



# International Journal of Innovative Research in Science Engineering and Technology (IJIRSET)

*(A Monthly, Peer Reviewed, Refereed, Scholarly Indexed, Open Access Journal)*



**Impact Factor: 8.699**

**Volume 14, Issue 6, June 2025**

# Transformative Multi-Agent AI Framework for Precision Sentiment Response Generation: Advancing Real-Time Customer Engagement through Large Language Models

Leela Naga Sai Vamsi Krishna Dogiparthi

University of Bridgeport, USA

**ABSTRACT:** The number of digital users across the different sectors has increased exponentially. More users these days are leaning towards shopping online for different types of products, enabling many businesses to expand their reach via internet. As the business grows it becoming more and more challenging for businesses to attain high satisfaction among the customers by addressing their queries and resolving them promptly to retain users, attract new customers and safeguard the business reputation. In this digital era, anything can travel fast to different corners of the world. So, it becomes an utmost importance for the businesses to safeguard their reputations. This exponential user growth generating even higher number of customer reviews are presenting an uphill task for businesses to address each customer. Especially, the ones which have negative experience with your products. Employing human force to address these issues are soon becoming a nearly impossible task and unrealistic. In this paper a novel “Multi-Agent Artificial Intelligent Automation” framework has been proposed to auto address the customer review with human-like responses, auto case resolutions with intelligent solutions to the customer problems to reduce cognitive load on company’s workforces which can be employed to take care of cases which needs human intervention. Our novel framework is estimated to produce high accuracy (~93%) in the sentiment analysis. It is expected to provide a significant boost to the system performance (~40% – 50%) with ~72% reduction in response latency.

**KEYWORDS:** Artificial Intelligence, Accessibility, Natural Language Understanding (NLU), Natural Language Processing (NLP), Sentiment analysis, Multi-Agent Architecture, Response generation, Text Classification, Large Language Model (LLM), Natural Language Generation (NLG), Reasoning Model.

## I. INTRODUCTION

Over the past two decades the digital landscape has taken an upward trend with millions and billions of users surfing on internet for various needs such as online shopping, gaming, job search, housing, banking etc. In 2000, the number of users online were measured approximately ~413 million [1], [2] generating about \$27.6 billion [3] in U.S retail ecommerce sales. In 2020, the digital population bloomed to ~4.7 billion users [1], [2] generating ~\$4.25 trillion in sales globally [4]. By 2025, the online users have reached about ~5.56 billion with a global ecommerce sale of ~\$6.86 trillion and among that \$1.29 trillion has been generated in US alone. As the number of users/customers are growing thus the number of reviews that are generated by are growing exponentially. In 2010 businesses like Trip Advisor has reported over ~50 million reviews globally, followed by yelp with volume of 10 million reviews, total review across the major platforms was reported to ~200 million reviews [5]. In 2015, it was reported that Google attracted over 73% of total reviews of ~1.2 billion reviews. In 2020, the total reviews have reached ~3.5 billion [5], [6]. By 2025, it was reported that 93% of users relied on the reviews to purchase products from the businesses and the total review count was predicted to be 10 billion. This huge flux of customer reviews, user reliance on them created an unending challenge for the business to address user queries especially the ones who has negative experience. Employing human force to tackle a huge volume of reviews has soon become near too impossible, prompting businesses to design more intelligent, scalable & cost-effective solution that can understand the customer review category based on the sentiment analysis followed by action items to address the customer problems that are trivial to resolve. As the volume of the data that will be fed into the LLMs by agent is huge, the solution must scale as needed and should have separation of concerns in place.

