iJRASET
International Journal For Research in Applied Science and Engineering Technology



INTERNATIONAL JOURNAL FOR RESEARCH

IN APPLIED SCIENCE & ENGINEERING TECHNOLOGY

Volume: 14 **Issue:** III **Month of publication:** March 2026

DOI: https://doi.org/10.22214/ijraset.2026.77313

www.ijraset.com

Call: 08813907089 | E-mail ID: ijraset@gmail.com

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 III Mar 2026- Available at www.ijraset.com

Aspect-Based Sentiment Analysis for E-commerce Insights

Prof. Madhavi Kulkarni[1], Akash Shinde[2], Pankaj Mane[3], Sanket Sonar[4], Shubham Jadhav[5]
[1]*Professor, Department of Computer Engineering, JSPM's BSIOTR, Wagholi*
[2, 3, 4, 5]*Department of Computer Engineering, JSPM's BSIOTR, Wagholi*

Abstract: ***This paper details the creation of a system for a deep analysis of e-commerce reviews, offering a more advanced alternative to traditional sentiment analysis. Conventional tools typically assign a blanket "positive" or "negative" label to an entire review, a method that overlooks the specific nuances within customer feedback. . Our work implements Aspect-Based Sentiment Analysis (ABSA), a technique that pinpoints distinct product attributes—such as "camera," "battery life," or "screen quality"—and then allocates an independent sentiment to each feature. By utilizing Natural Language Processing (NLP) and modern machine learning models like Transformers, the system produces a highly detailed map of consumer opinion.***
Keywords: Aspect-Based Sentiment Analysis (ABSA),E-commerce Reviews, Sentiment Classification, Aspect Extraction, Natural Language Processing (NLP), Transformers, Named Entity Recognition (NER),Deep Learning, BiGRU.

I. INTRODUCTION

(Uncovering Granular Customer Opinions from Product Reviews)

In the age of E-commerce, customer reviews have become integral part of online purchase. Consumers and businesses can't manually mine useful data because of the sheer amount of reviews on the site. [1] This is partially addressed in traditional sentiment analysis systems, where a review has an overall ranking for the possible positive or negative score, which would be misleading. [2] The crucial fact that a client compliments a phone's photography not at the price of its battery life should be preserved, but hidden by these labels.

This limitation is tackled by building an Aspect-Based Sentiment Analysis (ABSA) system. ABSA is a next-level of text analysis tool, which provides more accurate and deeper insights about customer feedback.

II. BACKGROUND REVIEW AND LITERATURE REVIEW

In the digital age of e-commerce, customer reviews have become a staple of any online shopping experience. But there is no way for consumers or analysts to sort through relevant pieces of information manually, since the number of reviews is just enormous. [4] Traditional sentimental analysis partly solves this problem by assigning an overall sentiment score to a review, but this is often deceptive. [6]

If a client praises the photography of a phone while denouncing its battery life, the review should be tagged as “mixed” or “neutral.” Important context would have been stripped away.

It overcomes this limitation by implementing an ABSA system. ABSAThis is an advanced text analysis technique that provides a more accurate and comprehensive insights into customer feedback. [3]

The classification is then performed by passing the features to classifiers like Naive Bayes, Support Vector Machines (SVM), and Logistic Regression. [6] These models work fine for coarse-grained classification however, they fail to handle subtler aspects of language.

Deep Learning : Advanced models such as the LSTMs and RNNs are more capable of understanding the context and sequence of the words in a sentence. [7] The primary cost of these models is the requirement for large, labeled training datasets. Fine-tuning pre-trained masked transformer models, such as BERT and DistilBERT6 they apply contextual embeddings to achieve high accuracy for tasks like aspect extraction and sentiment classification is a more recent and better approach.

