Journal of Computer Science

Original Research Paper

Diverse Trajectory on Consumer's Product Feedback Analysis by Confidence Interval

[1]Vithya Ganesan, [2]Kirubakaran Namaskaram, [1]Viriyala Sri Anima Padmini, [3]Subrata Chowdhury and [4]Hariharan Shanmugasundaram

[1]*Department of Computer Science and Engineering, Koneru Lakshmiah Education Foundation Andhra Pradesh, India*
[2]*Department of Computer Science and Engineering, Chennai Institute of Technology, Chennai, India*
[3]*Department of Computer Science and Engineering, Sreenivasa Institute of Technology and Management Studies, Chittoor Andra Pradesh, India*
[4]*Department of Artificial Intelligence and Data Science, Vardhaman College of Engineering, India*

Article history
Received: 09-12-2024
Revised: 21-01-2025
Accepted: 31-01-2025

Corresponding Author:
Hariharan Shanmugasundaram
Department of Artificial Intelligence and Data Science, Vardhaman College of Engineering, India
Email: mailtos.hariharan@gmail.com

Abstract: The current investigation addresses the possibilities and limitations of the sentiment analysis of consumer remarks in the context of economic needs. It makes it easier for businesses to comprehend how the public's opinion shapes their abilities to satisfy clientele. Sentiment analysis concerns the assessment and interpretation of various forms of evaluation, such as product ratings, reviewer responses, and question-and-answer sessions. Thus, it exposes attitudes, trends and emotions concerning products and services. Commercialization has gained a lot of momentum over the years, and platforms like amazon.com form an important basis for customer feedback for organizations that have a lot of content to analyze. In the context of this research, consumer reviews and Q and A sections are collected through web scraping, and the latter is studied using the BERT neural network. Owing to its sophisticated natural language processing capabilities, BERT renders an accurate analysis of text sentiment as it considers even the slightest variation in the meaning of words. It does not stop at simply tagging the reviews as positive, negative or neutral; rather, it looks for sentiments and related themes that may affect consumers' attitudes and behaviours. In order to complement these findings, confidence intervals are introduced to allow for the handicapping of customer opinion averages and the confidence attached to predicting the Sentiment. This merging of BERT with confidence levels in the BERT predictions helps to clear any grey areas regarding the comments given on the product, thus providing comprehensive coverage of how the product was received and customer satisfaction. The results of the study show that this twofold strategy assists companies in better comprehending customer reviews that, in turn, shape their tactical plans, especially in marketing, product enhancement and customer care activities. This approach reveals customer feelings more explicitly and enables companies to take a proactive stance towards addressing customers' needs, thus enhancing customers' satisfaction and the product's success.

Keywords: BERT, Sentiment Analysis, Web-Scrapping, Consumer Feedback, Confidence Interval on Consumer Feedback

Introduction

Sentiment analysis is a subfield of Natural Language Processing (NLP) that focuses on identifying and categorizing opinions expressed in text, particularly to determine whether the Sentiment is positive, negative or neutral. This process is crucial for understanding customer feedback in various domains, such as e-commerce, service industries and product development. By extracting the Sentiment from customer reviews, organizations can gain insights into consumer satisfaction, product quality and overall experience. Given the rapid growth of online reviews and consumer-generated content, sentiment analysis has become an essential tool for businesses to monitor brand reputation and customer preferences in real-time. This study aims to explore how sentiment analysis can be improved using BERT models and confidence intervals to ensure accuracy and reliability in consumer feedback analysis.

Importance of Consumer Feedback and Current Techniques in Business

Consumer feedback plays a vital role in shaping product strategies, marketing decisions and customer service improvements. In the digital age, platforms like Amazon, Yelp and TripAdvisor are filled with customer reviews, questions and answers. Analyzing this vast data can provide businesses with valuable insights into what customers think about their products or services. This feedback loop is critical for continuous improvement and customer retention. For example, companies like Apple monitor product reviews to understand the public's response to new releases, allowing them to make timely adjustments to product features and services. Sentiment analysis helps quantify these opinions, allowing for data-driven decisions.

Traditionally, sentiment analysis has relied on basic text analysis techniques, such as bag-of-words models, where the frequency of words associated with positive or negative sentiments is counted. However, these models often struggle with understanding the context, sarcasm and nuanced opinions. Recent advancements in machine learning, such as the Bidirectional Encoder Representations from Transformers (BERT), have revolutionized sentiment analysis by improving the understanding of the context in which sentiments are expressed. BERT models are pre-trained on a large corpus of text, allowing them to capture intricate relationships between words and their meanings.

Challenges and Research Gaps

Despite such developments, many unresolved issues exist with regard to the sentiment analysis of texts. The most important of them is how to handle mixed or conflicting sentiments in one review. For most customers, it happens when they share opinions that are a mixture of both negative and positive statements regarding various attributes of the same product or service. For example, a review might read, "The product quality is excellent, but the delivery was extremely delayed." The feedback here is positive in terms of the product itself, but it shows dissatisfaction with the delivery experience. The traditional models for sentiment analysis classify feedback as positive, negative or neutral and, therefore, fail to capture such nuances in conclusions.

Another major challenge deals with the nature of human language as being quite complicated. Most of the time, sarcasm, idiomatic expressions, and slang distort the exactness of a sentiment analysis model. For example, the phrase "Oh great, another update that broke the app!" may well be perceived in a positive context by those models that lack the sense to analyze sarcasm. Similarly, in the context of domain-specific terminologies, such as technical jargon in software reviews or medical terms in healthcare feedback, general-purpose models are inherently unsuitable. These subtleties require more advanced natural language processing techniques that understand the nuances of context.