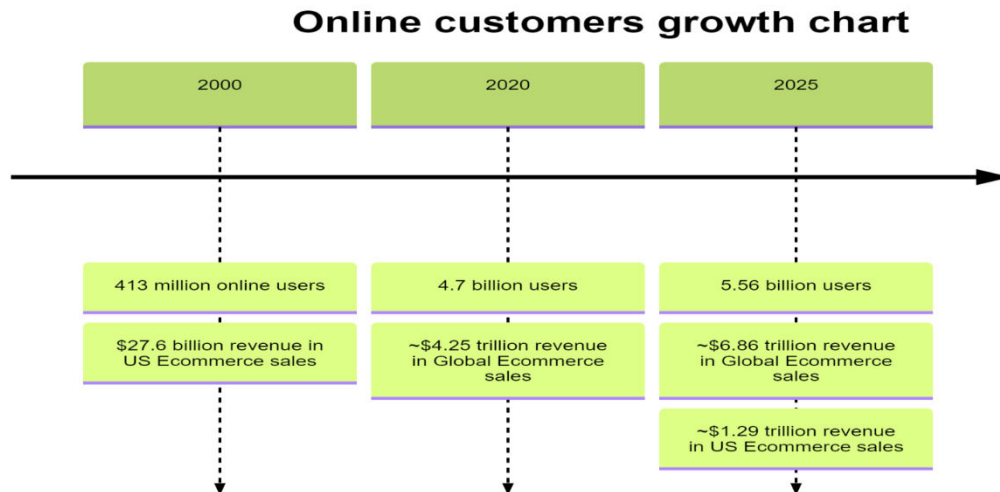


Fig - 1 Online User Growth Chart

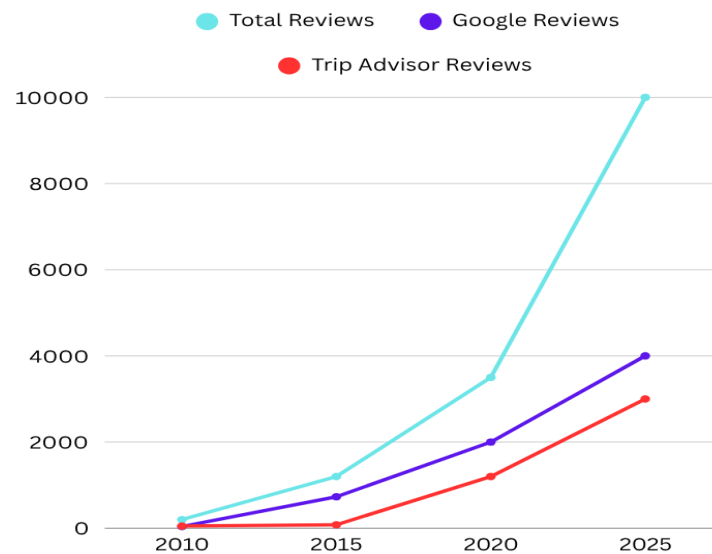


Fig - 2 Customer Review Distribution in millions

## II. LITERATURE REVIEW

Article 1 [7] emphasizes on sentiment analysis of the product reviews using sophisticated Machine Learning (ML) techniques such as Naïve Bayes (CLNB), Brand-Topic Models (dbTM) for overtime evolving insight generation into customer perspectives on the brand value. Even though the article does delve into in depth sentiment analysis of customer reviews, the proposed system lacks in terms of scalability, separation of concerns. It also has limitation in terms of training custom model over time and doesn't provide information about how we can leverage pre-trained LLMs to achieve similar results which are much cost-effective and requires less maintenance and updates. The paper doesn't also discuss about implementation of RAG pattern with LLM which can save huge amount of time for the businesses by injecting contextual information into the LLM for providing grounded responses.



Article 2 [8] conducted research on how to analyse the user sentiment on a certain product to enable companies to understand user sentiments to provide them feedback and guidance on future product development, marketing campaigns etc. The authors implement machine learning methodologies such as binary classification (0 & 1). They also emphasized on polarity analysis to identify the percentage of positive reviews vs negative reviews of the product over a range of  $[-1,1]$ . The proposed system in the article requires continuous training and maintenance of machine learning model to perform the sentiment analysis on the products. It also doesn't emphasize on the architecture of the system, which plays a critical role in the scalability of the business. It also doesn't address the automation of customer review response generation to automate and reduce cognitive workload on the human workforce.