Lexicon-Based Approaches: In these approaches, pre-defined word lists are used, where each word has an associated sentiment value. [8] Libraries like TextBlob and VADER are good examples of this approach. Even though they are fast, these methods often miss context and nuances of the text. [9]

©IJRASET: All Rights are Reserved

III. METHODOLOGY

1) Data Collection and Preprocessing: Starting from the raw, unstructured text data is collected from e-commerce platforms by a web scraping tool. [2] Then, a preparation pipeline with the following steps processes this raw data:

- Removing redundant data, special characters or HTML tags.
- Normalizing your text (to remove extra whitespaces, and lowercase it).
- Dealing with informal elements such as slang or emoticons, and lemmatization which reduces words to their canonical form.

2) Aspect Extraction: This building block of the module extracts specific "aspects" or features regarding the product from the cleaned text. There are two main methods to do this:

- Rule Based Approach: Identifying nouns and noun phrases describing entities through NLP libraries (spaCy) etc. [7]
- Machine Learning Approach: A customized model (Model customization) – e.g. constructed using a modified Transformer-based model – is trained with the ability of automatically classifying these attributes by formulating aspect extraction as a Named Entity Recognition (NER) task.

3) Aspect Extraction: This first layer extracts certain "aspects" or features of the product from the cleaned text. There is one of two ways you can do this:

- Rule-Based Approach: Identifying aspects as nouns and noun phrases using NLP libraries like spaCy. [7]
- Machine Learning Technique: A custom model, e.g. a fine-tuned Transformer-based model, is trained to automatically recognize these features by casting the aspect extraction as an NER task.

4) Backend API and Frontend Dashboard: The final step offers a convenient tool for effectively combining the entire pipeline of analyses.

- It is a REST API behind, where it's a process to coordinate, all they need is a user input (like an URL of product) to analysis.
- The integrated ABS data can be visualized on a user-friendly dashboard via interactive charts and tables. [1] This visualization is necessary to take complex data and turn it into actionable insights for businesses and consumers alike.

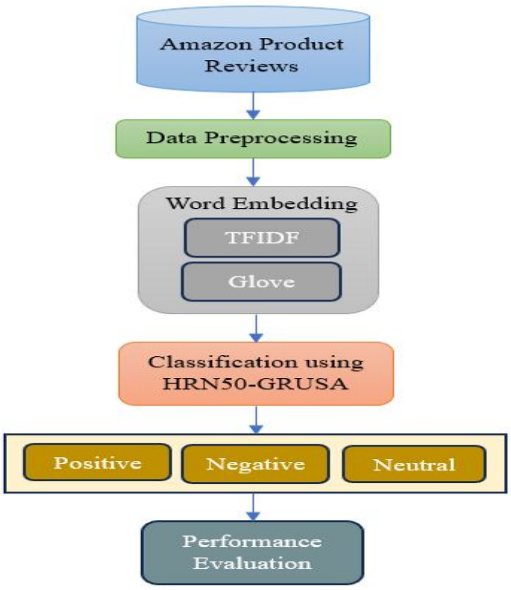


Figure: Workflow of the proposed methodology

A. Current Diagram

1) Shows the word embedding associated with it (TF-IDF, Glove) and is a popular feature representation technique used for general classification.

2) Shows "Classification using HRN50-GRUSA", which means that a single model assigns the entire review.

3) The resulting boxes contain one overall result: "Positive," "Negative" or "Neutral."

B. Improvement

1) Alter the "Word Embedding" to be "Aspect Extraction"(NER). The output of this stage is required to be the exact discussion topic, which must naturally precede sentiment labeling in the processing pipeline. It is the foundation of ABSA.

2) After all, let's not call this "Aspect-based Sentiment Classification," rather, lets call it 'Aspect-Specific Sentiment Classification'! To obtain sentiment with respect to that feature directly, model must be provided not only the reviews and aspect extracted. This is necessary because contradictory judgments may be present for the same review.

3) Full feedback map: we are supposed to get "Aspect-Sentiment Pairs" (e.g. {(Aspect 1, Sentiment), (Aspect 2, Sentiment)}, and then such a map would be supportive to the generationof aspect words. Businesses need this type of actionable feedback.

It is intended to outperform the accuracy of 80.54% set by the previous work, which we hope to achieve by reaching an accuracy of 85%.