A critical limitation of the existing sentiment analysis models is that they do not quantify uncertainty in their predictions. Although extensive research has been done to enhance classification accuracy, the importance of assessing confidence in the predictions has not been addressed well. In ambiguous or borderline feedback situations, it is important to indicate how reliable the sentiment prediction is. Without such clarity, businesses may make incorrect decisions based on potentially unreliable insights.

These gaps hold profound practical implications, especially for businesses whose strategies are decided based on sentiment analysis. To illustrate, if a company is developing its product roadmap based on customer feedback, ambiguity in such sentiments might lead to confusion about confidence levels. This could result in misaligned priorities, wasted resources, or ineffective strategies. This has significant risks to sentiment analysis when confidence metrics are absent; thus, a more robust approach needs to be put in place.

Addressing the Gaps

The novel framework introduced here integrates BERT models with confidence intervals to solve these problems. By incorporating confidence intervals, the proposed approach not only returns sentiment predictions but also measures the reliability or uncertainty of each prediction. This is valuable in cases where feedback is ambiguous, and the sentiments are not clearly defined. Confidence intervals will enable businesses to understand the robustness of the results in sentiment analysis and improve their usability for decision-making.

The advanced contextual understanding capabilities of BERT make the Model particularly fit for mixed-sentiment analyses. Identifying sentiments attached to specific aspects within a review about product quality or delivery experience means that this methodology offers more of a granular and detailed review of customer responses. Adding a confidence measure also increases overall interpretability. Businesses can give the highest priority to such high-confidence prediction results, going ahead and carefully analyzing low-confidence output results in subsequent analysis stages.

The practical implications of this enhanced approach are more relevant in those industries where sentiment analysis directly affects high-stakes decisions. For example, in retail, it can help to identify specific concerns that customers have about product features. In healthcare, it may provide clearer insights into patient feedback on treatments, while in finance, it can analyze investor sentiment with confidence levels. Addressing these challenges ensures the framework provides more reliable and actionable insights across various domains.

This study contributes significantly to the field of sentiment analysis by shifting the focus from improving classification accuracy to quantifying the uncertainty in predictions. This not only bridges the gap between theoretical advancements and real-world applicability but also provides businesses with tools to make more informed decisions. By integrating BERT models with confidence intervals, the study offers a practical solution to enhance both the accuracy and trustworthiness of sentiment analysis.

Literature Survey

E-commerce has recently experienced rapid growth. As a result, online purchasing has increased, which has resulted in an increase in online product reviews. Because the customer's opinion about the product is influenced by other consumers' recommendations or complaints, the implied opinions in customer reviews have a massive influence on the customer's purchasing decision. This study analyses the amazon reviews dataset and investigates sentiment classification using questions and answers and reviews (Acheampong *et al.*, 2021). Sentiment analysis, a tool for identifying customer sentiment, marketing initiatives, and product appraisals (Bellar *et al.*, 2023).

E-commerce remark sentiment analysis has become a research focus. The text context is not included in the current word vector representation paradigm. Xie *et al*. (2020) As a result, the SA-BERT pre-training language model based on transformer bidirectional encoder representation is proposed in this study. Bai (2011) Bert first encodes the word vector to reflect the semantic information of the remark text's context. The attention method is then utilized to extract text characteristics at a deeper level, grasp the semantics of text information and finish the e-commerce comment sentiment analysis assignment (Cao *et al*., 2011).

In this study, scraped the reviews, questions, and answers of the product from the Amazon website because, based on the customers' opinions, only the Sentiment of the product can be analyzed (Liu *et al*., 2020). This web scraping can be achieved using the Beautiful Soup package, which pulls out the required data from all HTML or XML files (Glez-Peña *et al.*, 2014). Scraped data reviews on a particular product, questions on that product, and answers to those questions and reviews (AlQahtani, 2021). Filter the data from Web-scraping for sentimental analysis. Geetha and Karthika Renuka (2021). The "nlptown/bert-base-multilingual-uncased-sentiment" model is a multilingual BERT model that analyzes Sentiment across languages, which is ideal for applications like customer feedback analysis and review classification (Maryam and Adel, 2024). This study reviews sentiment analysis techniques applied to customer feedback, emphasizing how sentiment insights enhance decision-making. It discusses key methods, challenges for brand ranking (Chen, 2022).

This study explores using Natural Language Processing (NLP) to analyze e-commerce reviews, providing insights into consumer sentiment. It highlights methodologies and implications for businesses (Kumar and Smith, 2022). The paper examines various machine learning techniques for analyzing customer feedback in e-commerce settings, focusing on their effectiveness in enhancing user experience and business strategies (Le *et al.*, 2024). This study proposes a deep learning approach for classifying product reviews and enhancing sentiment analysis accuracy. Wankhade explores consumer behaviour methods from e-commerce that highlight insights derived from analyzing customer reviews and feedback patterns (Wankhade *et al*., 2022).

