupervised models and LSTM(Long short term memory) related sentiment analysis to identify stratified feature sentiment for requirement evolution prediction. The main goal is to provide the product developers to capture and become acquainted with the requirements change in an unobstructed way among different product feature requirement granularities. The outcome of his work predominantly subsidized the efforts of automated text mining analysis for product requirements engineering.

Abhilasha Singh Rathor et Al. [4] elaborated the fact that product reviews are not only useful for customers to purchase the things but also for the companies to continue their trade. He also highlighted that textual product reviews will be disoriented and these are needed to be classified as positive, neutral and negative. His research also showed that the supervised machine learning algorithm,SVM(Support Vector Machine) can be considered as an entry-level learning algorithm for Amazon reviews as it has highest accuracy.

Aashutosh Bhatt et Al. [5] implemented the process of sentiment analysis using sentiment polarity methodology, this method basically returns positive value if the phrase has a positive underlying sentiment and a negative value if the phrase reflects a negative sentiment.To begin with, the phrases are tested for indirect opinions, the test phrase is examined for certain predefined words that were found during manual analysis of reviews.If the phrase test fails, the review is tested in to find the word-‘not’, if it exists then everything after that word is investigated to find whether they are positive words or negative words, and consecutive words polarity are added and finally negated.If both phrase test and not test fails, then test phrase is broken down into words and polarity of each word is found from a dictionary of sentiment words and collective polarity is calculated.

Najma Sultana et Al. [6] noticed that the blend of verb,adjective and adverb words emerged as the best union among different compositions of the parts of speech, this newly proposed technique is tested on benchmark Stanford Dataset using six well-known supervised classifiers. As it is well known that a sentence is a composition of different combinations of the parts of speech, these different associations produce different accuracy rates. The precision of these parts of speech is elaborated with respect to the wide variety of classifiers along with the accuracy rates of different POS also the execution time of training and testing dataset has been showcased.

Wanliang Tan et Al.[7] tried two different varieties of features and for these two varieties of features, they tried all the algorithms mentioned in the model part. From the experimental outcome, it is noticed that accuracy on the test set is optimum when LSTM(Long short-term memory) on the first type of feature is used. Data imbalance has affected the results in a greater way and it is one of the main reasons that accuracy is not upto the mark.It is ideal to obtain more data points from other sources before attempting the process of resampling and different weighting techniques with existing data.

Elshrif Ibrahim Elmurngi et Al. [8-14] has proposed Logistic Regression,Support Vector Machine and Naives Bayes algorithms to analyze Amazon reviews datasets and also elaborated sentiment classification methods.They carried out experiments using three divergent datasets of Amazon reviews by removing stopwords and the subsequent experimental results have shown the accuracy, precision and recall of sentiment classification algorithms. In addition to that, they are also able to identify unfair positive reviews, unfair negative reviews and unfair neutral reviews using the detection processes of this method[16-20]. The best part of this study is, it helped to find the best supervised learning algorithm with respect to its accuracy by comparing different sentiment classification algorithms which are used to classify Amazon reviews datasets into the categories of fair and unfair reviews.

III. System Design

4.1 Architecture

Following diagram depicts overall architecture of the system which includes several phases

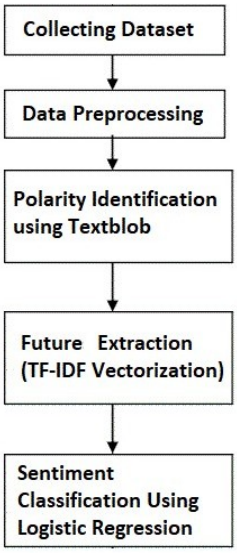

Figure 1. Architecture of Sentiment Analysis of Product Reviews

Collecting Dataset

We have collected Amazon musical instrument product reviews which consists of a total of 10261 data points and 9 features.

Data Preprocessing

When preprocessing, we have to perform the following:

1. Eliminate handles and URLs.
2. Tokenize the string into words.
3. Remove stop words.
4. Convert every word to its stem.
5. Convert all words to lowercase.

Polarity Identification using TextBlob

TextBlob is an actively used Natural Language ToolKit (NLTK) that supports complex analysis and operations on textual data. It returns Polarity and Subjectivity sentences.

Future Extraction (TF-IDF Vectorization)

For the purpose of transforming text into a meaningful representation of numbers, we make use of TF-IDF (Term Frequency - Inverse Document Frequency) Vectorizer.

It transforms text to feature vectors that can be fed to the estimator.

