International Journal For Advanced Research In Science & Technology
A peer reviewed international journal
www.ijarst.in
IJARST
ISSN: 2457-0362

Outlook Analysis of E-Commerce Product Reviews

Pavan.M[1], Assistant Professor, Department of Computer Science Engineering, Andhra Loyola Institute of Engineering and Technology, Vijayawada, India.
Deepak.Ch.V.V.R.N.S[2], Department of Computer Science Engineering, Andhra Loyola Institute of Engineering and Technology, Vijayawada, India.(19HP1A05C3).
Mahesh.Ch[3], Department of Computer Science Engineering, Andhra Loyola Institute of Engineering and Technology, Vijayawada, India.(19HP1A05A8).

Abstract
This study aims to conduct sentiment analysis of e-commerce product reviews using natural language processing (NLP) techniques. The analysis involves extracting and classifying the sentiment expressed in the reviews into positive, negative or neutral categories. By using machine learning algorithms, we seek to identify patterns and trends in the data to gain insights into customer satisfaction levels and perceptions of different products.

Keywords: Sentiment Analysis, reviews, patterns, Customer Satisfaction levels.

Introduction

In the era of e-commerce, online product reviews have become an essential source of information for customers looking to make informed purchasing decisions. The sheer volume of reviews available for any given product can be overwhelming, but by analyzing these reviews, businesses can gain valuable insights into customer attitudes and preferences. Outlook analysis of e-commerce product reviews involves analyzing the content of reviews to identify emerging trends and themes, as well as evaluating the sentiment expressed in the reviews. This type of analysis can help businesses understand customer needs and preferences.

Objective

The objective of outlook analysis of e-commerce product reviews is to gain insights into customer perceptions and attitudes towards different products by analysing the content of reviews. The analysis aims to identify emerging trends and themes in customer feedback, evaluate the sentiment expressed in reviews, and analyse the factors that influence customer satisfaction and dissatisfaction. The ultimate goal of this analysis is to provide businesses with actionable insights to improve their products and services, enhance customer satisfaction and loyalty, and develop effective marketing strategies.

Scope

The scope of outlook analysis of e-commerce product reviews is focused on gaining insights into customer perceptions and attitudes towards different products. This involves analyzing product reviews from one or more e-commerce platforms, focusing on specific product categories, and using both quantitative and qualitative analysis techniques to evaluate sentiment, identify key themes and patterns, and analyze factors that influence customer satisfaction and dissatisfaction.

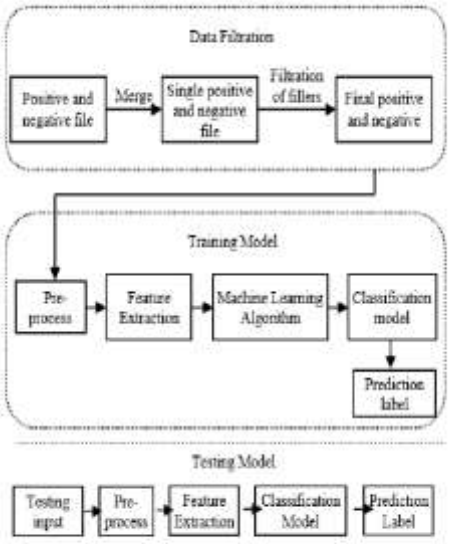
Methodology

A. Natural Language Processing

With the help of NLP, we can determine the review is positive or not. In python, we have the library nltk(natural language toolkit) which is used for NLP. In NLP, we have the technique named “word cloud” which is visual representation of data i.e., reviews. For generating word cloud in python, modules required are matplotlib, pandas and word cloud. The required modules are used for “splitting the words in reviews and checks with the positive words and negative words in word cloud. This proposed algorithm has three phases. They are

- Data Filtration
- Training model
- Testing model

International Journal For Advanced Research In Science & Technology
A peer reviewed international journal
www.ijarst.in
IJARST
ISSN: 2457-0362



B. Data Filtration

- The first step is to pre-process the data by removing irrelevant information such as punctuations, numbers, and emotions.
- The next step is to filter the data to remove any remaining irrelevant information.