Article 3 [9] Kumar's team tackled ambiguous feedback pools head-on, developing novel aspect-based classification strategies that sharpen decision-making in noisy sentiment environments. Their work specifically addresses uncertainty in unstructured feedback sources—a persistent challenge in opinion mining. The proposed system doesn't address the architectural importance of designing a scalable system for processing the unstructured feedback sources and doesn't emphasize on using pre-trained LLM models with RAG pattern for automation of customer negativity expressed via their feedback.

Article 4 [10] study highlights the approach of using lexicon based approach to construct the sentiment lexicon using words, degree verbs, & negative verbs. Sentiment intensity, sentiment polarity are identified by the constructed lexicon. By doing so the author model is proposed to have higher accuracy where sentiment words are compared input text to identify the sentiment scores. Although the model proposed improves the performance, their study doesn't accommodate how the model can be used in a real-time application and how it can scale to meet the demand.

### III. METHODOLOGY

In our study of creating scalable automated customer review response and case management agent using pre-trained LLM models with sentiment analysis to reduce the cognitive workload on company's human force and tackle unmanageable customer management. We have implemented below methodologies for implementing our system

**Multi-Agent Architecture** – Our system implements multi-Agent modular architecture built using microservices to divide each agent based on its concern and bounded context. This modularity helps our system scale up and down both vertically and horizontally when needed based on the workload. This helps the system to be robust and fault tolerant, allowing it to process any volume of customer reviews

**Sentiment Analysis** – Sentiment analysis is one of widely known and accepted methodology to extract the human opinions from their reviews, process them for understanding whether they had positive, negative or neutral experiences. Each review is scored against three sentiments (positive, negative, & neutral) and the dominated score becomes the overall sentiment of the user. This methodology allows us to tackle each of the review differently, making the system adaptable and understanding. This technique also allows organization to review the overall customer experience with their product(s) or service(s).

**Human-like Response Generation** – Our agentic architecture installs techniques such as Empathy, Natural Language Generation & Behavioural Analysis to generate automated responses to customer reviews which are more human like. This technique allows the organization to generate automated responses that connects them close to their end user.

**Retrieval-Augmented Generation (RAG)** – LLM models are trained on more generalized data and may not have the accurate, up to date context for it to provide grounded responses to the consumer systems. To overcome this challenge, we have implemented RAG pattern to provide LLM models with system specific information that is needed for it to provide the response that is grounded, apt to the system consuming it. This makes our system adaptable to different end environments and stop the LLM hallucinations.

**Prompt Engineering** – Prompt Engineering is one of the key techniques which allows us to fine-tune the behaviour of the LLM model as per the agent needs and avoid training custom models for each agent needs. We have followed 3S principles for designing our System Prompts to enable agents to perform with accuracy.

**Data Processing techniques** – The system implements data processing techniques such as stop word removal, tokenization, Entity extraction to clean the data, extract important information and prepare data for sentiment analysis and sentiment scoring

**Natural Language Processing (NLP) & Natural Language Understanding (NLU)** – NLP & NLU is used by the LLM to understand the user message/review and process it for further analysis such as sentiment scoring etc.

**Prompt filters** – System installs prompt filters to perform various operations such as PII information redaction, removal harmful prompt content to safeguard the LLM model and implement Data privacy and security around the LLM model and system. This is one of the system safeguard techniques

**Responsible AI Framework** – Ethical AI practices are a quite necessary when creating an AI system. The system Implements 6 key principles of Ethical AI [11] which are Fairness, Reliability & safety, Privacy and security, Inclusiveness, Transparency & Accountability to safeguard the system and its user from malpractices. It also helps to deliver a system that can give value to both ends of its spectrum.

**Reasoning Model** – The system is equipped with reasoning LLM models to perform deep reasoning before producing a response to customer reviews. These models increase the accuracy of the result with different steps it took to produce the result for better understanding and scoring the model performance.

