Industrial Engineering Journal
ISSN: 0970-2555
Volume: 53, Issue4, April : 2024

FEEDBACK ANALYSIS USING MACHINE LEARNING

Mr.R.Venkatesh[1],Sailada Achyutha[2],Routhu HariKrishna[3],Ponnamanda Hari Vardhana Shubhash[4],Rajana Bhanu Prakash[5]

[1]Assistant Professor ,Department of Computer Science & Engineering,Raghu Engineering College, Vishakhapatnam, Andhra Pradesh

[2,3,4,5] Student of B-TECH, RaghuEngineering College , Vishakhapatnam, Andhra Pradesh

Email:-20981a05f4@raghuenggcollege.in ,20981a05e6@raghuenggcollege.in ,20981a05d9@raghuenggcollege.in , 20981a05e3@raghuenggcollege.in.

ABSTRACT:

In the realm of project management and improvement, feedback analysis is pivotal. With the advent of machine learning, the ability to efficiently collect and analyze large volumes of feedback has significantly advanced. This abstract outlines a system that leverages machine learning models to perform sentiment analysis on project feedback.
The system is designed to process feedback data, which could be collected via forms or other means, and apply machine learning algorithms to classify the sentiments expressed. Techniques such as Naïve Bayes, Random Forest, or even deep learning methods like Convolutional Neural Networks (CNNs) and Long Short-Term Memory networks (LSTM) can be utilized for this purpose.
The goal is to categorize feedback into various sentiment classes, such as positive, neutral, or negative. This classification aids in understanding the general reception of the project and pinpointing areas that require enhancement. The analysis can also reveal patterns and trends in feedback that might not be immediately apparent through manual evaluation.
By implementing such a system, organizations can gain actionable insights from project feedback, leading to informed decision-making and continuous improvement of project outcomes.In this project loaded vectorizer has been used for data cleaning.

1.INTRODUCTION:

Data cleaning places a important role in this analysis. Loaded vectorizer is used for data cleaning of the reviews , means it will divide the positive and negative reviews by the customers and use them for further references of the development of the items in any organization

1.1IMPORTANCE OF DATA CLEANING:

Data cleaning is a crucial step in the sentiment analysis of reviews, as it directly impacts the accuracy and reliability of the results. Here’s why it’s so important:

Enhances Data Quality: Data cleaning involves filtering datasets to remove incorrect, duplicate, or missing values, which enhances the overall quality of the data.
Prevents Distorted Results: Unclean data can interfere with the analysis, leading to misleading or distorted results. Proper data cleaning ensures that the sentiment analysis reflects the true sentiment expressed in the reviews.

Improves Model Performance: A well-prepared dataset not only improves the performance of machine learning models but also ensures that the insights derived from sentiment analysis are reliable and can be used to make informed decisions.

Addresses Unstructured Data: Especially in social media reviews, like tweets, the unstructured nature of the data presents unique challenges. Data cleaning helps in structuring this data for effective sentiment analysis.

Data cleaning is an essential part of the sentiment analysis process, ensuring that the insights obtained are accurate and can be trusted for making business decisions, improving products, and understanding customer satisfaction.

1.2 SCOPE:

The scope of a feedback analysis project using machine learning (ML) is quite broad and can encompass various aspects:

- **Automated Sentiment Analysis**: ML can be used to automatically determine the sentiment of feedback, whether it's positive, negative, or neutral, which is essential for understanding customer satisfaction.
- **Trend Identification**: By analyzing feedback over time, ML can help identify trends and patterns, providing insights into how products or services are being received by customers.
- **Product Improvement**: Feedback analysis can guide product development by highlighting areas that need improvement or features that are particularly appreciated by users.
- **Customer Service**: ML can categorize feedback into different types, such as complaints or inquiries, which can then be routed to the appropriate department for action.
- **Business Strategy**: The insights gained from feedback analysis can inform business strategies, marketing campaigns, and other operational decisions.
- **Educational Systems**: In academic settings, feedback analysis can help improve teaching methods and educational content based on student responses.