C. Project Modules:

Module 1: Data Acquisition And Pre-Processing

Summary: This module is responsible for scrapping product reviews from online stores and pre-processing the text data for analysis.

Objectives: Implement a web scrapper to extract the reviews from specific product page.

Strip out the HTML, special characters and other rubbish:

Prepare your preprocessed data to be used in machine learning models.

Module 2: Aspect Extraction

Input: This is the key modules that discover the highlighted features of products from reviews.

Objectives: Utilize NLP to detect noun and noun phrases (“screen quality”, “battery”) that describe aspects of a product.

Train a custom Named Entity Recognition (NER) model to automatically and accurately categorize these features.

Module 3: Sentiment Classification

Summary: This module checks the sentiment (positive, negative or neutral), for each of the elements in Module 2.

Objectives:

Check the sentence to find out context of each item.

For fine-grained sentiment classification, a pretrained Transformer-based model (e.g., DistilBERT) could be employed.

Module 4: Backend API and Frontend Visualization

First things first: Displaying the results to the user Also, how those models can be used together in a real web application with operator select and thingy images: 2.

Objectives: Develop RESTFUL API for managing user requests and orchestrating the analysis.

A user-friendly front end dashboard with dynamic tables and charts to display combined aspect sentiments.

Model	ABSA Task/ Metric	Accuracy (%)	F1-Score
Data Acquisition and Preprocessing	Aspect-Specific Sentiment	85.56	0.90
Aspect Extraction	Classification (Hybrid)	82.01	93.00
Sentiment Classification	ABSA Classification	75.56	N/A
Backend API & Frontend Visualization	Aspect-Specific Sentiment	78.76	95.00

Table : Comparative Accuracy and F1-Scores in ABSA Research.

IV. ABBREVIATIONS AND ACRONYMS

1) ABSA - Aspect-Based Sentiment Analysis
2) NLP - Natural Language Processing
3) ML - Machine Learning
4) API - Application Programming Interface
5) NER - Named Entity Recognition
6) UI - User Interface
7) JSON: JavaScript Object Notation
8) URL: Uniform Resource Locator
9) NLTK: Natural Language Toolkit
10) SPA: Single-Page Application

V. TOOLS AND TECHNOLOGY

1) Libraries for web scrapping
2) Scraping: We're scraping product reviews from the e-commerce websites using Beautiful Soup, Scrapy or Selenium.
3) Manipulation of Data
   Summary: Pandas and NumPy are important for cleaning, processing, and storing the text data that you collected.
4) Regular Expressions
   Introduction: To Rip Off HTML tags and special characters from Data.
5) NLP Libraries
   Summary: For NLP tasks, espicially rule-based aspect extraction, we use NLTK and spaCy, respectively. [9]
6) Pre-trained Models
   Introduction: Hugging Face Transformers Library is employed to fine-tune BERT or DistilBERT for sentiment classification and aspect extraction (as a Named Entity Recognition task). [1]
7) Deep Learning Frameworks
   Summary: The machine learning models will be built and trained with TensorFlow or PyTorch.
8) Machine Learning
   Introduction Scikit-learn is an essential library for traditional machine learning algorithms and model evaluation.
9) API Communication
   Summary: Calls made between the front-end and back-end are controlled through JavaScript and libraries that include Axios.

VI. CHALLENGES WITH CURRENT TECHNOLOGIES

Information Loss: It's too simplistic to reduce a review to just one rating. Especially, if a client has some positive and negative thoughts about the aspect, it would not be able to grasp depth information in the text. [5] A review along the lines of, "Camera is fantastic, but battery drains fast!" could be considered either as "Neutral" or "Negative", then the camera won't have praise in this specific sense.

Too Little Information That Is of Use: A company may know that a product has some ratio of positive reviews, but they don't necessarily understand why. [9] Fundamental questions such as "What features work well?" "This is not a rhetorical question," wrote Jeremy O'Brien John, Michelle Zimmer, Jennifer Mankoff in their 2010 critique. and "Which ones are failing?" are beyond the response power of classical systems. That prevents companies from getting the data they need to make decisions about how their products can be improved.

