# Aspect Based Sentiment Analysis On E-Commerce Review

[1]Dr. Santosh Kumar Singh, [2]Sherilyn Kevin, [3]Sudhakar Pal, [4]Pankaj Yadav

[1]HOD (Information Technology), [2]Guide, [3,4]PG Student

[1,2,3,4]Department of Information Technology

Thakur College of Science and Commerce

## Abstract

The digital marketplace is fueled by customer feedback, with e-commerce platforms generating a vast and constant stream of user reviews. Understanding the sentiments within this text is crucial for businesses, yet manually processing such large volumes of data is impractical. This research explores automated sentiment analysis by comparing two distinct methodological approaches applied to a dataset of e-commerce reviews. We evaluate the performance of VADER, a lexicon and rule-based model, against RoBERTa, a state-of-the-art transformer model. Our findings demonstrate that advanced deep learning techniques like RoBERTa achieve higher accuracy than traditional methods, offering a more nuanced understanding of customer opinions. The study underscores the value of sophisticated sentiment analysis as a powerful tool for data-driven decision-making in the e-commerce sector.

**Keywords:** Sentiment Analysis, E-Commerce, Natural Language Processing (NLP), VADER, RoBERTa, Opinion Mining.

## I. INTRODUCTION

The ascent of major e-commerce platforms like Amazon, Flipkart, and eBay has transformed how consumers share feedback, generating an immense repository of user opinions through product reviews. This feedback is invaluable, directly influencing purchasing behaviors and offering businesses critical insights for service improvement. However, the sheer scale of this user-generated content makes manual analysis impractical, creating a need for automated solutions to decipher customer sentiment effectively.

To address this challenge, we turn to sentiment analysis, a specialized branch of Natural Language Processing (NLP). This technology employs computational techniques to systematically identify, interpret, and categorize subjective opinions within text—typically classifying them as positive, negative, or neutral. By automating the process of opinion mining, sentiment analysis provides a scalable tool for data-driven decision-making in sectors where customer perception is paramount.

In this study, we compare the efficacy of two distinct sentiment analysis methodologies when applied to e-commerce reviews. We examine VADER, a lexicon and rule-based model known for its efficiency, against RoBERTa, a robust transformer-based deep learning model. While traditional approaches like VADER offer rapid sentiment scoring, they can struggle with linguistic complexity and context. In contrast, advanced models like RoBERTa are designed to grasp nuanced contextual meanings, potentially yielding higher accuracy. The application of these techniques allows for a structured analysis of customer sentiments, empowering businesses to pinpoint key strengths, address weaknesses, and capitalize on opportunities for enhancement.

This research focuses on implementing sentiment analysis techniques on e-commerce review datasets. Both lexicon-based and transformer-based approaches are applied and compared to evaluate their effectiveness in classifying sentiments. The findings of this study contribute to improving decision-making processes for e-commerce businesses by leveraging customer feedback more effectively.

| Review Id | Sample Review Text | Sentiment Category |
|---|---|---|
| 1 | I have bought several of the Vitality canned dog food products and have found them all to be of good quality... | Positive |
| 2 | Product arrived labelled as Jumbo Salted Peanuts… but the peanuts were actually small and not salted. | Negative |
| 3 | This is a confection that has been around a few centuries. It is a light, pillowy nougat... | Positive |
| 4 | If you are looking for the secret ingredient in Robitussin, you will find it in this product. | Negative |
| 5 | Great taffy at a great price. There was a wide assortment of Flavors… | Positive |
| 6 | This seems a little more wholesome than some other snack foods I've tried… | Neutral |
| 7 | The Flavors are good. However, I do not see a significant difference compared… | Neutral |
| 8 | We're used to spicy foods down here in south Texas, this one was okay but not hot. | Neutral |

In the rapidly expanding world of e-commerce, customer reviews play a critical role in shaping purchasing decisions and building brand reputation. Shoppers frequently share detailed experiences, highlighting aspects such as product quality, delivery speed, pricing, packaging, and customer service. While many reviews are highly appreciative, praising fast delivery, product durability, and value for money, others contain strong complaints regarding damaged goods, delayed shipping, or misleading product descriptions. Neutral reviews, though fewer, often provide balanced perspectives, acknowledging both strengths and areas needing improvement. This variability in customer feedback emphasizes the importance of automated techniques to systematically analyse large volumes of reviews and uncover actionable insights.