Overall, the scope of such a project is to leverage ML to gain actionable insights from feedback data, which can drive improvements and strategic decisions across various domains.

1.3 MOTIVATION:

The motivation behind using machine learning (ML) for project feedback analysis is multifaceted:

- **Efficiency**: ML can process vast amounts of feedback data quickly and accurately, which would be time-consuming and laborious if done manually.
- **Predictive Insights**: ML models can predict future trends and behaviours by analysing past feedback, helping organizations to be proactive rather than reactive.
- **Personalization**: Feedback analysis using ML can lead to personalized responses and services, enhancing user experience and satisfaction.
- **Continuous Improvement**: By understanding feedback at a granular level, ML enables continuous improvement in products, services, and processes.
- **Objective Analysis**: ML provides an objective analysis of feedback, reducing the potential for human bias in interpreting sentiments.
- **Innovation**: The use of ML for feedback analysis can inspire new research and development in the field of psychology and education, particularly in understanding the impact of feedback on motivation.

These motivations drive the adoption of ML in feedback analysis to enhance decision-making, improve customer relations, and foster growth and innovation.

1.4 OVERVIEW:

Project Overview:

The project aims to develop a machine learning-based system for analysing feedback received on various projects. The core objective is to automate the process of interpreting large volumes of feedback data to extract meaningful insights that can drive improvements and inform decision-making.

Key Components:

- **Data Collection**: Gathering feedback from multiple sources such as surveys, online reviews, and social media comments.
- **Data Preprocessing**: Cleaning and preparing the data for analysis, which includes removing duplicates, correcting errors, and handling missing values.
- **Feature Extraction**: Identifying and extracting relevant features from the feedback that are indicative of sentiments, opinions, and topics.
- **Model Training**: Using algorithms like convolutional neural networks (CNNs), bidirectional LSTM, and attention mechanisms to train models that can accurately classify feedback sentiments[1]1.
- **Analysis and Interpretation**: Applying the trained models to new feedback data to categorize sentiments and extract trends, patterns, and actionable insights.
- **Visualization**: Presenting the analysis results in an understandable format through graphs, charts, and dashboards.

Expected Outcomes:

- **Sentiment Classification**: Categorizing feedback into positive, negative, or neutral sentiments.
- **Trend Detection**: Identifying trends over time to understand how the reception of projects changes.

- **Issue Identification**: Pinpointing specific issues or areas of concern highlighted in the feedback.
- **Performance Metrics**: Establishing metrics to measure the success of projects based on feedback analysis.

Applications:

- **Customer Service**: Improving customer support by understanding client feedback and addressing their concerns.
- **Product Development**: Guiding product improvements and feature enhancements through customer reviews analysis.
- **Business Strategy**: Informing strategic decisions and marketing initiatives based on customer sentiment trends.

2.LITERATURE SURVEY:

Conducting a literature survey for a feedback analysis project using machine learning (ML) involves reviewing various scholarly articles, journals, and conference papers to understand the current state of research in the field. Here's a synthesized overview based on recent literature:

Literature Survey Overview:

- **Machine Learning Algorithms**: A comprehensive study of ML algorithms is essential as they form the backbone of feedback analysis systems. This includes supervised, unsupervised, and reinforcement learning methods, each with its own set of applications and challenges.
- **Analysing Qualitative Data**: Research on using ML to analyse qualitative survey data, such as open-ended responses, is gaining traction. Techniques like Latent Dirichlet Allocation (LDA) and neural networks are employed for sentiment analysis and topic modelling.
- **Sentiment Analysis Methods**: The literature also covers various sentiment analysis methods and their applications, including the categorization of online reviews, predictive decision-making, and the detection of false reviews.
- **Education Domain**: Specific studies focus on sentiment analysis in the education domain, utilizing machine learning and deep learning methods to analyse student feedback and improve educational outcomes.
- **Decision Making in Healthcare**: Systematic reviews have been conducted on the use of ML methods in healthcare, particularly how they inform decision-making at the patient-provider level.