Misleading for Mixed Reviews: Sometimes the reviews are as mixed (postive and negative) as they get, but instead of saying "neutral" we label them that way. This masks the good along with the bad, making their overall analysis untrustworthy.

Everyone reads the same review differently. For instance, the phrase "The device feels bulky" is downside to some and neutral to others. Maintaining homogeneity of sentiment in interpretation is exceptionally hard without automated normalization.

No One-Size-Fits-All: Traditional methods do not provide personalized insights for different user needs. [7]Eg, a photographer care more about the quality of camera, but a gamer mainly focuses on performance; "positive" rating does not give them the feature specific aspect to help them make their choices.

VII. EASE OF USE

1) Ease of use: A measure of how easy a system is to learn and operate.
2) Simple Input: Enter the product URL only.
3) Usability: Highlights the user friendliness. Its user friendly nature demonstrates just how easy it is for viewers to maneuver through and understand.
4) Automatic Process: All complex from site scrape and preprocessing to model running is automatic and handled by the system. The user doesn't need to train the models or preprocess the data on his own.
5) Intuitive Visualization: The results are laid out using charts and summary tables. This eliminates the need to manually dig through raw text or data for the insights required.

VIII. APPLICATION

1) Product Development: A company can decide which specific parts of a product are satisfying customers, and formulation thereof. [8] This allows them to pick their own battles. [2] A smartphone corporation might find that people "love" a "camera," but they "hate" the "battery life", so they are forced to produce a better battery in the next generation of phone.
2) Marketing Strategy: ABSA helps marketing teams understand which elements of their ads consumers appreciate and value most, thus making their messaging even more powerful. On the flip side, they can also meet negative comments by launching campaigns that show how an issue has been addressed.
3) Customer Support and Satisfaction: Businesses can proactively identify reoccurring issues based on sentiment around topics such as " shipping," "delivery time," or "customer support". Increased customer satisfaction and loyalty could be the outcome of this targeted policy.
4) Competitive Analysis: Companies can also utilise the ABSA for comparing features of their products with their competitors. This is its competitive edge and produces a clear picture of what they are doing right or wrong in the market.

Buying Decisions – Consumers often struggle to understand the pros and cons of thousands of products available in the market/to perceive what benefits they seek from a product. To assist them in making better decisions based on their own requirements, the ABSA report can provide a standardised summary that shows what are favoured features, and what are not. [6]

IX. FUTURE DIRECTIONS AND RESEARCH OPPORTUNITIES

1) Handling implicit information: Another problem is to trace the implicit features or contexts which are not stated explicitly in the text. In the review it got really hot after a few minutes of use, for example, performance or heat dissipation is understood. 2 sheer, when users are not given the opportunity to skim through an article at all, it is important that more sophisticated context awareness and common sense reasoning be employed in retrieving these rather than raw keyword or noun phrase extraction.
2) Irony and Sarcasm Detection: It could also be the case that our algorithm does not work well with sarcasm, in which posative phrases are used to convey negative sentiment. [9]Further works could focus on algorithms that capture this subtlety better, as it is a major issue in sentiment analysis.
3) Sentiment Intensity and Causality: A more advanced approach might be to predict the strength of a sentiment (e.g., -1 to +1) instead of just its polarity. Note that this research could also focus on reasoning about cause and effect relationships among the components (e.g., knowing that low rating is because of short battery life). [5]
4) Multi-language: It supports more than one language Translation support by can be integrated. This important area of research is usually carried out by constructing new datasets and models to work for multiple languages, or for using cross-lingual pre-trained models.
5) D2: Create Robustness-Oriented Datasets The community should focus on creating robustnessoriented datasets that capture more realistic challenge by considering, for example, multi-aspect and multisentiment sentences rather than being stuck in the self-reinforcing loop of research by limitations of popular benchmark datsets like SemEval.
6) Cross-Domain Analysis: We need to adapt the system for different domains (restaurants, travel, and public opinion), as these can have very different elements and vocabulary. [7] Research is necessary to develop models which can efficiently perform such domain transfer so that they need not be retrained on a new task/domain.
7) Integration to External Data: In order to present a more complete view of consumer response, the system could be integrated with other data sources (e.g., social media mentions, customer service tickets and survey replies).
8) Real-Time: Also by tuning the system for real-time analysis on new uploaded evaluations, organisatios can quickly act on problems or trends as they are discovered.