*Formula :*tf-idf(t, d) = tf(t, d) * idf(t)

Where, tf(t,d) = no.ofoccurences of term(t) in documents(d)

idf(t) = no.of documents containing term(t)

n is the total number of documents

Sentiment Classification using Logistic Regression

Logistic regression is a classification algorithm that makes use of a weighted combination of the input features and passes them through a sigmoid function. In this research, we use One-vs-all strategy for multiclass classification that involves training N distinct binary classifiers, each designed to recognize a specific class.In this we consider one class as 1 and rest all as 0, we train the model and get the requisite weights. We store the value of weights in a dictionary format for each classifier. Then by the help of Sigmoid Function we calculate the probability.

IV. Results

The entire dataset of 10,227 reviews was divided into three target labels of positive reviews (8991 data points) , Neutral (771 data points) and Negative (465 data points). After resampling the data we got each of 8991 data points of Positive, Neutral and Negative target labels. The total dataset of 26973, is distributed into a training set of size 20229 (75%) and a test set of size 6744 (25%).After applying the Classification algorithm the results are expressed using a Confusion matrix.

Confusion Matrix

Confusion matrix is an array of values representing the true positives,false positives, true negatives and false negatives which can be used to estimate the accuracy of a classification model. It is simple to understand, but the related terminology can be confusing.

array([[2267, 0, 2],
[51, 2192, 10],
[120, 208, 1894]], dtype=int64)

Ture Positive (TP) = 2267
False Negative (FN) = 2
False Positive (FP) = 171
False Negative (FN) = 4304
Accuracy = (TP+TN) /(TP+TN+FP+FN)= 0.94

Classification Report

Classification report is helpful to compute the quality of predictions of a classifier.It is widely used to display the models' precision values, recall values, f1 score and support. It lays out an easy path to understand the overall performance capabilities of trained models.

Table 1. Classification Report

	precision	recall	f1-score	support
0	0.93	1.00	0.96	2269
1	0.91	0.97	0.94	2253
2	0.99	0.85	0.92	2222
accuracy	-	-	0.94	6744
macro avg	0.95	0.94	0.94	6744
weighted avg	0.95	0.94	0.94	6744

Table 1 shows the vital classification metrics such as precision values, recall values,f1-score and support. The metrics are calculated using the values of true positives, false positives, true negatives and false negatives.

ROC Curve

The ROC (receiver operating characteristic)curve decides on which threshold to set up based upon the objective criteria. Here we plotted ROC for different classes which can help us understand which class was classified better. Also we plot micro and macro averages on the roc curve.

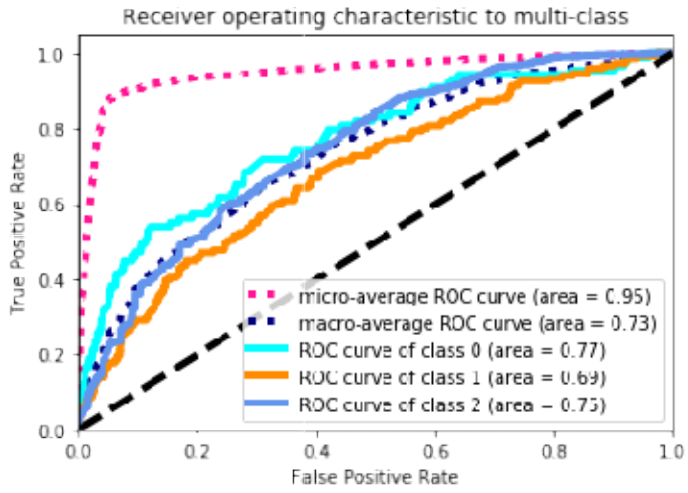

Figure 2. ROC Curve

V. Conclusion

In this study, we tried a traditional machine learning algorithm (Logistic Regression) to analyze customer feedback and classify all the classes by splitting the sentiments. By making the best of machine learning models and Natural Language Processing approaches, organizations can now interconnect with their customers both rationally and emotionallyand provide tailored assistance.

In summary, a brief classification of all the classes have been done by starting with text cleaning, customizing stop words, handling imbalance datawith smote, word vectorization and building the model.

Here are few insights from the study:

- We considered ngram in sentiment analysis as one word can't give proper results and stop words have to be manually checked as they have negative words. It is advised to avoid using stop words in sentiment analysis.
- Most of our neutral reviews were actual criticismof the product from the buyers, so amazon can consider these as feedback and give them to the seller to help them improve their products.
- Balancing the dataset gave us a better accuracy score. Without balancing, we got good precision but very bad recall and in-turn it affectedf1 score. So balancing the target feature is important.