Pinky Aktar, uses sentiment analysis in e-commerce to capture customer sentiments to improve business strategies by machine learning (Akter *et al*.,2025). Vaissnave explores convolutional inception techniques to e-commerce reviews to extract valuable insights, aiming to enhance marketing strategies. The study emphasizes leveraging customer sentiment for targeted marketing approaches and improved consumer engagement (Vaissnave *et al*., 2025). Singh, U analyze customer sentiment in mobile commerce applications, focusing on methods for capturing and interpreting user feedback. Their research aims to enhance user experience and inform business strategies through effective sentiment analysis techniques tailored for mobile platforms (Singh *et al*., 2022). Lan, S explores techniques for sentiment analysis in consumer electronics reviews, focusing on methods to extract valuable insights from customer feedback and enhance by mapping relationship between user feedback and production characteristics (lan *et al*., 2025).

Zhou explores on sentiment analysis of Cross-border e-commerce questions and answers, reviews of a product, and visualization with their sentimental results using consumer preferences for comparing consumer/customer uncertainty and product legitimacy. Zhou *et al*. (2025). From the above survey, the gap is identified on semantic review analyzed with confidence interval to identify the coordinates between questions and feedback. So data sets are used to compare the voices of customers and it is explained as follows.

Data Set

In the dataset, labels like questions, answers, sentiment-question, sentiment-answers, reviews-title, sentiment-titles, reviews, sentiment-reviews and it is shown in Table (1):

1. Questions
2. Answers

DOI: 10.3844/jcssp.2025.1129.1139

3. Sentiment-question
4. Sentiment-answers
5. Reviews-title
6. Sentiment-titles
7. Reviews
8. Sentiment-reviews

Inquiry approach: Questions (1-3, 5, 10) are excessively intriguing in certain qualities of a product, such as its specifications and features. Users, on the other hand, are keen on the accessories of the phone, its RAM, the quality of its camera and some design features such as the notch and battery capacity. User Needs and Preferences: Statements (4, 6-9) mostly depict observations or intentions to purchase the devices. For example, in statement 4, the respondent self-reported the camera quality as being satisfactory. In statements 6 and 7, this is further illustrated with a pending purchase and the mention of an older unit, which implies that a certain brand is preferred (iPhone or Samsung). In statement 8, there is a suggestion that the person is either looking for 5G devices or 5G functionality is a must for them. In statement 9, colour and model are likely based on the user's personal preference.

Information gaps: Query 2 about RAM, Query 3 about camera information and Query 10 about battery number are the specific criteria for which users require definitive responses; hence, the expectation of such details in primary product descriptions could be tricky. Decision-Making influencers: In the case of Statements 6 and 7, where the decision is which brand or Model to opt for, this could be determined by the years of existence of a certain brand (Samsung Note 8) or the highlighted features like 5G, RAM and battery.

Labels:

- Questions: This label indicates that all the preprocessed, scraped questions data on the particular product
- Answers: This label indicates that all the preprocessed, scraped answers data on the particular product against the questions
- Sentiment-question: This label indicates after performing sentimental analysis on those question data, it gave us a rating from 1-5, which indicates poor, unsatisfactory, satisfactory, good, outstanding
- Sentiment-answers: This label indicates after performing sentimental analysis on those question data, it gave us a rating from 1-5, which indicates poor, unsatisfactory, satisfactory, good, outstanding
- Reviews-title: This indicates the title of reviews on that product which is given by the customer
- Sentiment-titles: This label tells us the rating on review titles where we perform sentimental analysis on review titles
- Reviews: This gives us details view of the full review of the customer on the product
- Sentiment-reviews: It shows us the sentimental analysis result on the reviews data

Data Preprocessing

The dataset utilized had to be preprocessed before actual use to make sure that the operation was completed as effectively as feasible. So, first we have preprocessed the data of questions on the product shown in Table (1). Data curation is shown on Table (2).

The inquiries indicate a mixture of product-specific questions and general comments that encapsulate the scope as well as the user's priorities and existing information deficits. The queries are directed at the technical specifications (accessories, RAM, camera, notch, battery etc.) which reveal the need to obtain more in-depth information on the various features and functionalities of the device (Queries 1-3, 5 and 10). These questions imply that users have some concerns as to what comes with the device and its functionality.

Table 1: Data for sentiment analysis

ID	Question	Answer	Sent. Ques.	Sent. Ans.	Title	Sent. Title	Review
0	What all accessories come in the box?	USB C to lightning cable and SIM tool	1	5	Don't Purchase the iPhone	1	This product is duplicate of iPhone 12, only camera design changed. Very low-quality product; buy Samsung mobile phone
1	In this Model, how much capacity of RAM?	6 GB	3	5	Smil	5	You should be a big bot to buy this phone; you could get iPhone 12 Pro, that's the best
2	What is a Camera?	12 MP!	1	1	Damaged Product Reviewed	1	There was a hairline scratch on the screen, not clearly visible. When we tried contacting Amazon, they blamed Apple, and vice versa
3	The camera is very clear?	Yes, best camera available	4	5	Excellent upgrade (but costly)	5	Awful. Just buy iPhone 12! Not worth the money! Unless you're corrupt, then go ahead
4	Does it still have a notch?	Yes	3	5	Trash AF	1	I am poor and can't afford it; I'm bored, so here's my review while drinking coffee... can't study, came here to write this review... (truncated)

Table 2: Questions after data preprocessing

Query	Question/Statement	Type
1	What all accessories come in the box with the phone?	Question
2	Is this Model, how much capacity of RAM?	Question
3	What is the Camera?	Question
4	Its camera is very clear	Question
5	Does it still have a notch	Question
6	Hi friends, I want to buy a mobile, but confused between iPhone 13 Pro Max and Samsung Galaxy S22 Ultra	Statement
7	Present I'm using Samsung Galaxy Note 8. Plz help.	Statement
8	This Phone 5G in	Statement
9	iPhone 13 Pro Max Sierra Blue 128	Statement
10	How many batteries?	Question