Training Model

1. Load the dataset from a file (e.g., a CSV or text file) containing text reviews or comments along with their corresponding labels (e.g., positive or negative).
2. Tokenize the text data to extract individual words, punctuation marks, and other relevant features.
3. Filter the extracted words to remove stop words (e.g., "the", "and", "is", etc.) and retain only words that are relevant for sentiment analysis, such as adjectives, adverbs, and verbs.
4. Label the data as "pos" or "neg" based on their corresponding sentiment scores.
5. Compute the frequency distribution of the remaining words and select the top 5000 most frequent words as features for training.
6. Shuffle the labelled data using a random seed to ensure a random distribution of samples during training.
7. Split the labelled data into training and testing sets, typically using a 70-30% split.
8. Train several classification algorithms, such as Naïve Bayes, Linear Model, SVM, and Decision Tree, on the training data.
9. Evaluate the performance of each algorithm on the testing data and select the best-performing model.
10. Use the selected model to classify new, unseen text data based on their sentiment scores.
11. This approach is a standard method for sentiment analysis and can be adapted for other text classification tasks as well.

Testing Model

Once the model is trained and evaluated on the testing data, it can be used for predicting the sentiment of new, unseen text data. However, before feeding the new data into the model for prediction, it's important to pre-process the input data in a similar way as the training data to ensure consistency and accuracy in the predictions.

The pre-processing steps for new input data may include:

1. Removing any symbols, special characters, or numbers from the text.
2. Tokenizing the text into individual words and filtering out any stop words.
3. Mapping the remaining words to the features used during training, typically based on the top 5000 most frequent words.
4. Feeding the pre-processed input data into the trained model to generate predictions.
5. After predicting the sentiment of new input data, the results can be analyzed and evaluated to measure the accuracy and performance of the model.

Experimental Setup

1. Collect a large dataset of ecommerce product reviews related to the product(s) of interest.
2. Pre-process the collected dataset by cleaning and standardizing the text data.
3. Label the pre-processed data as positive or negative sentiment based on their corresponding review ratings.
4. Split the labelled data into training and testing sets using a 70-30% split.
5. Extract relevant features from the pre-processed data using techniques such as frequency analysis, TF-IDF, or word

International Journal For Advanced Research In Science & Technology
A peer reviewed international journal
www.ijarst.in
IJARST
ISSN: 2457-0362

embeddings.

6. Train a suitable machine learning model such as Naive Bayes, Logistic Regression, or Support Vector Machines (SVM) on the extracted features using the training data.
7. Evaluate the performance of the trained model on the testing data using metrics such as accuracy, precision, recall, and F1 score.
8. Apply the trained sentiment analysis model to the remaining unlabelled reviews to classify them as either positive or negative sentiment.
9. Analyse the classified reviews to identify the key factors influencing customer sentiment towards the product(s) of interest.
10. Present the results of the outlook analysis in a report or dashboard format, including visualizations such as word clouds, sentiment histograms, or time-series plots to help identify trends and insights.

Experimental Results

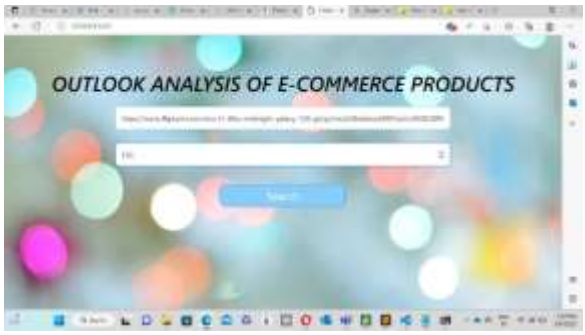
The experiment results for outlook analysis of ecommerce product reviews would depend on the specific dataset, pre-processing techniques, machine learning model, and evaluation metrics used. However, here are some possible outcomes that could be observed:

1. The trained sentiment analysis model achieves a high accuracy, precision, recall, and F1 score on the testing data, indicating that it is effective in predicting the sentiment of ecommerce product reviews.

2. The classified reviews are analysed to identify the key factors influencing customer sentiment towards the product(s) of interest. This may include identifying common positive and negative keywords, sentiment trends over time, or frequently mentioned complaints and suggestions for improvement.

3. The outlook analysis reveals insights into the overall sentiment towards the product(s) of interest, potential areas for improvement, and areas of strength. This information can be used to guide product development, marketing, and customer support strategies.