©IJRASET: All Rights are Reserved | SJ Impact Factor 7.538 | ISRA Journal Impact Factor 7.894 |

X. CONCLUSION

This project successfully moves beyond the limitations of traditional sentiment analysis by providing deep, granular insights from customer reviews. The proposed system dissects product reviews to extract specific features (aspects) and assign sentiment to each one, offering a detailed, feature-level analysis. The system's modular architecture, encompassing data acquisition, aspect extraction, sentiment classification, and a user-friendly frontend dashboard, provides a robust solution for a complex problem. This enables businesses to make data-driven decisions to improve their products and marketing strategies. It also empowers consumers to make smarter purchasing choices by quickly understanding the specific pros and cons of a product. The project's design allows for future improvements, such as adding support for more languages or detecting sarcasm, ensuring its adaptability and long-term value.

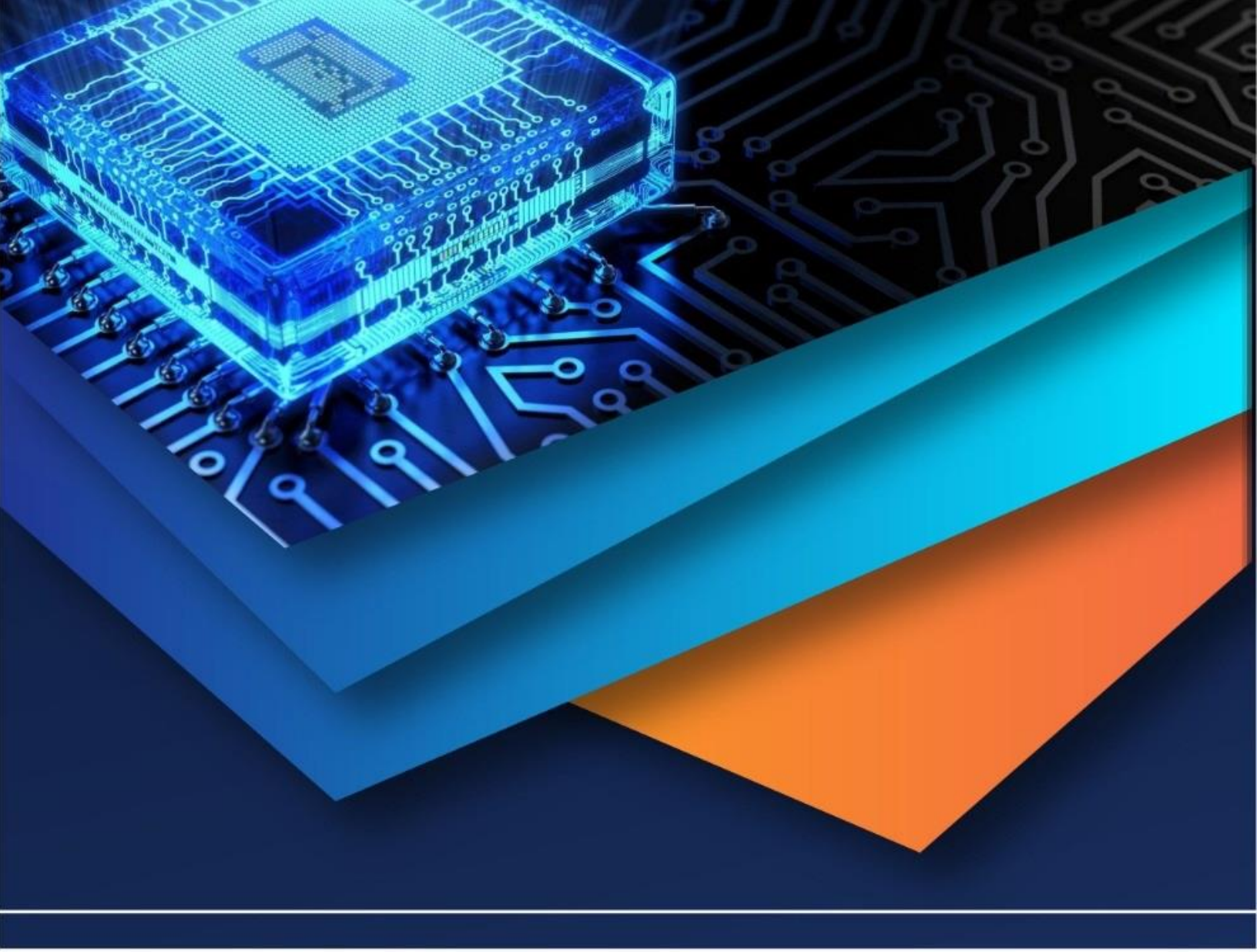
XI. ACKNOWLEDGEMENT

We extend our deepest thanks to Prof.Madhavi Kulkarni for their expert guidance, constructive feedback, and continuous encouragement throughout the development of this project. This project was a collaborative effort, and we are grateful to each team member for their hard work and dedication in their respective modules, from data acquisition to front-end visualization. We acknowledge the foundational work in the fields of Natural Language Processing and Sentiment Analysis. We are particularly thankful to the authors of key research papers and the creators of the open-source libraries that made this project possible. We are indebted to the open-source community for providing the powerful tools and libraries, such as Python, scikit-learn, Pandas, NumPy, spaCy, and Hugging Face Transformers, that were instrumental in the development and implementation of our system.

