Analysis of Product Reviews Using Sentimental Analysis

[1]Gollapalli Vinay Jyothi, [2]Katragadda Meghana, [3]Vemula Sai Samyukta, [4]Mrs.J.Janisha

[1,2,3]Students, [4]Professor
Bharat institute of higher education and research

Abstract: We have considered a data set from Kaggle Amazon Reviews Dataset. This data set contains data about the customer reviews of the amazon products. The data set is a .csv file containing reviewtext, review ratings, of the customers along with their review id, review province etc, and their coordinates. We have done all the preprocessing required on the data set, added a column sentiment which decides the positive or negative review to the dataset based on review rating and also we used nltk packages to text cleaning and removing unwanted words for deciding the reviews(stopwords). After the preprocessing phase we used bag of words strategy to convert the text content in reviews to numerical feature vectors, for that we used SciKit-Learn's CountVectorizer and also we use TfidfTransformer to overshadow the words which has lower average counts with same frequencies ,which has very little meaning We split the train and test data and train using machine learning models after defining the feature vector values with CountVectorizer and TfidfTransformer. To determine the model performance, use Naive Bayes analysis, Support Vector Machines, and Logistic Regression, and print the confusion matrix and accuracy

INTRODUCTION

Everyone can freely express his/her views and opinions anonymously and without the fear of consequences. Social media and online posting have made it even easier to post confidently and openly. These opinions have both pros and cons while providing the right feedback to reach the right person which can help fix the issue and sometimes a con when these get manipulated These opinions are regarded as valuable. This allows people with malicious intentions to easily make the system to give people the impression of genuineness and post opinions to promote their own product or to discredit the competitor products and services, without revealing identity of themselves or the organization they work for. Such people are called opinion spammers and these activities can be termed as opinion spamming. There are few different types of opinion spamming. One type is giving positive opinions to some products with intention to promote giving untrue or negative reviews to products to damage their reputation. Second type consists of advertisements with no opinions on product.

There is lot of research work done in field of sentiment analysis and created models while using different sentiment analysis on data from various sources, but the primary focus is on the algorithms and not on actual fake review detection. One of many other research works by E. I. Elmurngi and A. Gherbi [1] used machine learning algorithms to classify the product reviews on Amazon.com dataset [2] including customer usage of the product and buying experiences. The use of Opinion Mining, a type of language processing to track the emotion and thought process of the people or users about a product which can in turn help research work.

Opinion mining, which is also called sentiment analysis, involves building a system to collect and examine opinions about the product made in social media posts, comments, online product and service reviews or even tweets. Automated opinion mining uses machine learning, a component of artificial intelligence. An opinion mining system can be built using a software that can extract knowledge from dataset and incorporate some other data to improve its performance. One of the biggest applications of opinion mining is in the online and e-commerce reviews of consumer products, feedback and services. As these opinions are so helpful for both the user as well as the seller the e-commerce web sites suggest their customers to leave a feedback and review about their product or service they purchased. These reviews provide valuable information that is used by potential customers to know the opinions of previous or current users before they decide to purchase that product from that seller. Similarly, the seller or service providers use this information to identify any defects or problems users face with their products and to understand the competitive information to know the difference about their similar competitors' products.

There is a lot of scope of using opinion mining and many applications for different usages: Individual consumers: A consumer can also compare the summaries with competing products before taking a decision without missing out on any other better products available in the market. Businesses/Sellers: Opinion mining helps the sellers to reach their audience and understand their perception about the product as well as the competitors. Such reviews 3 also help the sellers to understand the issues or defects so that they can improve later versions of their product. In today's generation this way of encouraging the consumers to write a review about a product has become a good strategy for marketing their product through real audience's voice. Such precious information has been spammed and manipulated. Out of many researches one fascinating research was done to identify the deceptive opinion spam.

Nowadays, when customers want to draw a decision about services or products, reviews become the main source of their information. For example, when customers take the initiation to book a hotel, they read the reviews on the opinions of other customers on the hotel services. Depending on the feedback of the reviews, they decide to book room or not. If they came to a positive feedback from the reviews, they probably proceed to book the room. Thus, historical reviews became very credible sources of information to most people in several online services. Since, reviews are considered forms of sharing authentic feedback about positive or negative services, any attempt to manipulate those reviews by writing misleading or inauthentic content is considered as deceptive action and such reviews are labeled as fake [1]. Such case leads us to think what if not all the written reviews are honest or credible.