REFERENCES

1. Callen Rain, “Sentiment Analysis in Amazon Reviews Using Probabilistic Machine Learning” Swarthmore College, Department of Computer Science.
2. LijishaT ,ArchanaBalakrishnan , Sumayya K S , Arjun P K , Renjith Sunny, “An Approach towards Feature Specific Opinion Mining and Sentimental Analysis Across E-Commerce Websites”.
3. LinglingZhao andAnping Zhao, “Sentiment Analysis Based Requirement Evolution Prediction”.
4. Abhilasha Singh Rathora ,AmitAgarwalb , PreetiDimric, “Comparative Study of Machine Learning Approaches for Amazon Reviews”.
5. Aashutosh Bhatt, Ankit Patel, Harsh Chheda, KiranGawande, “Amazon Review Classification and Sentiment Analysis”.
6. Ayaluri MR, K. SR, Konda SR, Chidirala SR. 2021. Efficient steganalysis using convolutional auto encoder network to ensure original image quality. PeerJ Computer Science 7:e356 https://doi.org/10.7717/peerj-cs.356
7. A Mallikarjuna Reddy, VakulabharanamVenkata Krishna, LingamguntaSumalatha and AvukuObulesh, “Age Classification Using Motif and Statistical Features Derived On Gradient Facial Images”, Recent

Advances in Computer Science and Communications (2020) 13: 965. https://doi.org/10.2174/2213275912666190417151247.
8. A. Mallikarjuna Reddy, V. Venkata Krishna, L. Sumalatha," Face recognition based on stable uniform patterns" International Journal of Engineering & Technology, Vol.7 ,No.(2),pp.626-634, 2018,doi: 10.14419/ijet.v7i2.9922
9. Najma Sultana, PintuKumar , Monika Rani Patra , Sourabh Chandra, S.K. SafikulAlam, "Sentiment Analysis for Product Review".
10. Wanliang Tan, Xinyu Wang, XinyuXu, "Sentiment Analysis for Amazon Reviews".
11. Elshrif Ibrahim Elmurngi and AbdelouahedGherbi, "Unfair Reviews Detection on Amazon Reviews using Sentiment Analysis with Supervised Learning Techniques".
12. Sangani, Chirag, and SundaramAnanthanarayanan. "Sentiment analysis of app store reviews."
13. https://www.analyticsvidhya.com/blog/2018/04/a-comprehensive-guide-to-understand-and-implement-text-classification-in-python/
14. K.Yessenov and S. Misailovic, "Sentiment analysis of movie review comments".
15. https://towardsdatascience.com/document-feature-extraction-and-classification-53f0e813d2d3
16. M. S. Elli and Y.-F. Wang. "Amazon reviews, business analytics with sentiment analysis".
17. https://machinelearningmastery.com/random-oversampling-and-undersampling-for-imbalanced-classification/ Sentiment analysis using product review by Xing Fang and Justin Zhan.
18. R. Collobert, J. Weston, L. Bottou, M. Karlen, K. Kavukcuoglu, and P. Kuksa "Natural language processing from scratch".
19. Swarajya Lakshmi V Papineni, SnigdhaYarlagadda, HaritaAkkineni, A. Mallikarjuna Reddy. Big Data Analytics Applying the Fusion Approach of Multicriteria Decision Making with Deep Learning Algorithms International Journal of Engineering Trends and Technology, 69(1), 24-28, doi: 10.14445/22315381/IJETT-V69I1P204.
20. Srinivasa Reddy, K., Suneela, B., Inthiyaz, S.,Kumar, G.N.S., Mallikarjuna Reddy, A." Texture filtration module under stabilization via random forest optimization methodology "International Journal of Advanced Trends in Computer Science and Engineering, Volume 8, No.3, May - June 2019
21. K. Dave, S. Lawrence, and D. M. Pennock, "Opinion extraction and semantic classification of product reviews".
22. M. S. Elli and Y.-F. Wang, "Amazon reviews, business analytics with sentiment analysis".
23. S. Hota and S. Pathak, "Knn classifier based approach for multi-class sentiment analysis of twitter data".
24. B. Liu and L. Zhang, "A Survey of Opinion Mining and Sentiment Analysis".
25. R. Socher, A. Perelygin, J. Wu, J. Chuang, C. D. Manning, A. Ng, and C. Potts, "Recursive deep models for semantic compositionality over a sentiment treebank".