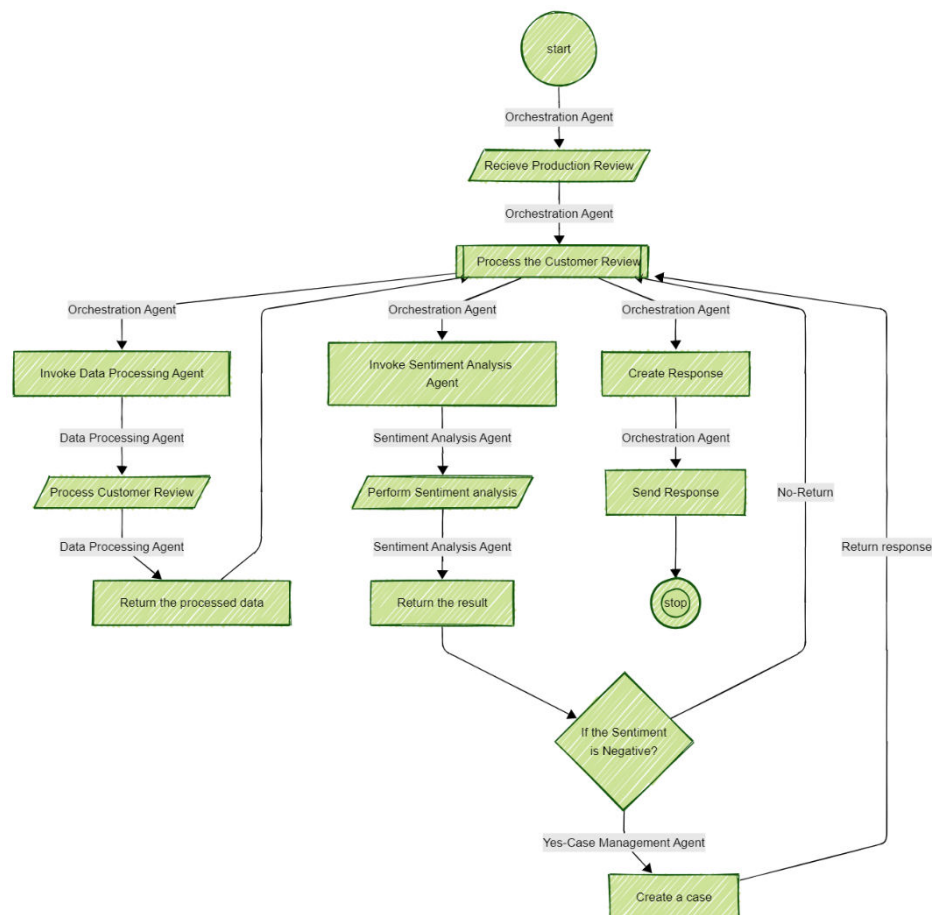


Fig - 3 Review Processing Flow Diagram

#### IV. PROPOSED SYSTEM

The proposed system utilizes Natural Language Processing (NLP), Natural Language Understanding (NLU) & sentiment analysis to understand the user dominated feeling out of the review. It implements Empathy and Natural Language Generation (NLG) to generate human like responses on to the user reviews. The system implements “Prompt Engineering” techniques to fine tune the behaviour of the LLM model for each of its agents. Tokenization & Entity extraction helps the system to perform sentiment scoring, extracting important information for case creation in case the review is deemed to be a negative experience for the end user. Reasoning models increases the accuracy of the result by injecting a deep reasoning & step by step process for deducing the user sentiment from his or her review and generate responses to them. Responsible AI principles are implemented to design system that can be trustworthy, safe and secure to use by the end users and organizations without exploiting their privacy and security. Prompt filters help us to safeguard the system from harmful prompts, redacting personal information from the prompts that are being fed into the LLM model. This is one of the most often used system safeguard techniques implemented at the system level. Retrieval Augmented Generation (RAG) provides the system the opportunity to inject contextual and up to date data into LLM model for producing more grounded result and avoid hallucinations. This technique provides a powerful flexibility and adaptability to the system by allowing it to be reused for multiple purposes without a need to retrain it for each use case.

