iJRASET
International Journal For Research in
Applied Science and Engineering Technology

INTERNATIONAL JOURNAL FOR RESEARCH
IN APPLIED SCIENCE & ENGINEERING TECHNOLOGY

Volume: 14 Issue: IV Month of publication: April 2026

DOI: https://doi.org/10.22214/ijraset.2026.78754

International Journal for Research in Applied Science & Engineering Technology (IJRASET)
ISSN: 2321-9653; IC Value: 45.98; SJ Impact Factor: 7.538
Volume 14 Issue IV Apr 2026- Available at www.ijraset.com

Polarity Detection and Aspect-Based Sentiment Analysis of E-Commerce Product Reviews

Prof. Vidya Yengur[1], Krishna Mundada[2], Sarthak More[3], Omkar Kute[4]
Department of Computer Engineering, JSPM University, Pune

Abstract: With the rapid growth of e-commerce platforms, customer reviews have become an essential source of informa- tion for both consumers and businesses. However, extracting meaningful insights from large volumes of textual data is a challenging task. Traditional sentiment analysis methods provide only an overall opinion, which fails to capture detailed feedback about specific product features. This paper presents an aspect- based sentiment analysis system that identifies product aspects and determines sentiment for each aspect individually. The system uses Natural Language Processing techniques along with a transformer-based model to improve accuracy and contextual understanding. The proposed approach provides detailed insights that help businesses improve products and assist customers in making informed decisions.
Index Terms: Sentiment Analysis, Aspect-Based Sentiment Analysis, NLP, DistilBERT, E-commerce

I. INTRODUCTION

In recent years, the growth of e-commerce platforms such as Amazon and Flipkart has significantly increased the amount of user-generated content. Customers frequently share their experiences in the form of reviews, which influence purchasing decisions of other users.

However, analyzing this large volume of textual data man- ually is time-consuming and inefficient. Traditional sentiment analysis techniques classify entire reviews as positive, nega- tive, or neutral. While useful, these methods fail to capture detailed opinions about individual product features.

For example, a customer may appreciate the design of a product but criticize its performance. Traditional systems cannot effectively represent such mixed opinions. Therefore, there is a need for a more advanced approach that can analyze sentiment at a finer level.

This paper proposes an aspect-based sentiment analysis system that identifies specific product features and assigns sentiment to each feature, providing a more detailed and meaningful analysis.

In addition to influencing purchasing decisions, customer reviews also act as a feedback mechanism for businesses. Companies continuously analyze reviews to identify patterns in customer satisfaction and dissatisfaction. However, due to the unstructured nature of textual data, extracting meaningful insights remains a complex task.

Another major challenge is the diversity in writing styles. Different users express opinions in different ways, including slang, abbreviations, and informal language. This variabilitymakes it difficult for traditional systems to interpret sentiment accurately.

Furthermore, the rapid growth of mobile commerce has increased the volume of short and informal reviews. These reviews often lack grammatical structure, making sentiment detection more challenging. Therefore, advanced models that can understand context and semantics are required.

Aspect-based sentiment analysis addresses these challenges by focusing on specific components of a product rather than treating the review as a whole. This allows for a more precise understanding of user opinions.

II. BACKGROUND AND MOTIVATION

Customer reviews play a critical role in shaping product perception in the online marketplace. Businesses rely on these reviews to understand customer needs and improve product quality.

The motivation behind this work is to overcome the limi- tations of traditional sentiment analysis. Existing systems do not provide detailed insights into which features are liked or disliked by users.

By using aspect-based sentiment analysis, businesses can identify strengths and weaknesses of a product more effec- tively. This also helps customers compare products based on specific features rather than overall ratings.

III. LITERATURE REVIEW

Several approaches have been proposed in the field of sentiment analysis.

Lexicon-based methods use predefined dictionaries of words with associated sentiment scores. These methods are simple but fail to understand context.

Machine learning techniques such as Naive Bayes and Support Vector Machines use statistical methods to classify sentiment. These methods improve performance but still strug- gle with complex language patterns.

Deep learning models such as Recurrent Neural Networks (RNNs) and Long Short-Term Memory (LSTM) networks cap- ture sequential information in text but require large datasets.

Recently, transformer-based models such as DistilBERT have shown significant improvements in performance. These models understand context better and provide more accurate results.

Recent research in sentiment analysis has focused on improving contextual understanding using transformer-based architectures. These models utilize attention mechanisms to capture relationships between words in a sentence, allowing them to understand meaning more effectively.