This survey indicates that the field is rich with diverse applications of ML for feedback analysis, spanning various domains from education to healthcare. The literature also suggests a growing interest in developing more sophisticated models to handle complex data and provide deeper insights into feedback.

3.IMPLEMENTATION STUDY:

An implementation study of a feedback analysis project using machine learning (ML) typically involves several key phases:

1. **Problem Definition**: Clearly defining what the feedback analysis aims to achieve, such as sentiment analysis, trend identification, or actionable insights extraction.
2. **Data Collection**: Gathering feedback data from various sources like surveys, online platforms, or customer service interactions.
3. **Data Preprocessing**: Cleaning and organizing the data to ensure it is suitable for analysis. This may include removing duplicates, handling missing values, and normalizing text.
4. **Feature Engineering**: Selecting or creating relevant features from the data that can be used to train the ML models. This could involve natural language processing techniques to extract sentiment-related attributes.
5. **Model Selection**: Choosing appropriate ML algorithms based on the nature of the data and the analysis goals. Common choices include Naïve Bayes, Support Vector Machines, or neural networks.
6. **Model Training and Testing**: Training the selected models on a portion of the data and testing their performance on another set to evaluate their accuracy and effectiveness.
7. **Deployment**: Integrating the trained models into a system where they can analyse new feedback data and provide ongoing insights.
8. **Monitoring and Maintenance**: Continuously monitoring the system's performance and making necessary adjustments or retraining the models with new data to maintain accuracy over time.

PROPOSED SYSTEM:

In the ever-evolving landscape of natural language processing and sentiment analysis, Mitra's 2020 paper titled "Sentiment Analysis Using Machine Learning Approaches (Lexicon based on movie review dataset)" published in the Journal of Ubiquitous Computing and Communication Technologies (UCCT) delves into the application of machine learning techniques for sentiment analysis.

Sentiment analysis, the process of discerning and quantifying emotional tone within text data, has gained prominence due to its wide-ranging applications, from market research to social media monitoring. Mitra's research focuses specifically on sentiment analysis using a lexicon-based approach, leveraging a dataset comprised of movie reviews.

The choice of the dataset is significant as movie reviews often contain rich and nuanced expressions of sentiment, making them a valuable resource for studying sentiment analysis methodologies. The paper is set against the backdrop of the increasing relevance of sentiment analysis in the digital age, where understanding public opinion and sentiment on various topics is crucial.

By harnessing machine learning techniques and a specialized lexicon, the study seeks to contribute to the advancement of sentiment analysis methodologies, shedding light on the practical applications of such techniques in the realm of movie reviews and potentially extending their utility to broader domains of text analysis. This paper's findings hold relevance for researchers, practitioners, and enthusiasts interested in the ever-expanding field of sentiment analysis and its real-world implications.

This paper not only presents an advanced technique but also provides insights into the potential applications and improvements in the field of medical image analysis. The research aims to contribute to the development of more accurate and efficient tools for the early diagnosis and quantification of retinal diseases, thereby potentially enhancing the quality of patient care and reducing the burden on healthcare systems

By utilizing advanced machine learning models like , the system achieves high accuracy in identifying offensive language, reducing the prevalence of harmful content.

1) Complex Understanding: 's ability to understand the context of words in sentences allows for a more nuanced detection of offensive language, surpassing simpler keyword-based filters.

2) Automated Moderation: Automates the content moderation process, significantly reducing the manual effort required and enabling real-time content filtering.

3) Scalability: Machine learning models can handle vast amounts of online content efficiently, making the system highly scalable and suitable for platforms of any size.

4) Continuous Learning: The algorithms can be trained on new datasets to adapt to evolving language and slang, ensuring the system remains effective over time.