#### V. SYSTEM ARCHITECTUE

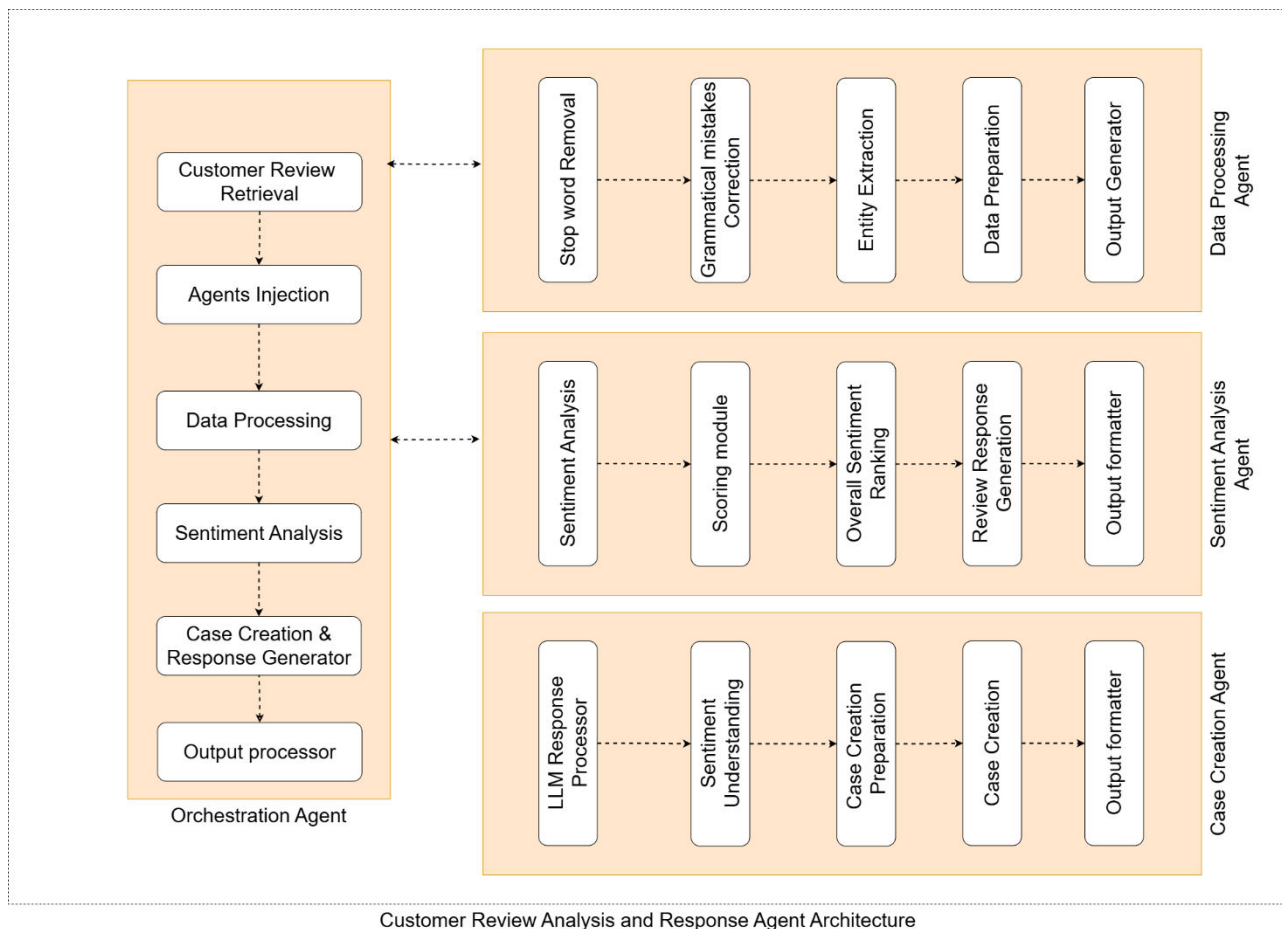


FIG - 4 SYSTEM ARCHITECTURE

## VI. RESULTS

Sentiment Analysis & Sentiment scoring – We have performed sentiment analysis over 100 customer reviews and produced the sentiment score for each three categories positive, negative and neutral for each review with a dominated sentiment being the overall customer review sentiment. Below are the results providing about 93% model accuracy.