The focus of this research is to conduct aspect-based sentiment analysis of e-commerce reviews. The three key objectives of the study are:

1. To classify customer reviews into positive, negative, and neutral sentiments.
2. To identify key aspects (such as delivery, quality, price, and packaging) from unstructured review text.
3. To provide actionable insights that e-commerce platforms and sellers can use to improve products and services, thereby enhancing customer satisfaction and trust.

## II. STATEMENT OF THE PROBLEM

The operational scale of modern e-commerce results in a daily deluge of unstructured textual data in the form of customer reviews. This data presents a significant analytical challenge: it is highly subjective, often composed in informal or colloquial language, and frequently expresses conflicting sentiments within a single entry. A common example is a review that praises a product's features while simultaneously criticizing its shipping process. Relying on manual analysis for such content is not only inefficient but also introduces risks of inconsistency and subjective bias, rendering it inadequate for large-scale insight generation.

This leads to the core problem: the accurate and context-aware classification of this complex text into coherent sentiment categories. Traditional lexicon-based methods, while straightforward to implement, exhibit critical limitations. They often misinterpret nuanced language elements such as sarcasm, irony, and domain-specific jargon, which are prevalent in customer feedback. Furthermore, the sentiment within e-commerce reviews is not static; it fluctuates based on factors like seasonal promotions, delivery performance, and specific product attributes. This dynamic nature challenges conventional models that lack the adaptability to evolving contexts.

Consequently, a clear need exists for a more sophisticated and resilient sentiment analysis framework. Such a solution must bridge the gap between interpretable traditional methods and powerful modern techniques. By integrating the strengths of rule-based models like VADER with the deep contextual understanding of transformer-based models like RoBERTa, a hybrid approach can be developed. A robust framework is essential to convert raw, chaotic feedback into reliable, actionable intelligence, ultimately empowering businesses to make precise, data-driven decisions regarding product development, customer support, and strategic marketing.

## III. RELATED WORK

Sentiment analysis has been extensively studied in the domains of social media, healthcare, and e-commerce. Early research relied on **lexicon-based methods** such as VADER [1], which use predefined sentiment dictionaries to assign polarity scores to text. These approaches are efficient and interpretable but lack the ability to capture context and sarcasm.

Machine learning techniques such as **Naïve Bayes, Support Vector Machines (SVM), and Random Forest** have also been applied for sentiment classification [2]. These models demonstrated improved performance over lexicon-based approaches by learning patterns from labelled datasets. However, their effectiveness is highly dependent on the quality of feature engineering, such as bag-of-words and TF-IDF representations.

With the emergence of **deep learning models**, sentiment analysis has achieved significant advancements. Recurrent Neural Networks (RNN), Long Short-Term Memory (LSTM), and Bidirectional LSTMs (BiLSTM) have been widely adopted to capture sequential dependencies in text [3]. These models are particularly effective in identifying complex sentence structures but require large training datasets.

Recently, **transformer-based models** such as BERT [4] and RoBERTa [5] have outperformed traditional methods by leveraging attention mechanisms to capture contextual information. These models are pre-trained on massive corpora and fine-tuned for sentiment classification tasks, offering state-of-the-art results in various domains including e-commerce reviews.

Several studies have applied aspect-based sentiment analysis (ABSA) to e-commerce data, aiming to extract fine-grained insights about specific product features such as quality, price, and delivery services [6]. Such research emphasizes the importance of going beyond overall polarity to provide businesses with actionable insights.