Another area of research involves hybrid approaches that combine rule-based and machine learning techniques. These systems aim to leverage the strengths of both approaches to improve accuracy and robustness.

Additionally, researchers have explored domain-specific sentiment analysis, where models are trained on data related to a particular product category such as electronics or clothing. This improves performance as the model becomes familiar with domain-specific vocabulary.

Despite these advancements, challenges such as sarcasm detection, implicit aspect identification, and multilingual sen- timent analysis remain open research problems.

IV. PROBLEM STATEMENT

Traditional sentiment analysis assigns a single sentiment label to an entire review. This leads to loss of important information, especially when reviews contain mixed opinions.

The problem addressed in this paper is to develop a system that:

- Extracts product features from reviews
- Assigns sentiment to each feature
- Provides structured and meaningful insights

V. PROPOSED SYSTEM

The proposed system is designed as a pipeline consisting of multiple stages. The proposed system is designed to handle real-world data efficiently. It follows a modular architecture where each component performs a specific task, making the system flexible and scalable.

One of the key strengths of the system is its ability to process reviews dynamically. Instead of relying on preloaded datasets, the system extracts live data from e-commerce plat- forms, ensuring that the analysis is always up-to-date.

Another important feature is the separation of aspect ex- traction and sentiment classification. This modular design allows independent improvement of each component without affecting the overall system.

The system also ensures efficient handling of large datasets by processing reviews in batches. This improves performance and reduces processing time.

A. Data Collection

Product reviews are collected from e-commerce websites using web scraping techniques. The system takes a product URL as input and extracts relevant reviews.

B. Text Preprocessing

The collected data is cleaned to remove unnecessary ele- ments such as HTML tags, special characters, and irrelevant text. This improves the quality of input data.

C. Aspect Extraction

Product features such as “battery”, “camera”, and “display” are identified using Natural Language Processing techniques. The system uses spaCy to extract meaningful aspects from the text.

D. *Sentiment Classification*

Each extracted aspect is analyzed using a transformer- based model (DistilBERT) to determine sentiment. The model classifies sentiment as positive, negative, or neutral.

E. *Visualization*

The results are presented using a user-friendly interface. Charts and tables are used to display aspect-wise sentiment analysis.

VI. SYSTEM ARCHITECTURE

The system architecture consists of the following compo- nents:

- User Input Module
- Web Scraping Module
- Preprocessing Module
- Aspect Extraction Module
- Sentiment Classification Module
- Visualization Module

The system ensures smooth flow of data from input to output,The architecture of the system is designed to ensure smooth communication between different modules. Each mod- ule operates independently but is connected through a well- defined pipeline. The data flow begins with user input and moves through various stages such as data extraction, preprocessing, analysis, and visualization. This structured approach ensures that errors in one module do not propagate to others.

The backend system manages data processing and model execution, while the frontend provides an intuitive interface for users. The separation of frontend and backend improves maintainability and scalability.

The system architecture also supports future enhancements, such as integration with cloud services or deployment on larger platforms.

VII. IMPLEMENTATION DETAILS

The system is implemented using modern technologies. The backend is developed using Python. Natural Language Processing tasks are performed using spaCy. Sentiment clas- sification is carried out using the DistilBERT model.

The frontend is developed using HTML, CSS, JavaScript, and Bootstrap. The system provides an interactive interface for users to view results. The implementation focuses on achieving a balance between performance and efficiency. The use of a lightweight transformer model ensures faster processing without compromising accuracy.

The backend handles tasks such as data extraction, pre- processing, and model inference. Efficient data structures and libraries are used to manage large volumes of text data.

On the frontend, responsive design techniques are used to ensure compatibility across different devices. The interface is designed to provide clear and concise visualization of results. Error handling mechanisms are also implemented to manage issues such as invalid URLs or missing data. This improves the reliability of the system.

VIII. WORKING OF THE SYSTEM

The working of the system can be described step-by-step:

1) User enters the product URL
2) Reviews are extracted from the website
3) Text is cleaned and processed
4) Aspects are identified from the reviews
5) Sentiment is assigned to each aspect
6) Results are displayed to the user

The system operates in a sequential and automated manner to ensure efficient processing of user input and generation of meaningful results.

Initially, the user provides a product URL from an e- commerce platform. The system validates the input to ensure that the link is correct and accessible. Once validated, the web scraping module extracts customer reviews associated with the product.