5) Precision and Recall Balance: Careful tuning of models ensures a balance between precision (minimizing false positives) and recall (minimizing false negatives), optimizing the moderation process.

6) Versatility: The system can be applied across various digital platforms, including social media, forums, and chat applications, to foster safer online communities.

7) Reduced Bias: Through rigorous training and evaluation, the models aim to minimize bias in content moderation, promoting fairness and equity in online interactions.

8) User Experience Improvement: By effectively filtering out offensive content, the system contributes to a more positive and respectful online environment, enhancing user satisfaction.

METHODOLOGY:

DATA-CLEANING:COUNT VECTORIZER:**CountVectorizer** is a technique used in the preprocessing step of text analysis, particularly within the domain of Natural Language Processing (NLP). It's part of the data cleaning process because it converts text data into a numerical format that machine learning algorithms can work with. Here's how it fits into the data cleaning pipeline:

1. **Tokenization**: CountVectorizer starts by tokenizing the text, which means breaking it down into individual words or terms.
2. **Building a Vocabulary**: It then creates a vocabulary of all the unique tokens from the entire text corpus.
3. **Counting Occurrences**: For each document or piece of text, CountVectorizer counts the number of occurrences of each term from the vocabulary.
4. **Creating a Document-Term Matrix**: The result is a document-term matrix where each row represents a document and each column represents a term from the vocabulary. The values in the matrix are the term frequencies.
5. **Handling Case Sensitivity**: By default, CountVectorizer converts all characters to lowercase before tokenizing, which helps in maintaining consistency across different uses of the same word (e.g., “House” vs. “house”).
6. **Removing Stop Words**: It can also remove stop words—common words that are usually filtered out before processing because they carry less meaningful information about the content of the text.
7. **Custom Preprocessing**: CountVectorizer allows for custom preprocessing of text data. This means you can define your own function to clean the text data in a way that’s specific to your needs before it’s vectorized.

By transforming text data into a numerical format, CountVectorizer facilitates the application of various machine learning models to text data, which is essential for tasks such as sentiment analysis, topic modelling, and document classification.

NLP ALGORITHM: Natural Language Processing (NLP) algorithms play a crucial role in the analysis of reviews, providing valuable insights into customer sentiment and preferences. Here’s an overview of how NLP algorithms are used in review analysis:

Understanding Customer Sentiment: NLP algorithms can process textual data from reviews to determine the sentiment behind them. This involves classifying the text into categories such as positive, negative, or neutral. Techniques like sentiment analysis help businesses understand how customers feel about their products or services[1].

Extracting Key Information: NLP algorithms can identify important information within the reviews, such as specific features of a product that customers frequently mention. This can involve the extraction of named entities or key phrases that highlight what aspects are being discussed[2].

Topic Modelling: Algorithms like Latent Dirichlet Allocation (LDA) are used to discover the abstract topics that occur in a collection of documents. In the context of reviews, topic modeling can reveal the main themes that customers talk about, providing businesses with a better understanding of customer concerns and interests[1].

Text Preprocessing: Before analysis, reviews often undergo preprocessing steps such as tokenization, stemming, and lemmatization. These steps help in reducing the complexity of the text and in standardizing the data for better analysis by NLP algorithms[2].

Feature Importance: NLP can also determine which words or phrases are most indicative of a particular sentiment or topic. This helps in understanding which aspects of a product or service are most impactful on customer opinions[3].

Predictive Analysis: Some NLP algorithms can predict future customer behavior based on past reviews. For example, they can forecast the likelihood of a customer recommending a product or the potential for a product to become popular[4].

In summary, NLP algorithms are essential for transforming raw review data into structured insights that can inform business strategies and improve customer experiences.