Table 3: Answers after data preprocessing

Query	Response	Type
1	Just the USB C to Lightning Cable and the SIM Removing Tool	Statement
2	6GB	Statement
3	12 MP!	Statement
4	Yes, best camera available	Statement
5	Yes	Statement
6	I have an iPhone, but I prefer Android over iOS. Simplicity is more important than security for me	Statement
7	Yes	Statement
8	Yes	Statement
9	1	Statement
10	Yes, one physical and one eSIM	Statement
11	Yes	Statement
12	Happiness followed by sadness	Statement
13	Exchange of old mobile is available in most of the pin codes... Even if you get the product as damaged (mostly it won't happen) they will replace the product after a clear check	Statement
14	Very nice	Statement
15	Yes, it supports	Statement
16	Yes	Statement
17	No, you have to cancel it/reject the delivery and place a new request/order for the 256 one!	Statement
18	Mobile and Charging cable	Statement
19	Cunnly chain iccusel	Statement(unclear)

Lastly, Query 11 ("This bhang in dual SIM?") does not translate well, although it seems to suggest an inquiry with regards to the dual SIM support, which is quite common in some regions. All in all, the data suggests that users are concerned with specifications as well as competitors and connectivity, which calls for making available information on the products very useful and easy to read.

Then preprocess the data of answers on that product which are given against to questions shown in Table (3).

This table gives an understanding of the customer regression and preference towards accessories, features and services for an iPhone model. Majority of the answers show that consumers are content with specific attributes such as camera resolution (Query 4) and eSIM availability (Query 10). Answers state that some features are turned down for instance the user's comfort with the Android operating system instead of Apple's iOS operating system (Query 6), hence showing different priorities for different customers.

Another recurring theme is the complaints made in connection with the service of exchange and even replacement which is respondent number thirteen (Query 13). It says that the customer is allowed to bring back the faulty item after inspection in return for the good quality one and therefore customers trust services especially on expensive items where smooth services are expected. However, respondents like "Cunnly chain iccusel" (Query 19) are non-informative and may suggest a communication barrier or differences in writing styles owing to possible error or ambiguity in expression. Generally, it can be stated that they like the major features of the product but there are also demands for enhancement in the services offered and the communication applied to the users to make their experience better.

Materials and Methods

Web-scrape the required data from the website (eg: Amazon website) and segment into reviews, questions and answers of a particular product. From this segmentation step, preprocess the data which is in the three segments and it is shown in Fig. (1).

The Hugging Face Transformers module is a collection of hundreds of models trained for a range of tasks and in a number of modalities, such as text, vision and audio. In particular, these models are very effective for performing various Natural Language Processing (NLP) tasks such as text classification, information retrieval, question answering, text summarization and machine translation engine that supports more than one hundred languages.

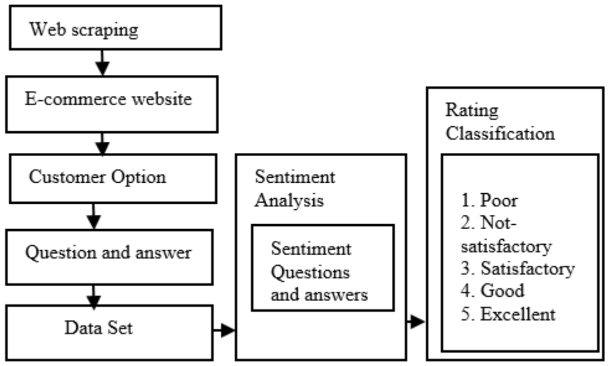

Fig. 1: Flow of customer feedback on product

Common examples of units present in the Transformers library are Auto Tokenizer which converts a string of text into numbers that can be fed into the models and Auto Model for Sequence Classification which imports pre-trained models that are suitable for classification purposes. All these tools enable patients easy to use the Model for any task from text input to the prediction of types or sentiments.

In addition, Secondly, there are other libraries such as Pytorch (torch), requests, re (which is regular expressions) and bs4 (notably, Beautiful Soup). These Libraries have delicate but important functions in the process. This library supports tensor computations and constructing deep neural networks trained on GPU. The requests library allows one to perform a web scraping technique by sending requests to web pages while re is used for the manipulation of text with the matching of certain patterns. This library belongs to the bs4 package and it is used to retrieve, infomincompleteneess data from HTML and XML documents. All these resources together create an efficient controlled environment for web scraping along with cleaning and processing data for the purpose of performing sentiment analysis making it possible to study answers and questions posted on sites like Amazon.

The torch (PyTorch) framework is a widely used deep learning framework that supports tensor manipulation, especially efficient processing in the GPU. It provides an intuitive interface for designing and training neural networks through what is referred to as a “tape”-which keeps track of the operations performed on the neural network towards automatic differentiation in backpropagation. PyTorch is preferred in most cases for its ability to aid in research development because developing complex systems in other frameworks is very easy. In this undertaking, Pytorch, once again, is used alongside requests and re this time where the former eases the process of HTTP requests such as web scraping or dealing with APIs and the latter deals with regular expression and cleaning of text.

