International Journal of
DATA SCIENCE AND IOT MANAGEMENT SYSTEM
Peer Reviewed, Referred & Indexed Journal
ISSN: 3068-272X www.ijdim.com Original Research Paper

A Robust Agentic Pipeline for Dual-Target Classification of Mobile User Reviews and Ratings

P. Vijay Goud[1*], G Navya[2], Rayala Navya sree[2], Basavaraju Neelesh[2], Gudem Praveen[2]

[1]Assistant Professor, [2]UG Student,[1,2]Department of Computer Science and Engineering

[1,2]Kommuri Pratap Reddy Institute of Technology, Ghanpur, Ghatkesar, 501301, Telangana, India.

*Correspondence: P. Vijay Goud (panjala.vijay123@gmail.com)

Abstract

Mobile reviews act as a key indicator of customer satisfaction, with over 90% of smartphone users referring to reviews before purchasing a device, and 72% of consumers indicating that positive feedback enhances their trust in a brand. However, manual examination of customer reviews and ratings is inefficient, susceptible to errors, and often fails to capture the subtle emotions present in unstructured text. To address these issues, this study proposes a framework based on Natural Language Processing (NLP) using an iPhone 14 dataset that includes user reviews, titles, and ratings. The workflow begins with NLP preprocessing and Exploratory Data Analysis (EDA) to clean, standardize, and visualize the data distribution. Next, Efficiently Learning an Encoder that Classifies Token Replacements Accurately (ELECTRA) is utilized for contextual feature extraction, enabling effective semantic representation of textual data. To handle class imbalance in review categories, K-Means Synthetic Minority Over-Sampling Technique (K-Means SMOTE) is applied to generate synthetic samples. Unlike existing approaches such as Adaptive Boosting Classifier (ABC) and Tao Tree Classifier (TTC), the proposed framework integrates an Extra Trees Classifier (ETC) to ensure more robust and scalable classification. The model predicts bi-variate outputs Review Title and Rating thereby improving both sentiment understanding and rating consistency. By automating the analysis of reviews, the system provides valuable insights into customer satisfaction, product performance, and brand perception, ultimately supporting better decision-making and enhancing the overall customer experience.

Keywords: Sentiment analysis, mobile reviews, Natural Language Processing (NLP), ELECTRA, review prediction, customer satisfaction

1. Introduction

In the highly competitive smartphone industry, customer reviews and ratings have become vital factors influencing purchasing decisions and shaping brand reputation. Studies indicate that nearly 95% of consumers read online reviews before buying a product, and smartphones with higher average ratings can achieve up to a 20% increase in sales conversions. With a massive volume of reviews being generated daily across multiple platforms, businesses encounter significant difficulties in manually analysing and interpreting customer sentiments. As a result, automated systems have become essential for extracting emotions from textual reviews, accurately classifying user feedback, and delivering meaningful insights for product enhancement and improved customer engagement. In the current digital age, mobile phones rank among the most widely used consumer products, with millions of users in India actively expressing their opinions through online reviews. According to Statista, India recorded over 750 million smartphone users in 2024, and this number continues to grow rapidly each year. Along with the growth of mobile phone usage, e-commerce platforms such as Flipkart, Amazon, and dedicated mobile marketplaces have witnessed an enormous surge in customer reviews. These reviews are not just star ratings but detailed experiences that reflect user satisfaction, frustration, and emotional responses. Understanding these emotions is crucial for mobile companies, retailers, and service providers to improve product quality and customer satisfaction. The analysis of mobile reviews involves two aspects: identifying the emotional tone behind the words and linking it with the numeric rating provided by the customer. NLP offers the ability to process these large volumes of textual data and convert them into structured insights. By decoding customer emotions, companies gain a deeper understanding of what drives user satisfaction and dissatisfaction, enabling them to make data-driven decisions in product development, marketing strategies, and after-sales service. The research focuses on applying NLP to analyse mobile reviews and decode customer emotions. By linking emotional patterns in text with review ratings, businesses gain insights into consumer satisfaction. The approach enhances decision-making in marketing, product design, and customer engagement. It also provides customers with transparent information before purchasing.