NAIVE BAYES: The Naive Bayes algorithm is a probabilistic machine learning model that's used for classification tasks, and it's particularly effective in the field of text analysis, such as review analysis. Here's an overview of how it's applied:

Principles of Naive Bayes: Naive Bayes is based on Bayes' Theorem, which uses probability to make predictions. In the context of review analysis, it estimates the likelihood that a review belongs to a certain class (e.g., positive, negative) based on the data provided.

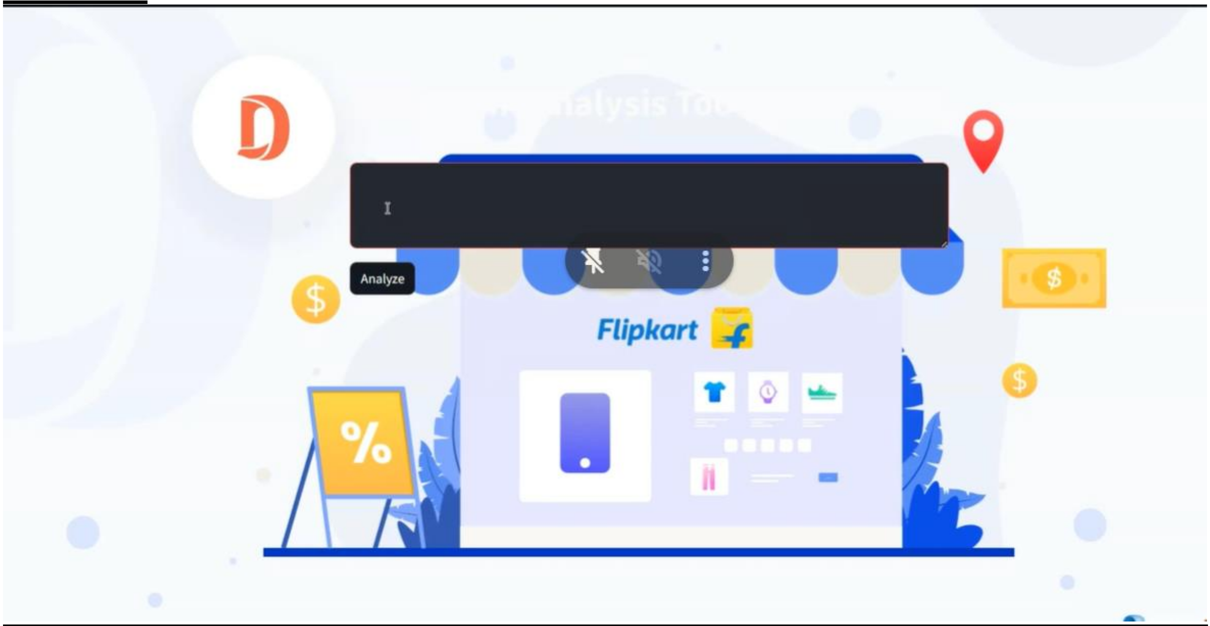
Application in Review Analysis: When analysing reviews, the Naive Bayes algorithm considers each word as a feature and calculates the probability of each class based on the frequency of these words. It operates under the 'naive' assumption that all features (words) are independent of each other, simplifying the computation of probabilities.

