



e-ISSN: 2319-8753 | p-ISSN: 2347-6710

IJIRSET

International Journal of Innovative Research in
SCIENCE | ENGINEERING | TECHNOLOGY



INTERNATIONAL JOURNAL OF INNOVATIVE RESEARCH

IN SCIENCE | ENGINEERING | TECHNOLOGY

Volume 13, Issue 4, April 2024

ISSN INTERNATIONAL
STANDARD
SERIAL
NUMBER
INDIA

Impact Factor: 8.423

 9940 572 462

 6381 907 438

 ijirset@gmail.com

 www.ijirset.com

Product Review Sentiment Analysis Using NLP

Dr.R.Nagarajan, A.Ramya

Assistant Professor, PG & Research Department of Computer Science, Sri Ramakrishna College of Arts and Science,
Coimbatore, India

UG Student, PG & Research Department of Computer Science, Sri Ramakrishna College of Arts and Science,
Coimbatore, India

ABSTRACT: This project investigates various learning algorithms through sentiment analysis of product reviews. By knowing the reviews of those products the customer gets to understand the tactics and can purchase easily. The increasing significance of sentiment analysis parallels the rise of social media platforms, critiques, forum discussions, blogs, micro-blogs, Twitter, and social networks. The realm of sentiment analysis is closely intertwined with natural language processing and text mining. Also known as opinion mining, sentiment analysis involves examining individuals' opinions to determine whether they are "happy," "sad," "angry," and so forth.

KEYWORDS: Product Review, Sentiment Analysis, NLP.

I. INTRODUCTION

Product reviews analyze the pros and cons of a specific product or service. Nowadays, these reviews are predominantly found online but may also be present in print media. Consumers rely on product reviews to aid them in making informed purchasing decisions.

Natural language processing (NLP) explores the interaction between humans and computers to enable computers to process written and spoken language in a manner akin to humans. This interdisciplinary field merges computer science, linguistics, and machine learning. At the core of NLP is the challenge of enabling computers to comprehend natural language. Unlike structured data such as spreadsheets and database tables, human languages, texts, and speech constitute unstructured data, posing significant challenges for computers to interpret. Hence, the necessity for Natural Language Processing arises.

II. LITERATURE WORK

Endsuy utilized Twitter datasets to perform exploratory data analysis on the 2020 US Presidential Election. They conducted a comparative analysis between the sentiment expressed in location-based tweets and the sentiment of on-the-ground public opinion. The study incorporated features such as latitude, longitude, city, country, continent, and state code, obtained using the OpenCage API and sciSpacy Named Entity Recognition (NER). Two datasets from Kaggle, concerning Donald Trump and Joe Biden, both dated November 18, 2020, were employed. For lexicon-based feature extraction, the study employed a valence aware dictionary for sentiment reasoning (VADAR), while logistic regression machine learning approaches were utilized for classification purposes.

Bibietal developed a Cooperative Binary-Clustering Framework for sentiment analysis on Indigenous datasets from Twitter. They employed majority voting to partition the data and integrated single linkage, complete linkage, and average linkage approaches. Utilizing the confusion matrix, they categorized the clusters into positive and negative sentiment categories. For feature selection, the study employed uni-gram, TF-IDF, and word polarity mechanisms. Results of the analysis indicate that the cooperative clustering approach outperforms individual partitioning techniques, achieving a 75% accuracy rate.