2. Literature Survey

Allouch et al. [1] described along with the main technologies that enable the creation of CAs. Capable of conducting ongoing communication with humans, CAs are encountered in natural-language processing, deep learning, and technologies that integrate emotional aspects. The technologies used for the evaluation of CAs and publicly available datasets are outlined. In addition, several areas for future research are identified to address moral and security issues, given the current state of CA-related technological developments. The uniqueness of our review is that an overview of the concepts and building blocks of CAs is provided, and CAs are categorized according to their abilities and main application domains. In addition, the primary tools and datasets that may be useful for the development and evaluation of CAs of different categories are described. Finally, some thoughts and directions for future research are provided, and domains that may benefit from conversational agents are introduced. Pota et al. [2] introduced a different approach for Twitter sentiment analysis based on two steps. Firstly, the tweet jargon, including emojis and emoticons, is transformed into plain text, exploiting procedures that are language-independent or easily applicable to different languages. Secondly, the resulting tweets are classified using the language model BERT, but pre-trained on plain text, instead of tweets, for two reasons: pre-trained models on plain text are easily available in many languages, avoiding resource- and time-consuming model training directly on tweets from scratch; available plain text corpora are larger than tweet-only ones, therefore allowing better performance. A case study describing the application of the approach to Italian is presented, with a comparison with other Italian existing solutions. The results obtained show the effectiveness of the approach and indicate that, thanks to its general basis from a methodological perspective, it can also be promising for other languages.

García-Madurga et al. [3] involved a comprehensive search in the Scopus and Web of Science databases, following PRISMA guidelines, resulting in 31 articles that met the inclusion criteria. The qualitative synthesis reveals that AI is being utilized in areas such as mental health monitoring, emotional support, personalized well-being programs, identification of psychosocial risk factors, and training and development. This review contributes to the existing literature by offering a detailed categorization of AI applications in workplace well-being, and it highlights the practical utility of AI in enhancing employee

International Journal of
DATA SCIENCE AND IOT MANAGEMENT SYSTEM
Peer Reviewed, Referred & Indexed Journal
ISSN: 3068-272X www.ijdim.com Original Research Paper

mental health and overall well-being. The findings suggest that AI has the potential to revolutionize the management of workplace well-being, providing actionable insights for both researchers and practitioners. Recommendations for future research are also discussed.

Chen et al. [4] presented a comprehensive literature review on the convergence of affective computing, interactive installation art, multi-dimensional sensory stimulation, and artificial intelligence (AI) in measuring emotional responses, demonstrating the potential of artificial intelligence in emotion recognition as a tool for sustainable development. It addresses the problem of understanding emotional response and measurement in the context of interactive installation art under artificial intelligence (AI), emphasizing sustainability as a key factor. The study aims to fill the existing research gaps by examining three key aspects: sensory stimulation, multi-dimensional interactions, and engagement, which have been identified as significant contributors to profound emotional responses in interactive installation art. The proposed approach involves conducting a process analysis of emotional responses to interactive installation art, aiming to develop a conceptual framework that explores the variables influencing emotional responses. This study formulates hypotheses that make specific predictions about the relationships between sensory stimulation, multi-dimensional interactions, engagement, and emotional responses. By employing the ASSURE model combined with experimental design, the research methodology ensures a systematic and comprehensive study implementation. The implications of this research lie in advancing the understanding of emotional experiences in interactive installation art under AI, providing insights into the underlying mechanisms that drive these experiences, and their influence on individual well-being from a sustainable perspective. The contributions of this research include bridging the identified research gaps, refining theoretical frameworks, and guiding the design of more impactful and emotionally resonant interactive artworks with sustainability in mind.

Alabdan et al. [5] reviewed of the approaches used during phishing attacks is presented. This paper comprises a literature review, followed by a comprehensive examination of the characteristics of the existing classic, modern, and cutting-edge phishing attack techniques. The aims of this paper are to build awareness of phishing techniques, educate individuals about these attacks, and encourage the use of phishing prevention techniques, in addition to encouraging discourse among the professional community about this topic. Leo et al. [6] proposed an overview of the cutting-edge approaches that perform facial cue analysis in the healthcare area. The document is not limited to global face analysis but it also concentrates on methods related to local cues (e.g., the eyes). A research taxonomy is introduced by dividing the face in its main features: eyes, mouth, muscles, skin, and shape. For each facial feature, the computer vision-based tasks aiming at analyzing it and the related healthcare goals that could be pursued are detailed.