The extracted reviews are often unstructured and may con- tain irrelevant information such as advertisements, emojis, or formatting tags. Therefore, the preprocessing module cleans the data and converts it into a structured format suitable for analysis.

After preprocessing, the system identifies important product features using linguistic patterns and dependency parsing techniques.

These features are treated as aspects for further analysis.
Each extracted aspect is then passed to the sentiment clas- sification module. The model evaluates the context in which the aspect appears and assigns an appropriate sentiment label. Finally, the processed data is aggregated and presented to the user in a clear and understandable format. The system ensures that users can easily interpret the results without requiring technical knowledge.
The overall workflow is designed to minimize manual intervention and provide accurate results in a short amount of time.

IX. RESULTS AND DISCUSSION

The system was tested on multiple product reviews collected from e-commerce platforms.
The proposed system achieved an accuracy of approxi- mately 90%. The results show that the system can effectively identify product aspects and assign appropriate sentiment.
Compared to traditional sentiment analysis, the proposed system provides more detailed insights. It helps in understand- ing user opinions at a deeper level. The experimental results demonstrate that the system performs well across different types of products. The model is able to correctly identify aspects and assign sentiment even in complex sentences.
It was observed that transformer-based models perform better than traditional methods in handling contextual informa- tion. This leads to improved accuracy and better classification results.
The system also shows good scalability, as it can handle multiple reviews without significant performance degradation. This makes it suitable for real-world applications.
However, certain challenges were observed in cases where reviews contained sarcasm or implicit meanings. These cases require more advanced techniques for accurate interpretation.

X. ADVANTAGES OF PROPOSED SYSTEM

1) Provides detailed feature-level analysis
2) Improves decision-making for users
3) Helps businesses identify product issues
4) High accuracy due to transformer-based model
5) Provides detailed understanding of customer opinions
6) Helps businesses improve specific product features
7) Reduces manual effort in analyzing reviews
8) Supports real-time data processing
9) Scalable architecture for handling large datasets
10) Improves customer satisfaction by identifying key issues

XI. APPLICATIONS

1) E-commerce product analysis
2) Customer feedback evaluation
3) Market research
4) Product improvement strategies
5) Enables customers to evaluate products based on spe- cific features such as performance, durability, and design rather than relying only on overall ratings.
6) Assists e-commerce companies in analyzing customer feedback to identify strengths and weaknesses of their products.
7) Supports market research by providing insights into con- sumer preferences and emerging trends across different product categories.
8) Can be integrated into recommendation systems to sug- gest products based on feature-level user preferences.
9) Useful in social media analytics to monitor public opinion about brands, products, or services.
10) Helps customer support teams automatically detect neg- ative feedback and prioritize issue resolution.
11) Assists manufacturers in improving product quality by identifying frequently reported issues in reviews.
12) Can be applied in business intelligence systems for data- driven decision-making.

XII. LIMITATIONS

1) Difficulty in detecting sarcasm
2) Limited support for multiple languages
3) Dependency on quality of data
4) The system may struggle to correctly interpret sarcastic or ironic statements, which can lead to incorrect sentiment classification. Difficulty in detecting implicit aspects where the feature is not explicitly mentioned in the review.
5) Performance depends heavily on the quality and com- pleteness of the extracted review data.
6) Limited ability to handle multiple languages if the model is trained only on English datasets.
7) Requires computational resources for processing large volumes of data, especially in real-time applications.
8) Domain-specific variations in language may reduce accu- racy when applied to unfamiliar product categories.
9) The system may misinterpret very short or ambiguous reviews due to lack of sufficient context.
10) Dependence on external websites for data extraction may cause issues if website structure changes.

XIII. FUTURE SCOPE

Future improvements can include: Future work can focus on improving the system's ability to handle complex language patterns such as sarcasm and irony.

Another important direction is multilingual sentiment anal- ysis, which will allow the system to analyze reviews written in different languages.

Integration with mobile applications can make the system more accessible to users. Additionally, real-time streaming of reviews can further enhance the system's capabilities.

Advanced deep learning models can also be explored to improve accuracy and performance.