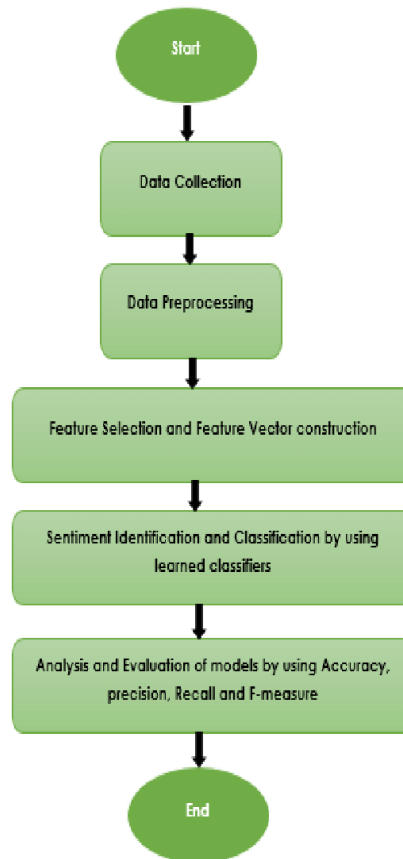
Cekik et al. (Year) utilized a filter-based feature selection method called Proportional Rough Feature Selector (PRFS) along with various classifiers such as SVM, DT, KNN, and Naive Bayes. PRFS, based on rough set theory, improved classifier performance significantly.



Peng et al. (Year) proposed a sentiment model adversarial learning method comprising a generator, discriminator, and sentiment classifier. They employed a multi-head attention mechanism in the generator to obtain efficient semantic and sentiment information. The discriminator measured the similarity of sentiment polarity in the generative vector and global vector produced by the generator and classifier, respectively, using gradients to update weights in the generator and enhance word embeddings.

Rodrigues et al. devised a pattern-based methodology for extracting aspects and analyzing sentiment. This approach leverages pattern analysis to extract explicit aspect syntactic patterns from product sentiments. It utilizes bigram features extraction and integrates Senti-Wordnet to ascertain the sentiment polarity of sentences. The study found that the multi-node clustering approach surpasses the single-node clustering approach in terms of performance.

III. METHODOLOGY



- Tokenization involves breaking down text into smaller units for analysis.
- Stop word removal entails eliminating common words from text to focus on unique words that provide essential information.
- Lemmatization and stemming involve reducing words to their root forms for processing.
- Part-of-speech tagging marks words based on their grammatical roles, such as nouns, verbs, and adjectives.
- After preprocessing the data, an algorithm is developed to analyze it. Two main types of natural language processing algorithms are commonly used:
- Rules-based systems rely on carefully crafted linguistic rules. This approach was prevalent in the early stages of natural language processing development and is still utilized today.
- Machine learning-based systems employ algorithms that learn patterns and relationships in data without explicit programming. These systems have gained prominence with advancements in machine learning techniques.



MACHINE LEARNING MODEL

Natural Language Processing (NLP) is a crucial component of artificial intelligence (AI), enabling computer programs to comprehend human language in its spoken and written forms. With roots in linguistics, NLP has been evolving for over 50 years and finds applications across various fields such as medical research, search engines, and business intelligence.

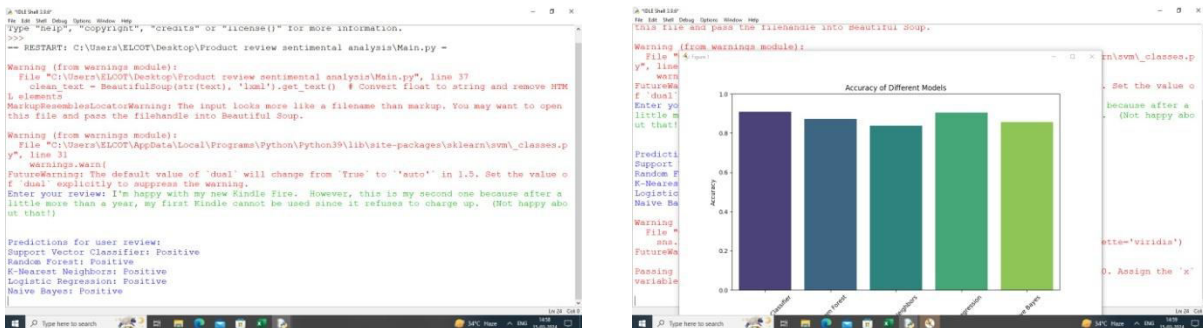
NLP empowers computers to grasp natural language akin to humans. Whether it's spoken or written, NLP utilizes artificial intelligence to receive real-world input, process it, and interpret it in a manner comprehensible to computers. Similar to how humans utilize different senses like hearing and sight, computers rely on programs to read and microphones to capture audio. Moreover, just as humans process input using their brains, computers employ programs for processing respective inputs. Eventually, during processing, the input is converted into code that the computer can comprehend.