Result Matrix				
	Positive	Negative	Neutral	SUM
Positive	35 35.00%	0 0.00%	4 4.00%	39 89.74% 10.26%
Negative	0 0.00%	33 33.00%	3 3.00%	36 91.67% 8.33%
Neutral	0 0.00%	0 0.00%	25 25.00%	25 100.00% 0.00%
SUM	35 100.00% 0.00%	33 100.00% 0.00%	32 78.13% 21.88%	93 / 100 93.00% 7.00%

Fig - 5 Sentiment Analysis Confusion matrix

Customer Review	Identified Sentiment	Agent Response
I'm very disappointed with the Smartphone. It broke after only a week. Don't buy it!	Negative	We're sorry to hear about your experience with the smartphone. Thank you for sharing your feedback. We'll be looking into this issue to prevent future occurrences. Your case number is CAS-20250608-1712. Please contact our customer service for further assistance with this case number.
The Keyboard is okay. It does what it's supposed to do, but nothing special	Neutral	Thank you for your feedback on the keyboard. We appreciate your input and strive to provide products that exceed expectations.
I love my new Monitor! It's everything I hoped for. Great product!	Positive	Thank you for your wonderful feedback! We're thrilled to hear that the monitor meets your expectations.

Table-1: Sample Customer Review and Agent Response Results

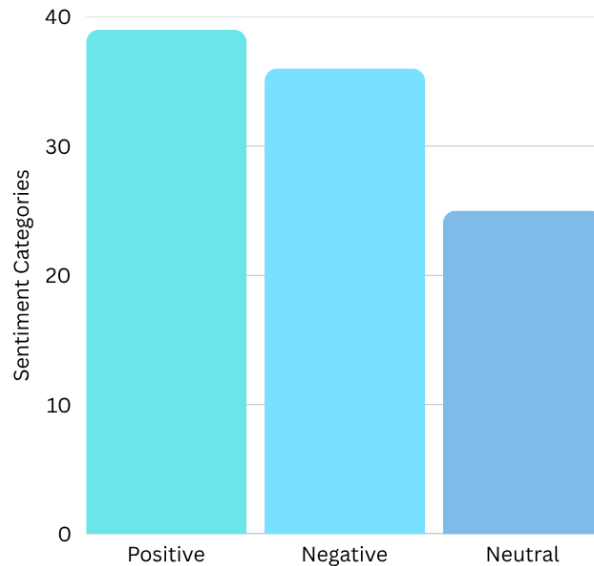


Fig - 6 Customer Reviews Distribution chart

## VII. DISCUSSION

Leveraging pre-trained LLMs with smart prompt engineering is such a game-changer. It offers practical, scalable solutions without demanding massive infrastructure investments or specialized teams.

We've found that layering RAG (Retrieval-Augmented Generation) on top of these models makes responses more reliable and clinically grounded. Think of it as giving the AI guardrails — it stays focused on relevant context instead of hallucinating answers

That said, pre-trained LLMs aren't a silver bullet. There are many known challenges of such system. Below are some of those challenges and possible solutions to overcome them and utilize the powerful capabilities of LLM models without any compromises.

### Challenges & Possible solutions:

“Hallucinations” – Using pre-trained LLM models comes with their challenges, one of the challenges is that as the LLM models are trained on more generalized data and the data used for training is no real-time or up to date. The LLM model also doesn't have contextual information of the current system, which prevents it to provide grounded response by default. To overcome this challenge we inject RAG pattern, where we instruct and provide the LLM model with system specific contextual information as a dictionary for the models to use and produce a response which is more realistic

“Responsible AI practices” – Any AI system is by default subjected to implementing Ethical AI principles to eliminate the privacy & security concerns of the end user. By implementing principles such as Inclusion, Fairness we eliminate bias from the model responses. By implementing the “Prompt Filters” we introduce another layer of security at the system level to detect any harmful content, PII information from the user prompt to safeguard both the system and its end user privacy.