4. The visualization of the outlook analysis results, such as word clouds, sentiment histograms, or time-series plots, can help to communicate the insights and trends in a clear and understandable way.



Conclusion

In conclusion, outlook analysis of ecommerce product reviews is a valuable tool for businesses to gain insights into customer sentiment towards their products. By collecting and pre-processing a large dataset of ecommerce product reviews, training a sentiment analysis model, and analyzing the classified reviews, businesses can identify key factors influencing customer sentiment and areas for improvement. The results of the outlook analysis can guide product development, marketing, and customer support strategies, and help businesses improve their overall customer experience.

References

[1] [1] Samaneh Moghaddam and Martin Ester, "Opinion Digger: An Unsupervised Opinion Miner from Unstructured Product Reviews", Proceedings of 19th ACM International Conference on Information and Knowledge Management, pp. 1825-1828, 2010.

[2] [2] Aurangzeb Khan, Baharum Baharudin and Khairullah Khan, "Sentiment Classification from Online Customer Reviews using Lexical Contextual Sentence Structure", Proceedings of International Conference on Software Engineering and Computer Systems, pp. 317-331, 2011.

[3] [3] M. Hu and B. Liu, "Mining and Summarizing Customer Reviews", Proceedings of 10th ACM International Conference on Knowledge Discovery and Data Mining, pp. 166-177, 2005.

[4] [4] A.M. Popescu and O. Etzioni, "Extracting Product Features and Opinions from Reviews", Proceedings of International Conference on Human Language Technology and Empirical Methods in Natural Language Processing, pp. 339-346, 2005.

[5] [5] G. Vinodhini and R.M. Chandrasekaram, "Sentiment Analysis and opinion Mining: A Survey", International Journal of Advanced Research in Computer Science and Software Engineering, Vol. 2, No. 6, pp. 28-35, 2012.

[6] [6] A. Collomb, C. Costea, D. Joyeux, O. Hasan and L. Brunie, "A Study and Comparison of Sentiment Analysis Methods for Reputation Evaluation", Available at: https://liris.cnrs.fr/Documents/Liris-6508.pdf.

[7] Mika V. Mantyla, Daniel Graziotin and Miikka Kuutila, "The Evolution of Sentiment Analysis-A Review of Research Topics", Computer Science Review, Vol. 27, No. 1, pp. 16-32, 2018.

[8] [8] F. Benamara, C. Cesarano and D. Reforgiato, "Sentiment Analysis: Adjectives and Adverbs are better than Adjectives Alone", Proceedings of International Conference on Weblogs and Social Media, pp. 1-7, 2006.

[9] [9] R.A. Hummel and S.W. Zucker, "On the Foundations of Relaxation Labeling Processes", Proceedings of International Conference on Computer Vision: Issues, Problems, Principles, and Paradigms, pp. 585-605, 1987.

[10] [10] Samaneh Moghaddam and Martin Ester, "ILDA: Interdependent LDA Model for Learning Latent Aspects and their Ratings from Online Product Reviews", Proceedings of 34th International ACM Conference on Research and Development in Information Retrieval, pp. 665-674, 2011.

[11] [11] Jorge Carrillo De Albornoz, Laura Plaza, Pablo Gervas and Alberto Diaz, "A Joint Model of Feature Mining and Sentiment Analysis for Product Review Rating", Proceedings of International Conference on Advances in Information Retrieval, pp. 55-66, 2011.

[12] [12] Sentiment Analysis, Available at: https://insightsatlas.com/sentiment-analysis/

[13] [13] Andrew L. Maas, Raymond E. Daly, Peter T. Pham, Dan Huang, Andrew Y. Ng and Christopher Potts, "Learning Word Vectors for Sentiment Analysis", Proceedings of 49th Annual Meeting of the Association for Computational Linguistics, pp. 1-7, 2011.

[14] [14] Understanding Sentiment Analysis: What It Is and Why It's Used Understanding Sentiment Analysis: What It Is and Why It's Used, Available at:https://www.brandwatch.com/blog/understanding-sentiment-analysis/

[15] [15] Sentiment Analysis Explained, Available at: https://www.lexalytics.com/technology/sentiment-analysis.