IV. RESULT AND DISCUSSION

Modern approaches to natural language processing rely heavily on deep learning, a subtype of artificial intelligence that analyzes and leverages patterns in data to enhance a program's comprehension. Deep learning models demand vast quantities of labeled data for training the natural language processing algorithm to recognize pertinent correlations. However, amassing such extensive datasets poses a significant challenge in natural language processing.

In contrast, earlier methods of natural language processing predominantly embraced a rules-based paradigm. Simple machine learning algorithms were instructed to identify specific words and phrases within text and deliver predefined responses upon encountering these phrases. However, deep learning represents a more adaptable and intuitive approach. In this framework, algorithms learn to discern speakers' intentions through exposure to numerous examples, resembling the way a child learns human language.

SCREENSHOTS



V. CONCLUSION ANF FUTURE WORK

In conclusion, our analysis using MultinomialNB classifier and NLP techniques showed that sentiment analysis of product reviews can be a powerful tool for gaining insights into the opinions and sentiments. MultinomialNB, which is specifically designed for text classification tasks, outperformed GaussianNB and BernoulliNB in our analysis in terms of accuracy, precision, recall, and F1-score. This highlights the importance of choosing the right classifier for the specific task at hand. Furthermore, NLP techniques such as word clouds and sentiment analysis using VADER provided valuable insights into the distribution of positive and negative sentiments in the movie reviews. This information can be used to improve the quality of movie reviews, recommend movies based on sentiment, and predict box office success.

The future work could include the integration of data into given model and increasing its performance level.

REFERENCES

1. M. Govindarajan conducted a study titled "Sentiment Analysis of Movie Reviews using Hybrid Method of Naive Bayes and Genetic Algorithm" published in the International Journal of Advanced Computer Research (ISSN (print): 2249-7277 ISSN 2277-7970), Volume-3 Number-4 Issue-13 in December 2013.
2. Apoorv Agarwal, Boyi Xie, Ilia Vovsha, Owen Rambow, and Rebecca Passonneau contributed to a paper on "Sentiment Analysis of Twitter Data."



3. Apoorv Agarwal and Jasneet Singh Sabharwal presented a paper on "End-to-End Sentiment Analysis of Twitter Data" at the Proceedings of the Workshop on Information Extraction and Entity Analytics on Social Media Data during COLING 2012 in Mumbai, in December 2012.
4. V.K. Singh, R. Piryani, A. Uddin, and P. Waila authored a paper titled "Sentiment analysis of movie reviews: A new feature-based heuristic for aspect-level sentiment classification," presented as a conference paper in March 2013.
5. Alexander Pak and Patrick Paroubek conducted research on "Twitter as a Corpus for Sentiment Analysis and Opinion Mining."
6. Ajinkya Ingle, Anjali Kante, Shriya Samak, and Anita Kumari explored "Sentiment Analysis of Twitter Data Using Hadoop," published in the International Journal of Engineering Research and General Science Volume 3, Issue 6, November-December 2015 (ISSN 2091-2730).



INTERNATIONAL
STANDARD
SERIAL
NUMBER
INDIA



INTERNATIONAL JOURNAL OF INNOVATIVE RESEARCH

IN SCIENCE | ENGINEERING | TECHNOLOGY

 9940 572 462  6381 907 438  ijirset@gmail.com



www.ijirset.com

Scan to save the contact details