What if some of these reviews are fake. Thus, detecting fake review has become and still in the state of active and required research area. Machine learning techniques can provide a big contribution to detect fake reviews of web contents. Generally, web mining techniques [3] find and extract useful information using several machine learning algorithms. One of the web mining tasks is content mining. A traditional example of content mining is opinion mining [4] which is concerned of finding the sentiment of text (positive or negative) by machine learning where a classifier is trained to analyze the features of the reviews together with the sentiments. Usually, fake reviews detection depends not only on the category of reviews but also on certain features that are not directly connected to the content. Building features of reviews normally involves text and natural language processing NLP. However, fake reviews may require building other features linked to the reviewer himself like for example review time/date or his writing styles. Thus the successful fake reviews detection lies on the construction of meaningful features extraction of the reviewers

LITERATURE REVIEWS

[1] New mobile concepts can transform the world. Thousands of applications for employment, commerce, entertainment, etc. are produced and disseminated online nowadays. Most app shops may have trouble suggesting apps to users. So, customers require app recommendations based on their tastes and other constraints. We created a mobile app recommendation system that uses ratings, size, and permission as factors. Apkpure.com is a popular Android app store that uses Web Crawler to gather website information and validate linkages. Then, the Clustering Algorithm groups programmes by popularity, permission, and security. This article provides a basic recommendation system without sacrificing rating, size, or permission. The advantages are Clustering Algorithm make the grouping faster .the Disadvantages are it is Depended on third parties vendors for main stream data

[2] Google app store collects user ratings and reviews. App reviews reflect users' contentment. This assists other users before downloading or buying programmes. The reviews' exponential increase makes manual extraction impossible. Sentiment analysis using NLP and machine learning uncovers and interprets emotions. This research classifies app reviews by emotion and examines university students' app market activity. We examined machine learning techniques employing ensemble learning and TF-IDF text representation. Our model was trained and evaluated on Google and student reviews. Trigram+TF-IDF yielded the highest accuracy (93.37%) and Fscore (0.88). LR and NB improved by 87.80% and 85.5% with bagging. The advantages are It has a highest accuracy of 93.37% and Fscore of 0.88. the Disadvantages are manual extraction impossible so over depended on the NLP

[3] Developers utilise user ratings and reviews to correct bugs and plan. This helps app sellers collect data. Redundancy and data volume are obstacles for machine learning. This study examines Shopify app reviews. We split user ratings into satisfied and unhappy groups and sanitise the data to avoid limitations. Later, bag-of-words, TF-IDF, and Chi2 store pertinent data. Random forest, AdaBoost, and logistic regression classify reviews as happy or unhappy. Accuracy, precision, recall, and f 1 score were evaluated. In this study, logistic regression combined with TF-IDF and Chi2 had an 83% true acceptance rate. The advantages are logistic regression combined with TF-IDF and Chi2 had an 83% true acceptance rate. The disadvantages are data Retrieval taken high storage and time

[4] Google Play contains almost every software and service app. App stores let people download and rate apps. User reviews may include issues, feature ideas, or word ratings. App review categorization is time-consuming. Classifying apps automatically may accelerate problem solutions. Proposed automatic categorization review. Non-textual app reviews aren't used. Deep-learning app categorization. Deep learning and app reviews classify text. Textual and non-textual information from each app review is preprocessed, app review sentiment is computed using Senti4SD, and the reviewer's history, including total reviews and submission rate, is calculated (i.e., what percentages of his review have been submitted for the associated app). Digitize app reviews. AI sorts app reviews. The proposed strategy advances the field, according to a public dataset evaluation. It raises accuracy to 95.49, recall to 93.94, and f-measure to 94.71. The advantages are It has more stability as Deep learning is used. the Disadvantages are Its accuracy is questioned

[5] Thousands review apps. Quickly responding to feedback boosts app ratings, popularity, and success. Many reviews can't be manually answered. Seq2seq training with review-response pairs automates response creation. Because training review-response pairings originate from several applications, such models struggle to create app-specific solutions. Various apps have different needs. Individual applications have restricted review-response pairings that lack critical information to respond to a new review. AARSYNTH synthesises app responses. AARSYNTH appseq2seq. AARSYNTH earns top-K app reviews and a user-review snippet. The data and user evaluation are merged into a machine reading comprehension model. Ratings and descriptions are clarified. Fusion produces app-specific code. AARSYNTH received Google Play reviews. AARSYNTH wins by 22.2%. Human testing shows AARSYNTH improves reaction quality above state-of-the-art technologies. The advantages are Evaluation are merged into a machine reading comprehension so data is classified. the Disadvantages are Improvement in user reaction can be done