Thirdly the requests library makes it possible to easily export and send HTTP requests in order to obtain the contents of a web page or an application interface, with an option built-in to handle data formats structured by JSON. On the other hand, the re (regular expressions) library has how example added parentheses alg interpret incvoroni inefficient ways to clean text and process it through finding and matching certain text regions in other texts. This especially applies to such cases as doing tokenization which is a pure NLP task or removing unnecessary data from the users` input which is basic cleaning. These aforementioned libraries complement each other during the phase when the data is received and prepared for some activities like sentiment scoring. After the data is cleaned, processed and the sentiment analysis is done, it would be made possible for the results to be presented in a manner depicting how the different questions alongside their respective responses were rated on a scale of 1-5.

Among many others, the Beautiful Soup package from the bs4 module persists as one of the popular technologies for scraping purposes as it is used basically to parse the HTML and XML documents. It enables the developers to navigate and search the layout of web pages by building a tree structure of the contents in the web page making it easier in the extraction of specific sections such as product Q and A in the case of e-commerce websites like Amazon. Beautiful Soup comes in full effect when there is a need to scrape data from html structures by specifying elements in the pages and getting the content, links or tags contained within those elements as required. This data may then be processed to remove any unwanted or redundant HTML tags within it to enhance readability and allow for efficient further manipulation of the data.

In Scraping of Amazon product’s Q and A section specifically, Beautiful Soup played a part in collecting the URLs, scraping their content and saving it in lists. Then those lists were transformed into pandas Data Frame for data arrangement and manipulations such as sentiment scoring. Sentimentzing was done by rating both the question and the answer on a scale of one to five according to customer satisfaction. After which, the data is presented in a way that illustrates how people question and answer each other with respect to a product. This example brings about the importance of web scraping in customer understanding and providing better products development.

Algorithm

Step 1: Import necessary libraries
Step 2: Collect Amazon product questions and answers
Step 3: Extract and clean questions and answers
Step 4: Create data frames for questions and answers
Step 5: Load the pretrained transformer model for sentiment analysis
Step 6: Assign sentiment ratings
Step 7: Count and analyze sentiment ratings

Step 8: Display the results

1. **Import Libraries:** Import pandas and a sentiment analysis tool (e.g., transformers or VADER)
2. **Data Collection:** Collect Amazon product questions and answers using the product ID
3. **Text Cleaning:** Define a function to clean and preprocess the collected text data
4. **Clean Questions:** Clean and preprocess the questions using the cleaning function
5. **Clean Answers:** Clean and preprocess the answers using the cleaning function
6. **Create DataFrames:** Create a DataFrame for the cleaned questions
7. **Create DataFrames (cont.):** Create a DataFrame for the cleaned answers
8. **Load Sentiment Model:** Load a pre-trained transformer model for sentiment analysis
9. **Apply Sentiment Analysis:** Apply the sentiment model to each question
10. **Assign Sentiment:** Store the sentiment scores for questions in the questions DataFrame
11. **Apply Sentiment Analysis (cont.):** Apply the sentiment model to each answer
12. **Assign Sentiment (con.):** Store the sentiment scores for answers in the answers DataFrame
13. **Count Sentiment Ratings:** Count the number of positive, negative, and neutral ratings in questions
14. **Count Sentiment Ratings (cont.):** Count the number of positive, negative, and neutral ratings in answers
15. **Summarize Results:** Summarize the sentiment analysis results for both questions and answers
16. **Display Questions Sentiment:** Print the Sentiment counts for questions
17. **Display Answers Sentiment:** Print the Sentiment counts for answers
18. **Data Validation:** Ensure there are no missing values in the data
19. **Save Results:** Optionally, save the sentiment results to a file or database
20. **End of Procedure:** Complete the analysis and prepare for next steps, if any

This pseudocode provides a structured overview of the algorithm in a straightforward format

The information contained in this table is related to the customer attitude and response towards an iPhone. Very often the feedback shows that they are upset with the product or services and this is reflected in the low score for both Sentiment and answers. As complaints where users experienced scratches on the screen also bother them with the customer care responders (Queries 2 and 5), the need for better service is clear… Negative feedbacks also make statements of a better deal being offered in earlier versions of the iPhone as compared to this one (Query 6), which indicates that other than providing the device, users would expect this latest device to be worth its price.

Positive appreciations, on the other hand, are limited to some features such as the camera quality (Query 3). This points out that some users may like certain features of the product but to the general sweep, the mood turns out to be unhappy owing to the unmet expectations concerning the construction width and the pricing. There are also elements of humor in some of the reviews (Query 8) which tend to suggest either a playful frustration or a serious one with respect to price and it is listed in Table (4).

Generally, the analysis of the feedback shows that there are users who love the product as well as those who do not love it, with the two extreme views being largely shaped by the cost of production and expected creativity and after-sale services.

Table 4: Dataset after sentiment analysis

Query	Question	Answer	Sentiment Question	Sentiment Answer	Review Title	Sentiment Title	Review Sentiment
1	What all accessories come in the box with the phone?	Just the USB C to lightning cable and the SIM removing tool	0	1	Do not purchase this iPhone	1	This product is a duplicate of iPhone 12, very low quality
2	Is this model, how much capacity of RAM?	6GB	3	5	Damaged product received, excellent upgradein context of being 90k poorer	4	Very dissatisfied, bad customer service
3	What is the camera?	12 MP!	1	1	It camera is very clear, yes best camera available	5	Very satisfied with the camera quality
4	Does it still have a notch?	Yes	4	5	You should be a bigggggg bot to buy this phone, you could get iPhone 12 Pro instead	1	Unhappy with the phone, recommending an older model