“Prompt Token Limitations” – Often LLM models are known to struggle with token limitations as they implement tokenization of the incoming data to understand the user needs and receive system specific contextual information for it to provide grounded responses. This challenge can be easily managed by implementing certain well-known and widely accepted tokenization techniques such as “Truncation”, “Chunking”, “Text Summarization”, “Removal of Redundant Terms”, “Stop word removal” etc. Sometimes the system must implement more than one of the above-mentioned techniques to reduce the tokens and decrease the cost associated with this.

## VIII. CONCLUSION

Our study on novel Transformative Multi-Agent AI Framework for Precision Sentiment Response Generation enables companies to automate the customer review management and reduce the workload on the human force. This novel framework is designed to provide real-time user review processing with high scalability & adaptability. Our system establishes a new paradigm in customer relation management saving thousands of dollars, processing time for companies and helps their human force to concentrate on more complex cases. The ability to use pretrained LLM models for different use cases and multiple agents saves money and effort for training, maintaining and scaling custom models. Along with the advantages, the system still has some limitations such as multilingual support for processing reviews in different languages, analysing reviews that cannot be categorized into any of the categories because of their non-relevance factor. All these limitations are planned to be addressed in our future study.

## REFERENCES

- [1] “Digital Around the World,” DataReportal – Global Digital Insights. Accessed: May 01, 2025. [Online]. Available: <https://datareportal.com/global-digital-overview>
- [2] “Digital 2025: Global Overview Report,” DataReportal – Global Digital Insights. Accessed: May 01, 2025. [Online]. Available: <https://datareportal.com/reports/digital-2025-global-overview-report>
- [3] J. Alexandria, “Ecommerce statistics 2025: key findings and shopping trends [latest research],” Hostinger Tutorials. Accessed: May 27, 2025. [Online]. Available: <https://www.hostinger.com/tutorials/ecommerce-statistics>
- [4] “51 ECommerce Statistics In 2025 (Global And U.S. Data) | SellersCommerce.” Accessed: May 27, 2025. [Online]. Available: <https://www.sellerscommerce.com/blog/ecommerce-statistics/>
- [5] “81 Online Review Statistics (New 2024 Data).” Accessed: May 27, 2025. [Online]. Available: <https://explodingtopics.com/blog/online-review-stats>
- [6] “Stats Proving the Value of Customer Reviews,” ReviewTrackers. Accessed: May 27, 2025. [Online]. Available: <https://www.reviewtrackers.com/reports/customer-reviews-stats/>
- [7] S. S. Ranjani and S. Kalaichelvi, “Sentiment Analysis of Product Reviews for Ecommerce Recommendation based on Machine Learning,” vol. 14, no. 5, 2025.
- [8] M. Loukili, F. Messaoudi, and M. E. Ghazi, “Sentiment Analysis of Product Reviews for E- Commerce Recommendation based on Machine Learning”.
- [9] V. V. Kumar, K. M. K. Raghunath, V. Muthukumaran, R. B. Joseph, I. S. Beschi, and A. K. Uday, “Aspect based sentiment analysis and smart classification in uncertain feedback pool,” Int. J. Syst. Assur. Eng. Manag., vol. 13, no. 1, pp. 252–262, Mar. 2022, doi: 10.1007/s13198-021-01379-2.
- [10] L. Yang, Y. Li, J. Wang, and R. S. Sherratt, “Sentiment Analysis for E-Commerce Product Reviews in Chinese Based on Sentiment Lexicon and Deep Learning,” IEEE Access, vol. 8, pp. 23522–23530, 2020, doi: 10.1109/ACCESS.2020.2969854.
- [11] “Responsible AI Principles and Approach | Microsoft AI.” Accessed: Jun. 08, 2025. [Online]. Available: <https://www.microsoft.com/en-us/ai/principles-and-approach>



INTERNATIONAL  
STANDARD  
SERIAL  
NUMBER  
INDIA



# INTERNATIONAL JOURNAL OF INNOVATIVE RESEARCH

IN SCIENCE | ENGINEERING | TECHNOLOGY

 9940 572 462  6381 907 438  [ijirset@gmail.com](mailto:ijirset@gmail.com)



[www.ijirset.com](http://www.ijirset.com)

Scan to save the contact details