[6] Apps can handle more user demands, but their access to sensitive data creates privacy issues. Existing strategies suggest automatically identifying explanatory lines from app descriptions to alert users of sensitive behaviours. However, many sensitive behaviours aren't addressed in app descriptions. General approaches transform code to sentences. These strategies lack the language to describe sensitive data usage and fail to address the context (app functions) of sensitive actions. We offer Describectx, a context-aware description synthesis strategy that trains a neural machine translation model using popular applications and creates app-specific descriptions for sensitive actions. Describectx encodes privacy policy vocabularies, call graph behaviour summaries, and GUI text context. Describectx creates better accurate descriptions (24.96 in BLEU) and higher user ratings compared to manually recognised reference sentences in 1,262 Android applications. The advantages areDescribectx creates better accurate descriptions. the Disadvantages are Precise column of information is only used but there are more available

[7] App store mining helps companies maintain and improve current applications, making it a viable demand-collection strategy. Despite advances in requirements mining, nothing is known about how to extract new application needs from current (similar) solutions. In the proposed project, we leverage app store data to discover app properties. These qualities and others (e.g. ratings) may assist developers identify app features. We asked practitioners about this method. This report presents first research findings

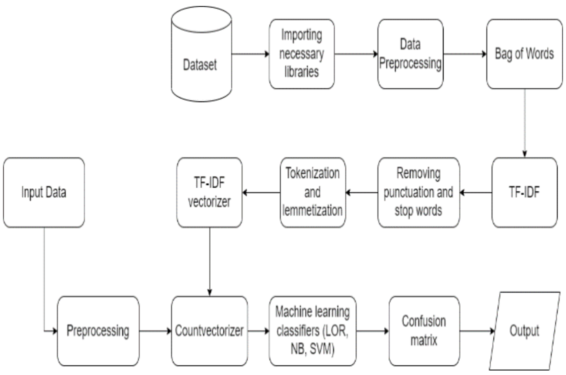
within a bigger strategy. Our survey shows practitioners desire our method. Conceptually, our method is contested. This study discusses how machine learning may be used to extract automatic needs from crowd-sourced data. The advantages are Information can be extracted automatically according to the needs from crowd-sourced data. the Disadvantages are Representation of unistructural data representation is not mention

[8] Google Play and Apple AppStore both have over 3 million apps. Billions use and rate these apps. User reviews give app vendors and developers with useful information about difficulties, new feature ideas, and implemented functions, according to recent research. This survey categorises app assessments as bug reports, feature requests, user experiences, and ratings. We use star rating, tense, text classification, NLP, and sentiment analysis. We compared accuracy using string matching. Metadata lowers classification accuracy. NLP achieved 70-95% classification accuracy and 80-90% recall. Multiclass classifiers outperformed binary classifiers. Our results influence the design of review analytics systems that help app providers, developers, and users filter and distribute critical evaluations. The advantages areNLP achieved 70-95% classification accuracy and 80-90% recall the Disadvantages are Filtering of data structured under stars and review can be improved

[9] Mobile apps and distribution channels have grown in recent years. Users may rate and review applications in app stores to provide creators feedback. User evaluations may enhance software quality, fix issues, and add functionality. Due to the expanding number of textual reviews, this data is difficult to manage. This study suggests using cognitive computing to solve this difficulty by constructing a smart agent to mine bug reports, feature proposals, and mobile app evaluations. This paper's key contributions include the creation of a cognitive agent to help developers manage user interactions, the use of machine learning methods to identify bugs and feature requests, and the agent's deployment in a real situation. The advantages are Functionality is provided with Machine Learning. the Disadvantages are Improvement in management textual review