Borna et al. [7] aimed to explore how artificial intelligence can help ease the burden on caregivers, filling a gap in current research and healthcare practices due to the growing challenge of an aging population and increased reliance on informal caregivers. We conducted a search with Google Scholar, PubMed, Scopus, IEEE Xplore, and Web of Science, focusing on AI and caregiving. Our inclusion criteria were studies where AI supports informal caregivers, excluding those solely for data collection. Adhering to PRISMA 2020 guidelines, we eliminated duplicates and screened for relevance. From 947 initially identified articles, 10 met our criteria, focusing on AI's role in aiding informal caregivers. These studies, conducted between 2012 and 2023, were globally distributed, with 80% employing machine learning. Validation

International Journal of
DATA SCIENCE AND IOT MANAGEMENT SYSTEM
Peer Reviewed, Referred & Indexed Journal
ISSN: 3068-272X www.ijdim.com Original Research Paper

methods varied, with Hold-Out being the most frequent. Metrics across studies revealed accuracies ranging from 71.60% to 99.33%. Specific methods, like SCUT in conjunction with NNs and LibSVM, showcased accuracy between 93.42% and 95.36% as well as F-measures spanning 93.30% to 95.41%. AUC values indicated model performance variability, ranging from 0.50 to 0.85 in select models. Our review highlights AI's role in aiding informal caregivers, showing promising results despite different approaches. AI tools provide smart, adaptive support, improving caregivers' effectiveness and well-being.

Pu et al. [8] proposed a stance detection approach based on multi-task learning that considers the influence of emotional features to tackle these challenges. Their method utilizes a RoBERTa pre-trained model in the shared layer to extract textual features for both stance detection and sentiment analysis. In the stance detection module, a BiLSTM model captures deeper temporal information, followed by three independent modules dedicated to extracting semantic features for specific stances. Concurrently, the sentiment analysis module employs a BiLSTM model to predict emotional polarity. The experimental results on the NLPCC2016-task4 dataset demonstrate that our approach outperforms existing methods, highlighting the effectiveness of integrating sentiment analysis with stance detection to enhance both accuracy and reliability, ultimately contributing to the security of social networks.

Li et al. [9] developed a two-stage nonlinear user satisfaction decision model (USDM). First, we use word2vec technology and lexicon-based sentiment analysis to mine the sentiment polarity of each product attribute in the reviews. Then, we develop KANO mapping rules using utility functions to classify consumer preferences based on attribute importance. Based on this, a two-stage nonlinear USDM is developed to describe post-purchase evaluation behaviour. In the first non-compensatory stage, consumers determine their initial satisfaction level based on the performance of basic attributes. If the performance of these attributes is poor, it is almost impossible for users to be satisfied. In the compensatory stage, the performance of the remaining attributes collectively affects final satisfaction through participation in user utility calculation. With the use of reviews from JD.com, we develop a genetic algorithm to determine feasible solutions for the USDM and verify its validity and robustness. The USDM is proven to be effective in predicting user satisfaction compared to other classic models and machine learning algorithms. This study provides a universal pattern for user satisfaction decisions and extends the study on preference analysis.

Silva et al. [10] reviewed the application of ontologies and knowledge graphs in cancer research. In total, our review encompasses 141 published works, which we categorized under 14 hierarchical categories according to their usage of ontologies and knowledge graphs. They also review the most used ontologies and newly developed ones. Our review highlights the growing traction of ontologies in biomedical research in general, and cancer research in particular. Ontologies enable data accessibility, interoperability and integration, support data analysis, facilitate data interpretation and data mining, and more recently, with the emergence of the knowledge graph paradigm, support the application of Artificial Intelligence methods to unlock new knowledge from a holistic view of the available large volumes of heterogeneous data.

3. Proposed Methodology

The proposed sentiment analysis system for mobile phone reviews is developed as a multi-stage automated pipeline, as illustrated in Figure 1. This pipeline systematically processes raw review data, extracts relevant features, addresses class imbalance, and predicts both review titles and ratings. The process begins with the collection of a comprehensive iPhone 14 dataset that includes review text, titles, and ratings. In the first step, data preprocessing is performed, which involves removing null values, normalizing text, tokenizing content, and applying label encoding to transform categorical data into machine-readable formats. The second step focuses on EDA to analyse sentiment distribution and detect any class imbalance within the dataset. In the third step, baseline models such as ABC and TTC are implemented for performance comparison. The fourth step involves building the proposed model, which combines ELECTRA embeddings for contextual feature extraction, K-Means SMOTE for generating synthetic samples to balance minority classes, and ETC for achieving robust and scalable predictions of both review titles and ratings. In the fifth step, the model's performance is evaluated using metrics such as accuracy, F1-score, and mean absolute error to ensure consistency and reliability.