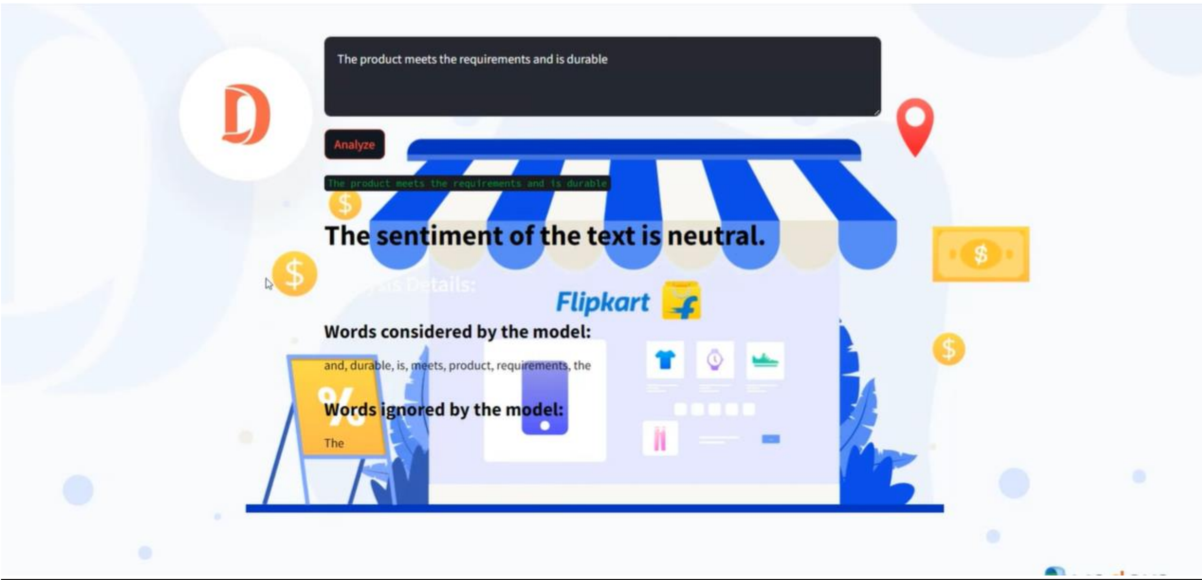
Training the Model: To prepare the Naive Bayes classifier for review analysis, it must be trained on a dataset of reviews that have already been labelled with their sentiments. This training involves learning the probability of each word occurring in positive and negative reviews.

Predicting Sentiments: Once trained, the Naive Bayes model can predict the sentiment of new, unlabelled reviews. It does this by applying the probabilities learned during training to the words in the new review and classifying it based on the highest probability.

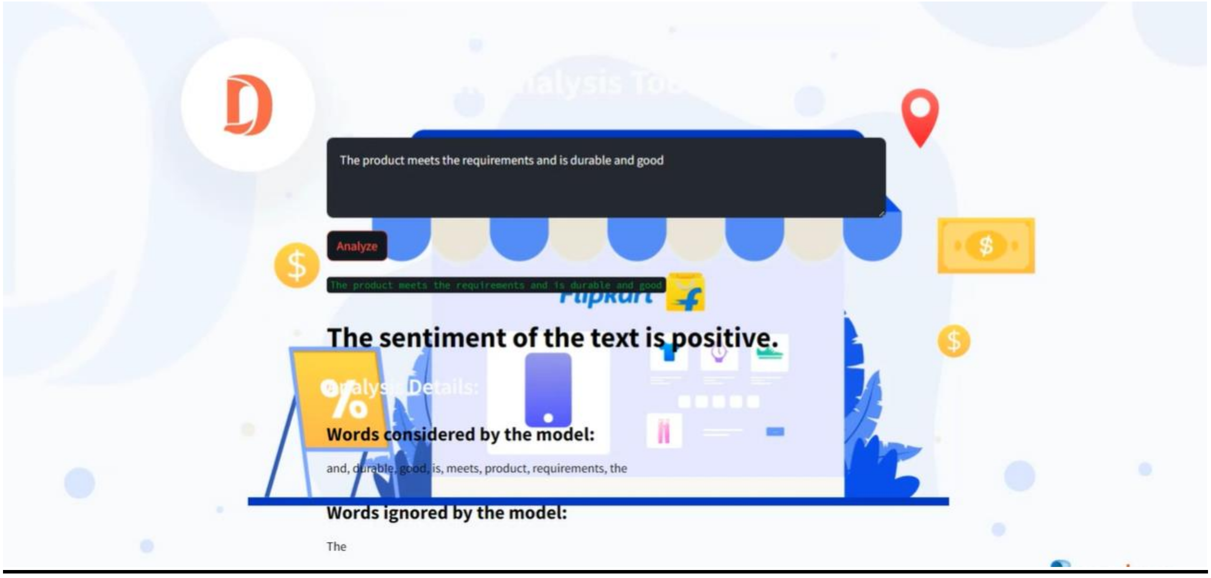
In review analysis, Naive Bayes is often used as a baseline model due to its simplicity and effectiveness. It's particularly good at handling large volumes of text data and providing a quick, initial understanding of the sentiments expressed in customer reviews.