[10] We give a hybrid (qualitative and quantitative) technique to estimate and predict QoE of WiFi or cellular mobile applications. 33 living lab participants evaluated their mobile applications' QoE for four weeks, yielding 5663 ratings. mQoL-Log recorded network, user, battery, and other data. Chrome, Maps, Spotify, Instagram, Facebook, Messenger, and WhatsApp were examined. We used machine learning (Extreme Gradient Boosting) to estimate application QoE. Our model accurately anticipated QoE. Walking and app objective were most predictive after session duration, battery level, and network QoS. Longer app sessions reduce QoE. The advantages areHighly complex data such as battery usage performance are handled easily. the Disadvantages are Session management is lacking when user try to filter out categories

System Architecture



EXISTING SYSTEM

The existing systems are based on RNN, recurrent neural networks. This approach disregards sentence positional information and reduces the complexity involved in building sentence sequences. The existing systems lack efficiency and training in the model. They have low accuracies. The existing systems are known to depict long-term dependency problems. In the existing systems, a major problem is the model's inclination towards positive reviews. the length of a text review tends to be larger when the person is writing a chunk criticizing the product. this can be clearly seen in the EDA, and therefore the dataset itself tends to carry some bias towards positive reviews.

PROPOSED SYSTEM

Previous iterations of our work contain testing multiple methods of data cleaning and preparation so as to get the cleanest possible text data. In addition to the fundamental pre-processing and data cleaning we had done to the text data, lemmatizing, tokenization, and removal of unclean data(tags, punctuation, stop words, etc.) were performed. This alone ended up increasing the accuracy by over 18 points. EDA remains identical for the bulk. The main changes between previous and current iterations of our work are changes to the hyperparameters of all models tested to enhance accuracy even by touch, the inclusion of cross-validation testing for all models, and an increase in dataset size. Comparing our results (accuracy and the precision) with the results of the works cited in the section related works and that use the same machine learning algorithms, demonstrate that the best accuracy and precision is given by our results.

MODULES

MODULE 1 : SENTIMENT CLASSIFICATION

Sentiment categorization is an important module in sentiment analysis of product reviews. The major goal of this module is to categorise the emotion of a product evaluation as good, negative, or neutral.

This lesson employs a number of methodologies, such as machine learning algorithms and lexicon-based approaches. Machine learning approaches entail training a model using a labelled dataset of product evaluations and their associated feelings, allowing

the model to reliably predict the sentiment of fresh, unlabeled reviews. Lexicon-based techniques, on the other hand, rely on pre-defined lists of positive and negative terms to assess a review's emotion.
The content of the reviews is examined in both techniques, and the frequency of positive and negative terms is determined to get the overall mood. Natural language processing methods and sentiment analysis algorithms are required for this module.
This module generates a sentiment score for each product review, indicating whether the review is good, negative, or neutral. The sentiment scores may be used to acquire insights into client feedback and sentiment about a product, allowing companies to make data-driven choices to improve their goods and increase customer happiness.

MODULE 2 : FEATURE EXTRACTION

Feature extraction is a module used in sentiment analysis of product evaluations. Its major goal is to find the important qualities or elements of a product mentioned by consumers in their evaluations. These qualities are then utilised to give insights into the product's strengths and limitations, as well as opportunities for development.
Several text mining approaches, such as frequency analysis and theme modelling, are used in this module to extract the most significant aspects from product evaluations. The frequency of certain words or phrases used in the reviews is counted in frequency analysis, while topic modelling includes finding clusters of similar words and phrases that indicate a single theme.
After the extraction of the attributes, they are classified and examined to identify their influence on consumer sentiment. For example, if many consumers express satisfaction with the product's customer service, this might be a strong selling feature for the product. Yet, if consumers repeatedly complain about the product's durability, this might be an area for improvement.
The programme requires understanding of text mining and data analysis methodologies, as well as product and industry knowledge. This module produces a list of the most essential features mentioned in product reviews, together with the emotion ratings associated with them.
Overall, feature extraction offers organisations with useful insights into user feedback and attitude about a product, allowing them to discover areas for development and make data-driven choices to boost customer happiness.