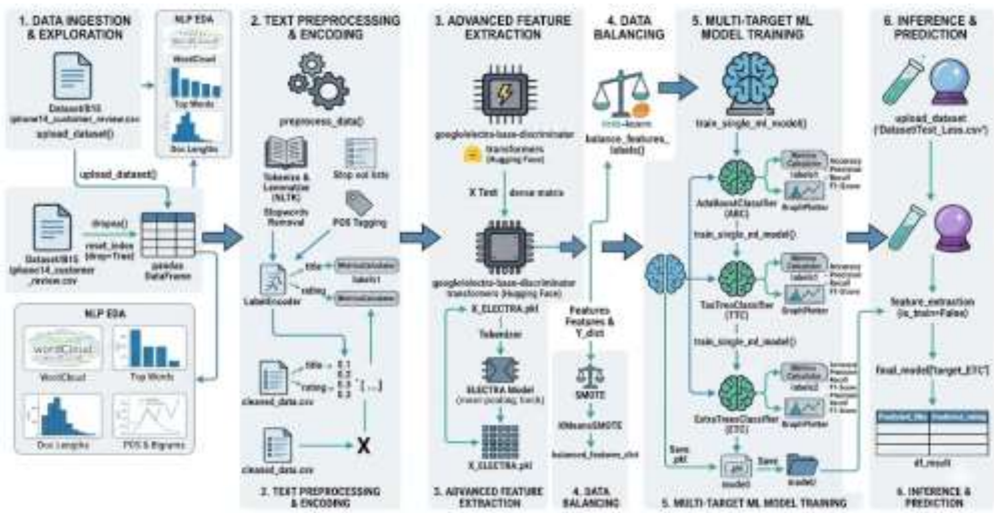


Figure. 1: Proposed system architecture.

Dataset Collection: The first step of the proposed system involves collecting raw customer reviews for the iPhone 14 from various e-commerce platforms. Each review is carefully examined to ensure it contains essential fields, including the review text, the review title, and the numerical rating. The dataset is validated to ensure diversity and representativeness, capturing a wide range of customer sentiments such as positive, negative, and neutral opinions. Additionally, the size of the dataset is checked to ensure it is sufficient for machine learning training, while also maintaining enough coverage of nuanced customer feedback to support robust analysis.

Dataset Preprocessing: Once the dataset is collected, preprocessing is carried out to clean and structure the data for analytical modelling. Any null or missing entries are removed, and inconsistencies in the data are addressed to improve quality. The review text undergoes normalization by converting all characters to lowercase, removing punctuation and special characters, and trimming extra whitespace. Tokenization is applied to split the text into meaningful words or subwords, creating a structured input format. Furthermore, categorical targets, such as review titles and ratings, are transformed into numeric labels through label encoding, resulting in a clean and consistent dataset ready for feature extraction and model training.

International Journal of
DATA SCIENCE AND IOT MANAGEMENT SYSTEM
Peer Reviewed, Referred & Indexed Journal
ISSN: 3068-272X www.ijdim.com Original Research Paper

EDA and Baseline Modelling: The pre-processed dataset is then subjected to exploratory data analysis to gain insights into sentiment distributions, word usage patterns, and potential class imbalances. Visualization techniques, including word clouds, frequency plots, and histograms, are employed to highlight common terms and the distribution of review lengths. To establish a benchmark, baseline models such as AdaBoost Classifier and Tao Tree Classifier are trained on the data. The performance of these baseline models provides reference metrics, which serve as a standard against which the proposed hybrid model will be evaluated in later stages.

Proposed Hybrid Model Construction: The core of the proposed system involves constructing a hybrid model that integrates advanced NLP embeddings with ensemble learning. Contextual and semantic features are extracted from the review text using ELECTRA embeddings, capturing subtle nuances in language and sentiment. To handle class imbalance in the dataset, K-Means SMOTE is applied to generate synthetic samples for underrepresented sentiment categories. These enriched features are then input into an ETC, which is capable of predicting multiple targets simultaneously, including review titles and ratings. This integration of embedding-based features, class balancing, and a robust ensemble classifier ensures a scalable and generalized dual-output predictive model.

Performance Evaluation: Once the hybrid model is trained, its performance is rigorously evaluated on a validation dataset. Metrics such as accuracy for review titles, F1-score for sentiment prediction, and mean absolute error (MAE) for ratings are calculated to measure predictive effectiveness. Cross-validation is employed to assess the generalization ability of the model and prevent overfitting to the training data. The performance results are compared against the baseline models to demonstrate the improvements achieved by the proposed hybrid approach, highlighting its ability to accurately capture sentiments and predict multiple review attributes simultaneously.