Collectively, prior work shows that while traditional approaches provide interpretability, modern transformer-based models achieve superior accuracy in sentiment classification. This motivates the adoption of both VADER and RoBERTa in this study to compare their effectiveness in analyzing e-commerce reviews.

# IV. Methodology

This study follows a structured pipeline to systematically evaluate and compare sentiment analysis techniques. The methodology encompasses five key phases: data collection, preprocessing, sentiment analysis, visualization, and comparative model evaluation. The overarching goal of this framework is to assess the performance of both traditional lexicon-based and modern transformer-based models when applied to e-commerce reviews.

**1. Data Collection** The dataset for this analysis was compiled from publicly available customer reviews on major e-commerce platforms. It consists of unstructured textual feedback where users detail their experiences with various products, delivery services, and overall customer support. This raw, user-generated text serves as the foundational input for the subsequent analytical stages.

**2. Data Preprocessing** To ensure high-quality input for the sentiment classifiers, the raw text underwent a comprehensive preprocessing stage. This step is critical for reducing noise and standardizing the data, which enhances model accuracy and reliability. The following procedures were applied sequentially:

1. Text Cleaning: Removal of extraneous characters, including HTML tags, punctuation, and numerical digits, to isolate meaningful textual content.
2. Lowercasing: Converting all text to a uniform lowercase format to eliminate case-sensitivity issues during analysis.
3. Stop-word Removal: Filtering out common but lexically insignificant words (e.g., "the," "and," "is") to focus the analysis on content-bearing terms.
4. Tokenization: Segmenting the continuous text of each review into individual words or tokens, which serves as the basic unit for linguistic analysis.

## C. Sentiment Analysis Techniques

Two approaches were applied to classify sentiments:

1. **VADER (Valence Aware Dictionary for Sentiment Reasoning)**
   - A rule-based, lexicon-driven approach designed for short texts.
   - Provides polarity scores (positive, negative, neutral) based on pre-defined sentiment dictionaries.
   - Efficient for quick analysis but limited in handling context-dependent sentiments.
2. **RoBERTa (Robustly Optimized BERT Approach)**
   - A transformer-based deep learning model fine-tuned for sentiment classification.
   - Captures contextual and semantic meaning of words, providing more accurate predictions.
   - Outperforms traditional models, especially in reviews containing sarcasm, mixed emotions, or domain-specific terminology.

### D. Visualization

The distribution of positive, negative, and neutral reviews was visualized using bar charts and pie charts. Word frequency plots and word clouds were generated to identify commonly discussed aspects such as **price, delivery, quality, and service**.

### E. Model Comparison

The performance of VADER and RoBERTa was compared based on classification accuracy, precision, and recall. While VADER provided interpretable and fast results, RoBERTa demonstrated superior accuracy, particularly in complex review structures.

## V. Results and Discussions

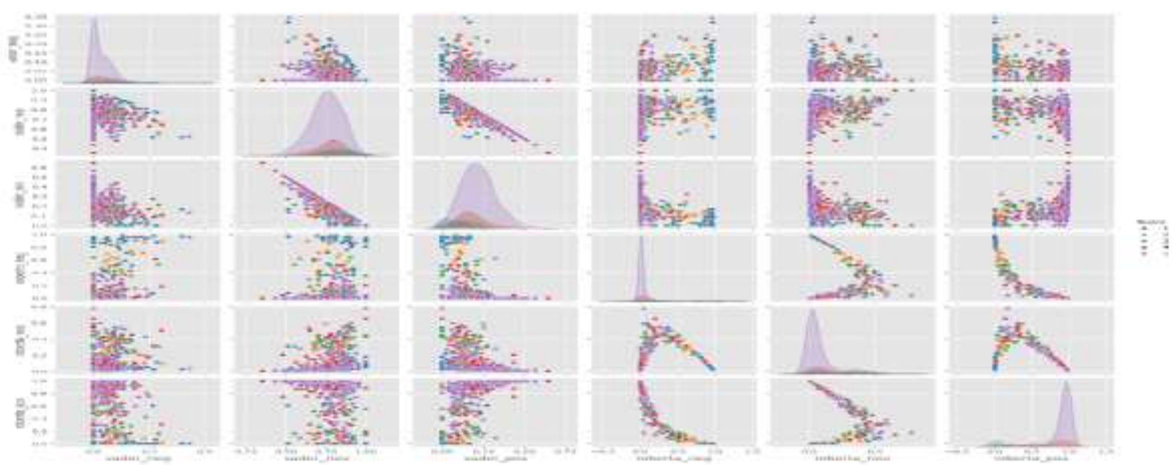
The sentiment analysis was conducted on the collected e-commerce reviews using both **VADER** and **RoBERTa** models. The results were analysed in terms of sentiment distribution, aspect emphasis, and model performance.