MODULE 3 : VISUALIZATION AND REPORTING

Visualization and reporting is a module used in the sentiment analysis of product evaluations. Its major goal is to show the sentiment analysis and feature extraction findings in a visually attractive and easy-to-understand way.
This subject entails using data visualisation tools and methods to generate charts, graphs, and dashboards that depict the sentiment scores and significant attributes collected from product evaluations. These visualisations may be used to analyse trends and patterns in consumer feedback, as well as providing insights into the product's strengths and faults.
The module also includes the creation of reports that summarise the sentiment analysis and feature extraction results. These reports give an in-depth look into consumer sentiment about the product, including the most often cited features, positive and negative sentiment, and ideas for improvement.
Expertise in data visualisation tools and methodologies, as well as report writing and presentation abilities, are required for this session. This module produces a visually attractive and simple representation of the sentiment analysis and feature extraction findings, as well as a thorough report with insights and suggestions for product enhancement.
Overall, visualisation and reporting provide organisations a clear and succinct picture of consumer sentiment towards their goods, allowing them to make data-driven choices to increase customer happiness and product quality.

Conclusion:

The findings of this study back with our prior data exploration findings, which found that the data is heavily skewed toward positive evaluations, as seen by the lower support numbers in the categorization report. Furthermore, both positive and negative evaluations have a high standard deviation with low frequencies, which we do not deem relevant, as seen by the lower accuracy, recall, and F1 scores in the categorization report. Analysis of product review using sentiment analysis 28 Despite the fact that the Neutral and Negative findings are not particularly good predictors in this data set, the sentiment analysis is predicted with a 92.23 percent accuracy rate. As a result, we feel at ease with the skewed data set. Additionally, when we continue to input additional datasets that are more balanced in the future, this model will re-adjust to a more balanced state. Note : That the first row will be disregarded because we replaced all NAN with " " earlier. When we originally imported the raw data, we tried to eliminate this row, however Pandas Dataframe didn't like it when we tried to drop all NANs (before stratifying and splitting the dataset). As a consequence, the best remedy was to replace the NAN with " ", and the first row would be omitted in this analysis. Finally, the overall result indicates that the products in this dataset are typically well-received. From the analysis above in the classification report, we can see that products with lower reviews are not significant enough to predict these lower rated products are inferior. On the other hand, products that are highly rated are considered superior products, which also perform well and should continue to sell at a high level. As a result, we need to input more data in order to consider the significance of lower rated products, in order to determine which products should be dropped from Amazon's product roster. The good news is that despite the skewed dataset, we were still able to build a robust Sentiment Analysis machine learning system to determine if the reviews are positive or negative. This is possible as the machine learning system was able to learn from all the positive, neutral and negative reviews, and fine tune the algorithm in order to avoid biased sentiments. In future we will try to include more robust algorithms and also attempt to integrate it into a web app so that it can become user friendly. In conclusion, we were able to correctly link positive, neutral, and negative feelings for each product in Amazon's Catalog, despite the fact that we needed additional data to balance out the lower rated goods in order to weight their relevance.

REFERENCES

1. Pang, B., & Lee, L. (2008). Opinion mining and sentiment analysis. Foundations and Trends® in Information Retrieval, 2(1-2), 1-135.
2. Liu, B. (2012). Sentiment analysis and opinion mining. Synthesis Lectures on Human Language Technologies, 5(1), 1-167.

3. Haddi, E., Liu, X. Z., & Shi, Y. (2013). The role of text pre-processing in sentiment analysis. Procedia Computer Science, 17, 26-32.
4. Cambria, E., & Hussain, A. (2012). Sentiment analysis: Concepts, techniques and applications. IEEE Intelligent Systems, 27(6), 78-84.
5. Kim, S. M., & Hovy, E. (2004). Determining the sentiment of opinions. In Proceedings of the 20th international conference on Computational Linguistics (pp. 1367-1373).
6. Hu, M., & Liu, B. (2004). Mining and summarizing customer reviews. In Proceedings of the 10th ACM SIGKDD international conference on Knowledge discovery and data mining (pp. 168-177).
7. Jindal, N., & Liu, B. (2006). Opinion spam and analysis. In Proceedings of the International Conference on Web Search and Web Data Mining (pp. 219-230).
8. Liu, B. (2015). Sentiment analysis: A multidisciplinary research area. IEEE Intelligent Systems, 30(3), 3-4.
9. Nasukawa, T., & Yi, J. (2003). Sentiment analysis: Capturing favorability using natural language processing. In Proceedings of the 2nd International Conference on Knowledge Capture (pp. 70-77).
10. Vlachos, A., & Meek, C. (2008). Evaluating the effectiveness of features for unsupervised lexical sentiment classification. In Proceedings of the 22nd International Conference on Computational Linguistics-Volume 1 (pp. 827-834).