REFERENCES

[1] Aspect-based sentiment classification of user reviews to understand customer satisfaction of e-commerce platforms — L. Davoodi, J. Mezei, M. Heikkilä (2025).
[2] Explainable Aspect-Based Sentiment Analysis Using Attention + SHAP — W. Toms (2025, SSRNpreprint).
[3] Fake review detection in e-Commerce platforms using Aspect-Based Sentiment Analysis (FakeDetectionGCN) — P. Phukon et al. (2025, MDPI / Applied Sciences).
[4] AI-Driven Sentiment Analytics: Unlocking Business Value in the E-Commerce Landscape — Q. Wu, C. Xia, S. Tian (2025, arXiv).
[5] Leveraging Deep Neural Networks for Aspect-Based Sentiment Classification (EEGCN) — C. Li, D. Cheng, Y. Morimoto (2025, arXiv).
[6] Aspect-Based Sentiment Analysis on E-commerce Reviews using BiGRU and Bi-Directional Attention Flow (BiGRU+BiDAF) — (Apr 2025 / ResearchGate / local journals).
[7] A Study On Improving Aspect-Based Sentiment Analysis — a generative instruction-tuned framework — (2024, arXiv).
[8] A systematic review of aspect-based sentiment analysis — Y. C. Hua et al. (2024, Springer).
[9] Aspect-Based Sentiment Analysis of Clothing Reviews in Vietnamese — P. Quoc-Hung et al. (PACLIC 2024 / ACL Anthology).
[10] Aspect based sentiment analysis of consumer reviews using an Unsupervised Attention Deep Neural Framework — A. Dey et al. (Dec 2024 / Applied Soft Computing or similar).

©IJRASET: All Rights are Reserved | SJ Impact Factor 7.538 | ISRA Journal Impact Factor 7.894 |



10.22214/IJRASET

45.98

IMPACT FACTOR: 7.129

IMPACT FACTOR: 7.429

INTERNATIONAL JOURNAL FOR RESEARCH

IN APPLIED SCIENCE & ENGINEERING TECHNOLOGY

Call : 08813907089 (24*7 Support on Whatsapp)