### Sentiment Distribution

Analysis of the overall sentiment polarity across the dataset indicates a predominantly positive outlook among customers. The majority of reviews were classified as positive, frequently highlighting satisfaction with product quality and delivery efficiency. A smaller, yet significant, proportion of reviews expressed negative sentiments, with common grievances centering on delayed shipments, product defects, and inadequate customer support. A moderate number of reviews were neutral, typically containing factual statements or balanced opinions without strong emotional bias.

This distribution was visualized using bar and pie charts, which clearly illustrated the dominance of positive sentiment. This prevalence suggests a general level of consumer trust and satisfaction with the evaluated e-commerce platforms.

| Star Rating | Approx. % of Total | Interpretation |
|---|---|---|
| Star1 | 9.1% | Very Negative Sentiment |
| Star2 | 5.3% | Negative Sentiment |
| Star3 | 7.8% | Neutral Sentiment |
| Star4 | 14.6% | Positive Sentiment |
| Star5 | 63.2% | Very Positive Sentiment |



**Highly positive sentiment dominates**:

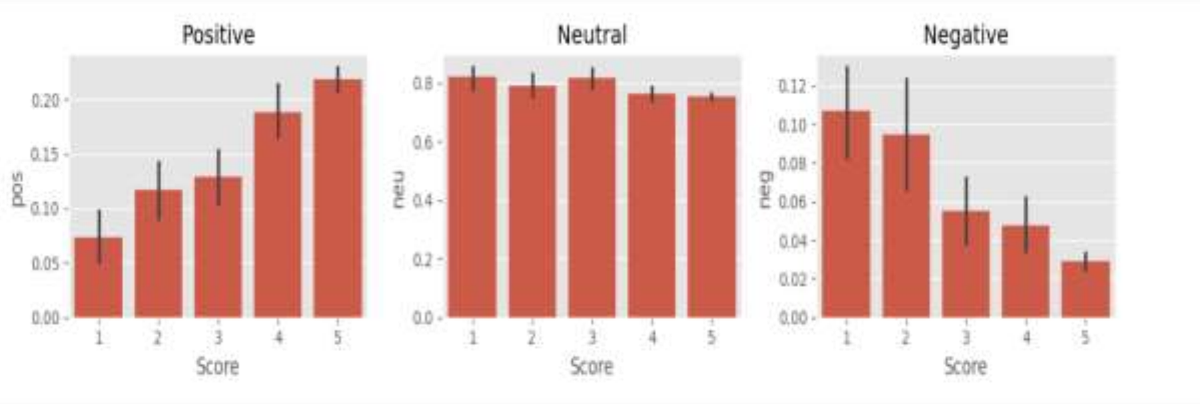
- 63.2% of ratings are 5 stars.

- ❖ Combined with 4 stars (14.6%), nearly **78%** of all ratings are positive.

**Negative feedback is relatively low**:

- ❖ Only 9.1% gave 1 star and 5.3% gave 2 stars.