DOI: 10.3844/jcssp.2025.1129.1139

5	Yes	1	5	There was a hairline scratch on the screen, got stuck in a blame game between Amazon and Apple	1	Negative review due to poor customer service and product damage
6	Yes	1	5	Just buy iPhone 12, not worth the 5 money unless you're from BJP	3	Disappointed with the value of the phone
7	Yes	1	1	Trash af	1	Very negative review, harsh criticism
8	Yes	1	1	I am poor and can't afford it, so if you have money to flex, buy it, bored while writing	1	Negative, written inahumoroustone

Results and Discussion

The confidence interval is calculated for question and answer and reviews and it is shown in the following Tables (5-6). The information in the table aids the consumer's understanding of how much valuations of the product differ among answerers for each query that the questioner posed together with the mean score which indicates the generalized trends in Sentiment. In some cases, in which the mean and confidence interval scores are higher, (e.g., Question 4 with mean score 4.375), it implies that the respondents gave very good ratings. On the other hand, questions which had low mean ratings like question 1 invariably cut across several categories and exhibit more or less neutral or mixed attitude. The standard deviation of the means depicts the varying degree of uniformity of the responses which reflects the different levels of satisfaction concerning the features and the aspects that still require improvements.

The table analyzing sentiment review indicates an overall positive trend in the scores, with average ratings that range from 3.75-4.25 in every category which demonstrates overall satisfaction. The mean for category 4 is the highest in all the categories 4.25 and it has a standard deviation of 1.276 which shows that the responses were very positive. The other categories have slightly wider confidence intervals, showing variability, but they still exhibit positive Sentiment.

Figures (2-7): Shows visualization of sentiment analysis for questions and reviewer.

Table 5: Confidence interval of Sentiment question

Sentiment question	Mean	Count	Std Dev	CL95 Hi	CL95 Lo
1	3.111	45	1.761	3.626	2.597
2	3.000	12	1.758	3.995	2.005
3	3.805	41	1.520	4.270	3.340
4	4.375	8	0.916	5.010	3.740
5	3.457	35	1.669	4.010	2.904

Table 6: Confidence interval of sentiment reviews

Sentiment reviews	Mean	Count	Std Dev	Cl95 Hi	Cl95 Lo
1	3.783	23	1.704	4.479	3.086
2	3.750	12	1.603	4.657	2.843
3	4.000	11	1.549	4.916	3.084
4	4.250	40	1.276	4.645	3.855
5	3.891	55	1.487	4.284	3.498