OUTPUT:


Analyze
Flipkart


The product meets the requirements and is durable
Analyze
The sentiment of the text is neutral.
Flipkart
Words considered by the model:
and, durable, is, meets, product, requirements, the
Words ignored by the model:
The

Industrial Engineering Journal
ISSN: 0970-2555
Volume: 53, Issue4, April : 2024



CONCLUSION:

In Conclusion, Key outcomes of the project included the ability to quantify abstract concepts such as customer loyalty and satisfaction. The ML models provided predictive insights into customer behaviour, enabling businesses to anticipate market trends and customer needs with unprecedented accuracy.

Moreover, the project highlighted the importance of continuous learning and adaptation. The iterative nature of ML ensured that the models evolved with the data, capturing the ever-changing dynamics of customer feedback.

In conclusion, the feedback analysis project demonstrated the transformative potential of machine learning in interpreting customer feedback. It has set a precedent for future endeavours in the realm of data analytics, paving the way for more informed and customer-centric business strategies.

REFERENCES:

• S. ChandraKala1 and C. Sindhu2, “OPINION MINING AND SENTIMENT CLASSIFICATION: A SURVEY,”.Vol .3(1),Oct 2012,420-427

•G.Angulakshmi , Dr.R.ManickaChezian ,”An Analysis on Opinion Mining: Techniques and Tools”. Vol 3(7), 2014 www.iarcce.com.

• Callen Rain,”Sentiment Analysis in Amazon Reviews Using Probabilistic Machine Learning” Swarthmore College, Department of Computer Science.

• Padmani P .Tribhuvan,S.G. Bhirud,Amrapali P. Tribhuvan,” A Peer Review of Feature Based Opinion Mining and Summarization”(IJCSIT) International Journal of Computer Science and Information Technologies, Vol. 5 (1), 2014, 247-250 ,www.ijcsit.com.

Industrial Engineering Journal
ISSN: 0970-2555
Volume: 53, Issue4, April : 2024

•Carenini, G., Ng, R. and Zwart, E. Extracting Knowledge from Evaluative Text. Proceedings of the Third International Conference on Knowledge Capture (K-CAP'05), 2005.

• Dave, D., Lawrence, A., and Pennock, D. Mining the Peanut Gallery: Opinion Extraction and Semantic Classification of Product Reviews. Proceedings of International World Wide Web Conference (WWW'03), 2003.

• Zhu, Jingbo, et al. "Aspect-based opinion polling from customer reviews." IEEE Transactions on Affective Computing, Volume 2.1,pp.37-49, 2011.

• Na, Jin-Cheon, Haiyang Sui, Christopher Khoo, Syin Chan, and Yunyun Zhou. "Effectiveness of simple linguistic processing in automatic sentiment classification of product reviews." Advances in Knowledge Organization Volume9, pp. 49-54, 2004.

•Nasukawa, Tetsuya, and Jeonghee Yi. "Sentiment analysis: Capturing favorability using natural language processing." In Proceedings of the 2nd international conference on Knowledge capture, ACM, pp. 70-77, 2003.

• Li, Shoushan, Zhongqing Wang, Sophia Yat Mei Lee, and Chu-Ren Huang. "Sentiment Classification with Polarity Shifting Detection." In Asian Language Processing (IALP), 2013 International Conference on, pp. 129-132. IEEE, 2013.