1) Support for multiple languages
2) Real-time sentiment analysis
3) Improved aspect extraction techniques
4) The system can be extended to support multiple lan- guages, allowing analysis of reviews written in regional and international languages.
5) Future work can focus on improving the detection of sarcasm and irony, which remain challenging problems in sentiment analysis.
6) Integration with real-time data streaming can enable con- tinuous monitoring of customer feedback as new reviews are posted.
7) Advanced deep learning models can be explored to fur- ther improve accuracy and contextual understanding.
8) The system can be enhanced to identify implicit aspects where features are not directly mentioned but inferred from context.
9) Security and privacy features can be added to ensure safe handling of user data and reviews.
10) The system can be extended to analyze multimedia con- tent such as images and videos along with textual reviews.
11) Automated summarization techniques can be incorpo- rated to generate concise summaries of large volumes of reviews.
12) Continuous learning mechanisms can be introduced to up- date the model based on new data and evolving language patterns.

XIV. CONCLUSION

This paper presents an aspect-based sentiment analysis system that provides detailed insights into customer reviews. The use of advanced NLP techniques and transformer-based models improves accuracy and performance.

The system helps both businesses and customers by provid- ing meaningful and structured information. It can be further improved to handle more complex scenarios in the future. This paper presented a comprehensive approach to aspect- based sentiment analysis for e-commerce product reviews. The system focuses on extracting meaningful insights from unstructured textual data by identifying product features and analyzing sentiment at a granular level.

The use of Natural Language Processing techniques com- bined with transformer-based models enables the system to achieve high accuracy and better contextual understanding compared to traditional methods.

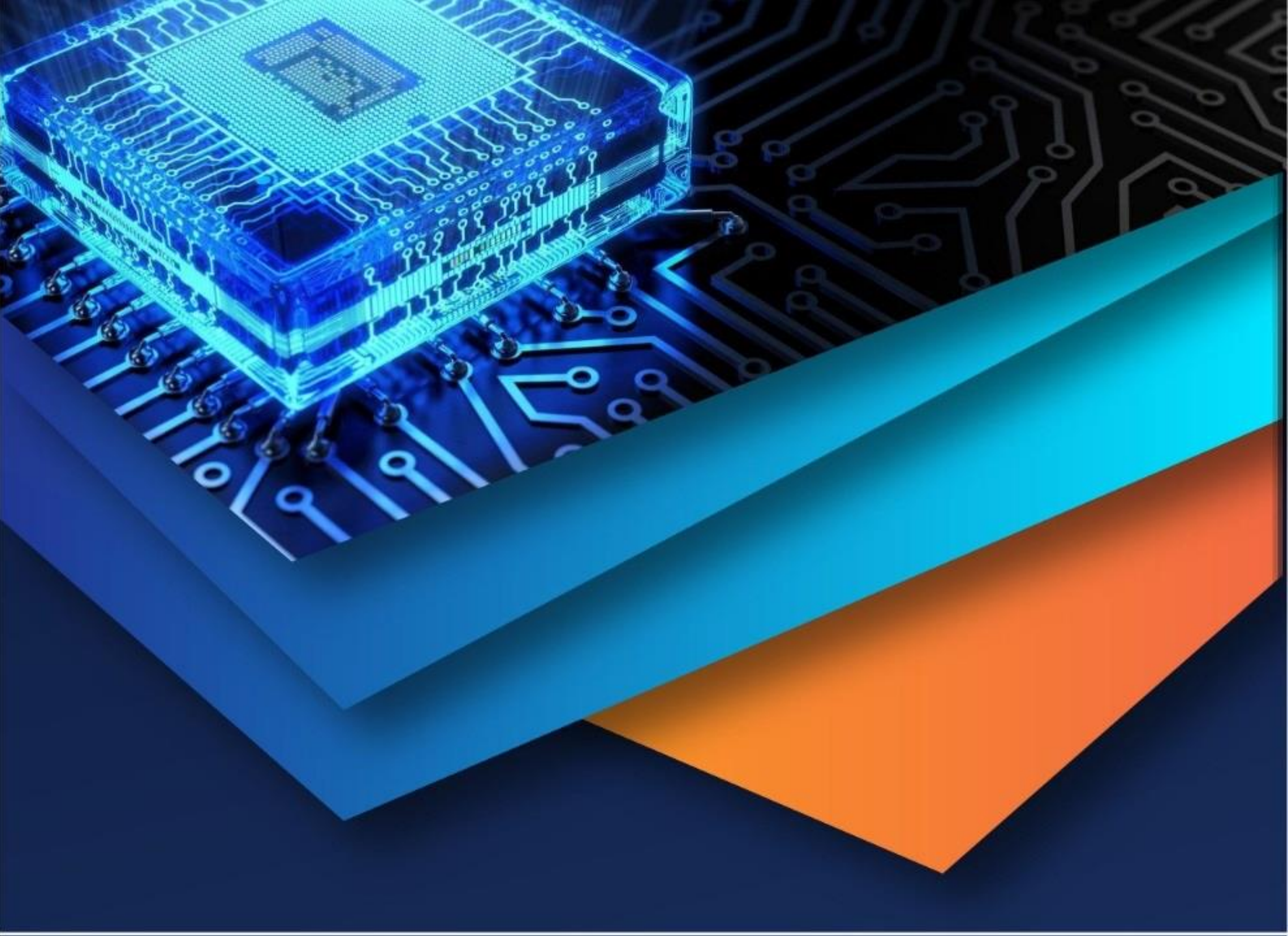
The proposed approach not only improves the quality of sentiment analysis but also provides valuable insights for both customers and businesses. Customers benefit from detailed product evaluations, while businesses gain a deeper under- standing of user feedback.