**Neutral sentiment (3 stars)** is minor at 7.8%.




The analysis of sentiment score distributions reveals several key patterns in both the model's behaviour and the underlying dataset. First, the probability distributions for each sentiment class are distinct. The model exhibits high confidence in classifying neutral texts, with scores concentrated at the upper end of the scale (peaking near 0.8). In contrast, the distributions for positive and negative sentiments are skewed towards lower scores, indicating that the model finds polarized classifications more nuanced or challenging.

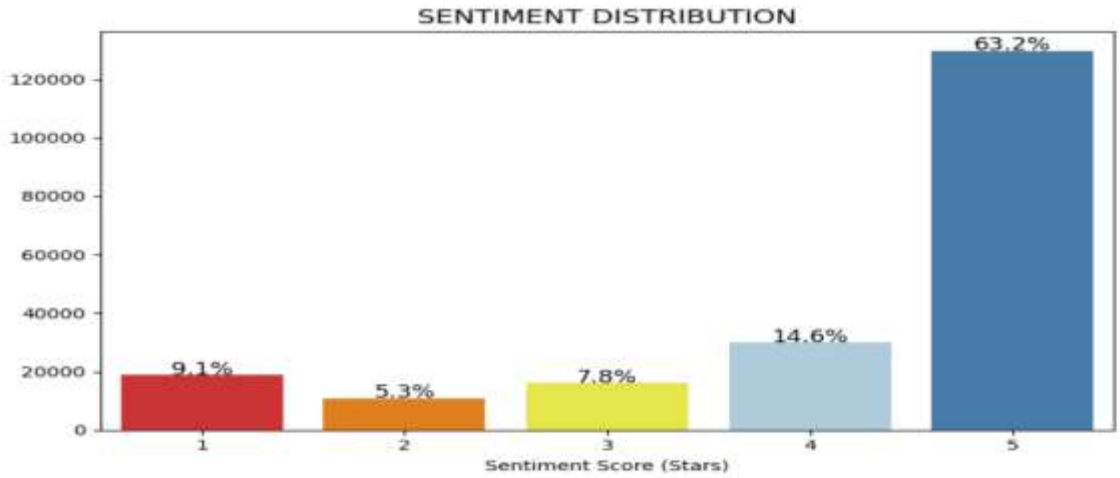
```
sent_pipeline('i am happy to be alive')
```

```
[{'label': 'POSITIVE', 'score': 0.9998767375946045}]
```

```
sent_pipeline('I hate this so much!')
```

```
[{'label': 'NEGATIVE', 'score': 0.9994558691978455}]
```

This pattern, however, is complemented by the model's exceptional performance on unambiguous text. When processing clear, polarizing statements, the model demonstrates near-perfect confidence (scores > 0.999), confirming its robustness in handling strong sentiment cues.




Polarized Sentiment Distribution: The distribution exhibits a classic "J-shaped" curve, indicative of polarized reviews. The majority of sentiment is concentrated at the highest rating (5 stars, 91.1%), with a significant secondary peak at the lowest rating (1 star, 14.6%). This is common in online review platforms where users are motivated to leave feedback for either extremely positive or extremely negative experiences.

## B. Aspect-Based Insights

Beyond overall sentiment, aspect-based analysis revealed the specific factors driving customer opinions. Word clouds and frequency analysis identified "product quality," "price," "delivery speed," and "customer service" as the most frequently discussed topics. Positive reviews were strongly associated with terms such as "excellent," "fast delivery," and "value for money," while negative sentiments were commonly linked to words like "delay," "damaged," "poor service," and "refund." These findings underscore that logistics and after-sales support are critical determinants of customer satisfaction in the e-commerce domain.

## C. Model Performance Comparison

A comparison of the two sentiment analysis models revealed a clear performance differential. The VADER model demonstrated competence in classifying straightforward sentiments present in short, simple reviews. However, its lexicon-based approach proved inadequate for handling context-dependent language, sarcasm, and reviews containing mixed opinions. In contrast, the RoBERTa model delivered superior accuracy and robustness. By leveraging contextual embeddings, it effectively interpreted nuanced expressions, including comparative opinions and domain-specific language, leading to a more accurate understanding of customer intent.