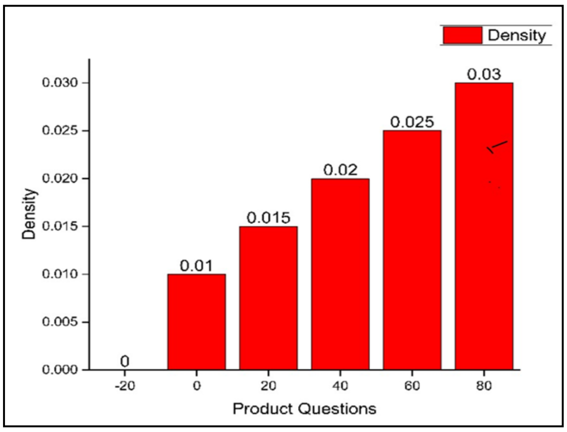

Fig. 2: Distribution function for product questions

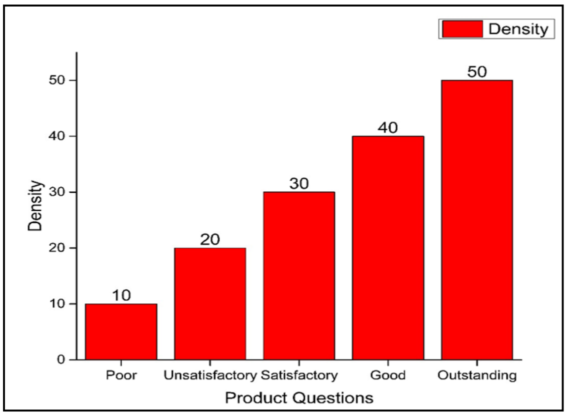

Fig. 3: Sentimental analysis of questions

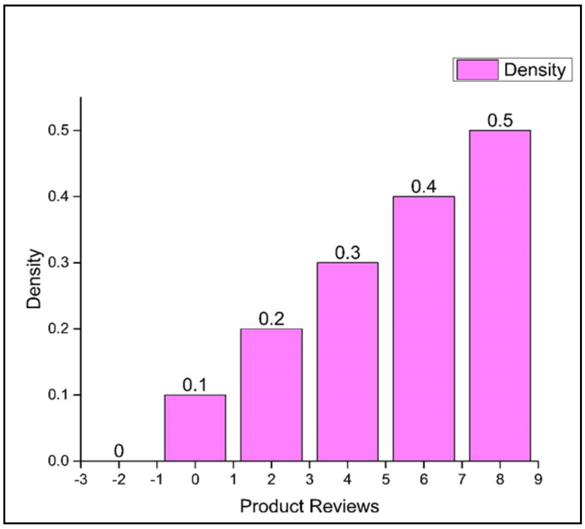

Fig. 4: Distribution function for product reviews

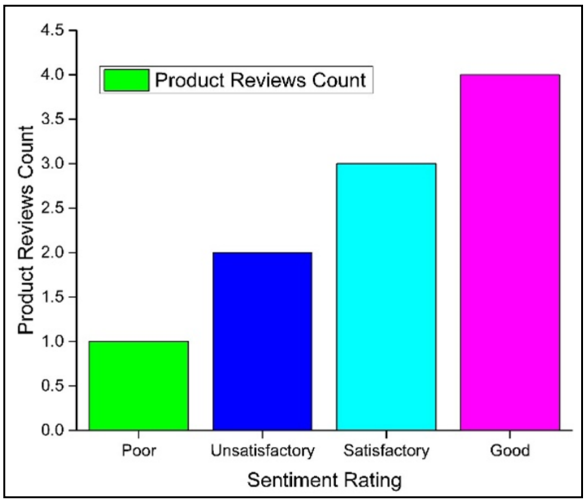

Fig. 5: Reviews of sentimental analysis

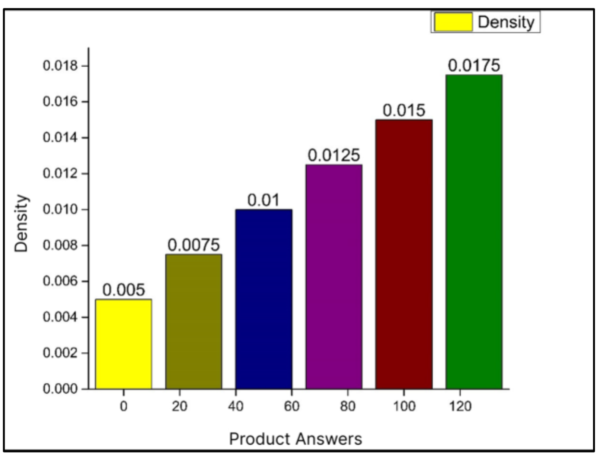

Fig. 6: Distribution analysis of product answers

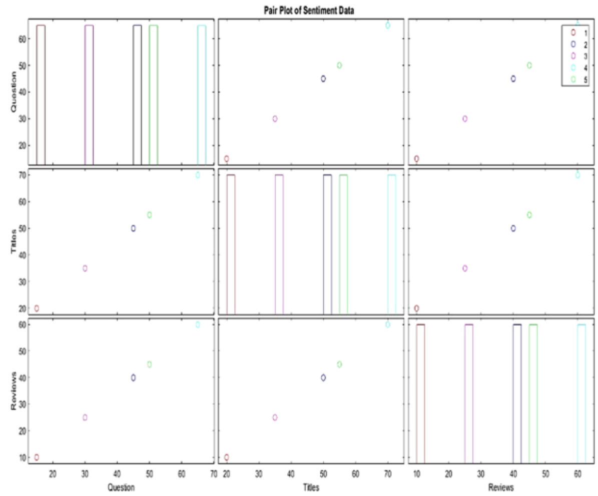

Fig. 7: Representation of data frame with pair plot

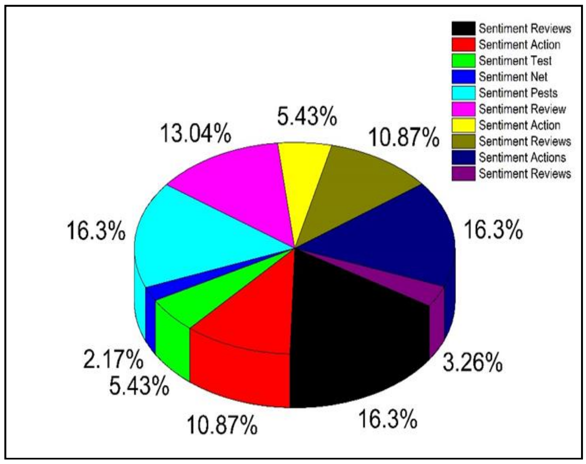

Fig. 8: Confidence-interval of product question and reviews

Figure (8): Shows the confidence interval between the Question answer and reviews of consumer/customer feedback on products. When % is high, it shows the legitimacy of consumer feedback. If it is less, it shows ambiguity on the consumer's mind and the product. Figure (8) shows subplot representation for sentiment analysis.

Conclusion

To conclude, the goal of the proposed research is to solve the semantic analysis of consumer/customer feedback which is based upon their opinions on the product. Web scrapped data has been taken for analysis to identify the authenticity of consumer /customer feedback. Web-scrapped data of the product iPhone13 Reviews, Questions and Answers section of that product is taken from the Amazon website for comparison and identify the trade-off between the reviews. A sentimental analysis using BERT to analyze and visualize. This study improves marketing and customer service and this strategy is helpful to policymakers and marketing planners.

Acknowledgment

Thank you to the publisher for their support in the publication of this research article. We are grateful for the resources and platform provided by the publisher, which have enabled us to share our findings with a wider audience. We appreciate the efforts of the editorial team in reviewing and editing our work, and we are thankful for the opportunity to contribute to the field of research through this publication.

Funding Information

The authors have no support or funding to report.

Author's Contributions

Vithya Ganesan: Critical révision of the article.
Kirubakaran Namaskaram: Design of the works.
Viriyala Sri Anima Padmini: Drafted the article.
Subrata Chowdhury: Data analysis and interpretation.
Hariharan Shanmugasundaram: Final version of the research article.

Ethics

No ethical issues that may arise after the publication of this manuscript.