The modular design of the system ensures flexibility and allows for future enhancements. With continuous advance- ments in machine learning and natural language processing, the system can be further improved to handle more complex scenarios.

Overall, the proposed system demonstrates the potential of advanced sentiment analysis techniques in transforming the way customer feedback is analyzed and utilized.

REFERENCES

[1] M. Hu and B. Liu, “Mining and Summarizing Customer Reviews,” 2004.
[2] J. Devlin et al., “BERT: Pre-training of Deep Bidirectional Transform- ers,” 2018.
[3] spaCy Documentation.
[4] Hugging Face Transformers Documentation.
[5] B. Pang and L. Lee, “Opinion Mining and Sentiment Analysis,” Foun- dations and Trends in Information Retrieval, 2008.
[6] A. Vaswani et al., “Attention Is All You Need,” 2017.
[7] I. Goodfellow et al., “Deep Learning,” MIT Press, 2016.
[8] T. Liu, “Sentiment Analysis and Opinion Mining,” Morgan & Claypool Publishers, 2012.
[9] E. Cambria, D. Das, S. Bandyopadhyay, and A. Feraco, “A Practical Guide to Sentiment Analysis,” Springer, 2017.
[10] K. Ravi and V. Ravi, “A survey on opinion mining and sentiment anal- ysis: tasks, approaches and applications,” Knowledge-Based Systems, vol. 89, pp. 14–46, 2015.
[11] Z. Zhang, D. Robinson, and J. Tepper, “Detecting hate speech on Twitter using a convolutional neural network,” in European Semantic Web Conference, 2018.
[12] R. Socher et al., “Recursive deep models for semantic compositionality over a sentiment treebank,” in Proceedings of EMNLP, 2013.
[13] A. Severyn and A. Moschitti, “Twitter sentiment analysis with deep convolutional neural networks,” in Proceedings of SIGIR, 2015.
[14] G. Hinton, L. Deng, D. Yu, et al., “Deep neural networks for acoustic modeling in speech recognition,” IEEE Signal Processing Magazine, 2012.
[15] Y. Kim, “Convolutional neural networks for sentence classification,” in Proceedings of EMNLP, 2014.
[16] J. Brownlee, “Deep Learning for Natural Language Processing,” Ma- chine Learning Mastery, 2017.
[17] S. Hochreiter and J. Schmidhuber, “Long short-term memory,” Neural Computation, 1997.
[18] A. Joulin et al., “Bag of tricks for efficient text classification,” in Proceedings of EACL, 2017.
[19] H. Schu¨tze, C. D. Manning, and P. Raghavan, “Introduction to Informa- tion Retrieval,” Cambridge University Press, 2008.
[20] Research articles on sentiment analysis from IEEE Xplore Digital Library.
[21] Online resources and tutorials on transformer-based NLP models.
[22] S. Bird, E. Klein, and E. Loper, “Natural Language Processing with Python,” O’Reilly Media.
[23] F. Chollet, “Deep Learning with Python,” Manning Publications.
[24] Research papers on sentiment analysis from IEEE Xplore Digital Li- brary.
[25] Online documentation and tutorials on web scraping techniques.
[26] Recent studies on aspect-based sentiment analysis in e-commerce.



10.22214/IJRASET

INDEX COPERNICUS 45.98

IMPACT FACTOR: 7.129

TOGETHER WE REACH THE GOAL

IMPACT FACTOR: 7.429

INTERNATIONAL JOURNAL FOR RESEARCH

IN APPLIED SCIENCE & ENGINEERING TECHNOLOGY