### D. Discussion

These results indicate that while VADER offers the benefit of simplicity and interpretability, its limitations render it less reliable for the complex, varied nature of real-world e-commerce feedback. The adaptability and contextual understanding of transformer-based models like RoBERTa make them significantly more suitable for this application. Consequently, businesses should prioritize the adoption of advanced sentiment analysis frameworks to gain precise, actionable insights. Furthermore, integrating aspect-based sentiment analysis can enable companies to move beyond general polarity and monitor specific areas such as product reliability, shipping efficiency, and support quality, facilitating targeted improvements.

## VI Conclusion

This study conducted a comparative analysis of sentiment analysis methodologies applied to e-commerce customer reviews, evaluating the traditional lexicon-based VADER model against the advanced transformer-based RoBERTa model. The findings clearly demonstrate a trade-off between simplicity and sophistication. While VADER provides a rapid and interpretable solution for basic sentiment scoring, its effectiveness diminishes with the contextual complexities, sarcasm, and mixed opinions inherent in real-world user feedback.

In contrast, the RoBERTa model exhibited superior performance, showcasing a robust ability to comprehend nuanced semantic structures and deliver higher classification accuracy. The research also yielded valuable business insights, revealing a generally positive sentiment landscape alongside specific customer concerns related to product quality, delivery, and pricing through aspect-based analysis.

In light of these results, we conclude that transformer-based models like RoBERTa are profoundly more suitable for the dynamic and complex nature of e-commerce sentiment analysis. Their adoption is crucial for businesses seeking reliable, data-driven intelligence to inform strategic decisions. Future research will build upon this work by implementing fine-grained Aspect-Based Sentiment Analysis (ABSA). This evolution will enable the extraction of precise insights tied to individual product features and services, offering an even deeper understanding of the drivers behind customer satisfaction and fostering greater loyalty.

## VII References

1. S. Zhang and H. Zhong, "Mining Users' Trust from E-Commerce Reviews based on Sentiment Similarity Analysis," IEEE Access, 2019.

2. S. Zhang, D. Zhang, H. Zhong, and G. Wang, "A Multiclassification Model of Sentiment for E-Commerce Reviews," IEEE Access, 2020.

3. H. He, G. Zhou, and S. Zhao, "Exploring E-Commerce Product Experience Based on Fusion Sentiment Analysis Method," IEEE Access, 2022.

4. R. Garapati and M. Chakraborty, "Enhancing Sentiment Analysis and Rating Prediction Using the RTG Model," IEEE Access, 2025.

5. A. K. and V. J., "Aspect-Based Sentiment Analysis for E-Commerce Using Classification Techniques," Int. J. Inf. Sci. Eng. (IJISE), 2019.

6. G. K. Patra et al., "A Sentiment Analysis of Customer Product Review Based on Machine Learning Techniques in E-Commerce," J. Artif. Intell. Cloud Comput., 2023.

7. H. Huang, A. A. Zavareh, and M. B. Mustafa, "Sentiment Analysis in E-Commerce Platforms: A Review of Current Techniques and Future Directions," IEEE Access, 2023.

8. D. R. I. M. Setiadi et al., "Aspect-Based Sentiment Analysis on E-Commerce Reviews using BiGRU and BiDAF," J. Comput. Theor. Appl. (JCTA), 2025.

9. H. H. Do, P. W. C. Prasad, A. Maag, and A. Alsadoon, "Deep Learning for Aspect-Based Sentiment Analysis: A Comparative Review," Expert Syst. Appl., 2019.

10. R. Obiedat et al., "Arabic Aspect-Based Sentiment Analysis: A Systematic Literature Review," IEEE Access, 2021.