References

Acheampong, F. A., Nunoo-Mensah, H., & Chen, W. (2021). Transformer Models for Text-Based Emotion Detection: A Review of BERT-Based Approaches. *Artificial Intelligence Review*, *54*(8), 5789–5829. https://doi.org/10.1007/s10462-021-09958-2

AlQahtani, S. M. A. (2021). Product Sentiment Analysis for Amazon Reviews. *International Journal of Computer Science and Information Technology*, *13*(3), 15–30. https://doi.org/10.5121/ijcsit.2021.13302

Bellar, O., Baina, A., & Ballafkih, M. (2024). Sentiment Analysis: Predicting Product Reviews for E-Commerce Recommendations Using Deep Learning and Transformers. *Mathematics*, *12*(15), 2403. https://doi.org/10.3390/math12152403

Bai, X. (2011). Predicting Consumer Sentiments from Online Text. *Decision Support Systems*, *50*(4), 732–742. https://doi.org/10.1016/j.dss.2010.08.024

Cao, Q., Duan, W., & Gan, Q. (2011). Exploring Determinants of Voting for the "Helpfulness" of Online User Reviews: A Text Mining Approach. *Decision Support Systems*, *50*(2), 511–521. https://doi.org/10.1016/j.dss.2010.11.009

Chen, N. (2022). E-Commerce Brand Ranking Algorithm Based on User Evaluation and Sentiment Analysis. Frontiers in Psychology, 13, 907818. https://doi.org/10.3389/fpsyg.2022.907818

Geetha, M. P., & Karthika Renuka, D. (2021). Improving the Performance of Aspect Based Sentiment Analysis Using Fine-Tuned Bert Base Uncased Model. *International Journal of Intelligent Networks*, *2*, 64–69. https://doi.org/10.1016/j.ijin.2021.06.005

Glez-Peña, D., Lourenço, A., López-Fernández, H., Reboiro-Jato, M., & Fdez-Riverola, F. (2014). Web Scraping Technologies in an API World. *Briefings in Bioinformatics*, *15*(5), 788–797. https://doi.org/10.1093/bib/bbt026

Le, H.-S., Do, T.-V. H., Nguyen, M. H., Tran, H.-A., Pham, T.-T. T., Nguyen, N. T., & Nguyen, V.-H. (2024). Predictive model for customer satisfaction analytics in E-commerce sector using machine learning and deep learning. *International Journal of Information Management Data Insights*, *4*(2), 100295. https://doi.org/10.1016/j.jjimei.2024.100295

Kumar, P., & Smith, L. (2022). E-Commerce Review Analysis Using NLP for Consumer Sentiment Insights. *IEEE Conference on Computational Intelligence*. IEEE Conference on Computational Intelligence. https://doi.org/10.1109/CIC59092.2022.10093674

Liu, Y., Lu, J., Yang, J., & Mao, F. (2020). Sentiment Analysis for E-Commerce Product Reviews by Deep Learning Model of Bert-BiGRU-Softmax. *Mathematical Biosciences and Engineering*, *17*(6), 7819–7837. https://doi.org/10.3934/mbe.2020398

Lan, S., Jiang, Y., Guo, T., Li, S., Yang, C., Edwin Cheng, T. C., Sethanan, K., & Tseng, M.-L. (2025). Personalized product design and user review and experience analysis: A data-driven hybrid novel approach. *Computers & Industrial Engineering*, *202*, 110939. https://doi.org/10.1016/j.cie.2025.110939

Maryam, A. A.-B., & Adel, I. A.-A. (2024). Sentiment Analysis on Customer Feedback for Improved Decision Making: A Literature Review. *2024 ASU International Conference in Emerging Technologies for Sustainability and Intelligent Systems (ICETSIS)*, 207–212. https://doi.org/10.1109/ICETSIS61505.2024.10459452

Akter, P., Hossain, S., Siddique, M. T., Ayub, M. I., Nath, A., Nath, P. C., Rasel, M., & Hassan, M. M. (2025). Sentiment Analysis of Consumer Feedback and Its Impact on Business Strategies by Machine Learning. *The American Journal of Applied Sciences*, *07*(01), 6–16. https://doi.org/10.37547/tajas/volume07issue01-02

Singh, U., Saraswat, A., Azad, H. K., Abhishek, K., & Shitharth, S. (2022). Towards Improving e-Commerce Customer Review Analysis for Sentiment Detection. *Scientific Reports*, *12*(1), 21983. https://doi.org/10.1038/s41598-022-26432-3

Wankhade, M., Rao, A. C. S., & Kulkarni, C. (2022). A Survey on Sentiment Analysis Methods, Applications, and Challenges. *Artificial Intelligence Review*, *55*(7), 5731–5780. https://doi.org/10.1007/s10462-022-10144-1

Xie, S., Cao, J., Wu, Z., Liu, K., Tao, X., & Xie, H. (2020). Sentiment Analysis of Chinese E-Commerce Reviews Based on BERT. *2020 IEEE 18th International Conference on Industrial Informatics (INDIN)*, 713–718. https://doi.org/10.1109/indin45582.2020.9442190

Vaissnave, V., Birunda, S. S., Dharani, V., Lalitha, R., & Muthamil Sudar, K. (2025). An Optimised Deep Learning Model with an Atrous Convolutional-Based Inception System for the Sentiment Analysis of Customer Online Reviews. *Journal of Engineering Design*, 1–26. https://doi.org/10.1080/09544828.2025.2475426

Zhou, R., Zhu, Y., Zheng, R., & Zhou, J. (2025). Differences of Consumer Reviews and Consumer Satisfaction on Cross-Border E-Commerce Platforms: A Text Mining Analysis Based on Cross-Cultural Perspective. *Journal of Organizational Computing and Electronic Commerce*, 1–23. https://doi.org/10.1080/10919392.2024.2447689