Prediction on New, Unseen Test Data: In the final stage, the trained model is applied to new, unseen review data to simulate a real-world scenario. The system generates predictions for both review titles and ratings, providing actionable insights into customer satisfaction and product performance. The outputs are visualized through a GUI dashboard, which allows stakeholders to interpret trends, identify key issues, and make informed, data-driven decisions. This step ensures that the entire pipeline is operational, scalable, and ready for practical deployment in business analytics and sentiment analysis applications.

3.1 ELECTRA

The ELECTRA architecture is a transformer-based model designed to enhance pretraining efficiency and contextual understanding in NLP tasks. Unlike models such as BERT, which rely on masked language modelling (MLM) to predict missing words, ELECTRA adopts a replaced token detection (RTD) strategy that trains the model to distinguish between original and synthetically replaced tokens in a sequence. This method significantly improves learning efficiency while maintaining or even surpassing the representational power of larger models. The ELECTRA architecture comprises two major components the Generator and the Discriminator. The process begins with an input sentence (e.g., "The chef cooked the meal."). During training, certain tokens in this sentence are randomly masked (e.g., "<M> chef <M> the meal."). The Generator, a smaller masked language model, predicts plausible replacements for the masked tokens (e.g., "The chef ate the meal.").

International Journal of
DATA SCIENCE AND IOT MANAGEMENT SYSTEM
Peer Reviewed, Referred & Indexed Journal
ISSN: 3068-272X www.ijdim.com Original Research Paper

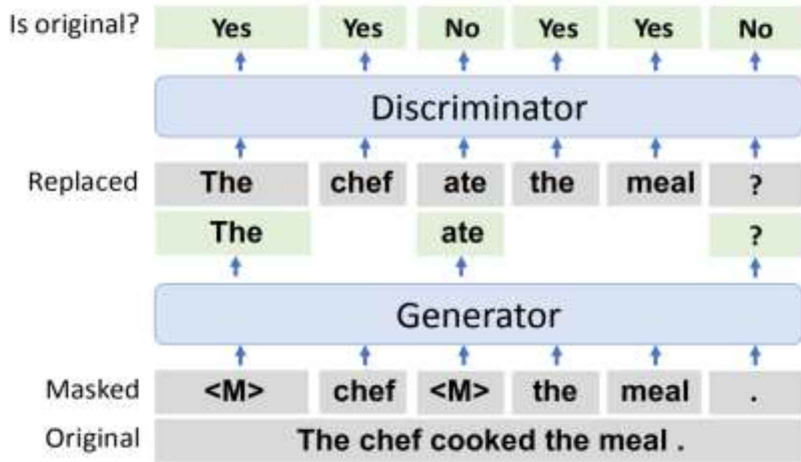


Figure. 2: ELECTRA architecture generator discriminator based contextual feature extraction framework

These replaced tokens form a new sequence, which is then passed to the Discriminator. The Discriminator's task is not to predict what the masked words were, but rather to determine for each token in the sequence whether it is original or replaced. For example, it outputs "Yes" for original tokens and "No" for the ones substituted by the generator. Through this adversarial setup, the Discriminator learns highly contextualized token representations by evaluating the authenticity of each token in its sentence context. The model thereby captures both local and global dependencies across words, enabling nuanced comprehension of language structure, sentiment, and meaning. The Generator is trained briefly to provide challenging replacements, but the Discriminator receives most of the learning signal making ELECTRA more sample-efficient and computationally faster than conventional MLM models. In this research, ELECTRA is utilized for contextual feature extraction from customer reviews. Its architecture allows it to recognize subtle emotional cues, negations, and context-dependent meanings that traditional models often overlook. By leveraging the Generator–Discriminator dynamic, the model learns discriminative embeddings that accurately represent sentiment polarity and intensity. These embeddings are then fed into the ETC for downstream classification of review titles and ratings, forming a powerful hybrid framework that unites the interpretability of ensemble learning with the contextual power of transformer-based NLP.

4. Results Description

Result analysis is a crucial stage in any study or experiment, as it involves interpreting the collected data to draw meaningful conclusions. It helps in identifying patterns, relationships, and trends within the results. Through analysis, raw data is transformed into useful information that supports or refutes the research objectives. This process also allows researchers to evaluate the accuracy and reliability of their findings. By comparing expected outcomes with actual results, one can determine the success or limitations of the study. Additionally, result analysis aids in making informed decisions and recommendations for future research or practical applications.

International Journal of
DATA SCIENCE AND IOT MANAGEMENT SYSTEM
Peer Reviewed, Referred & Indexed Journal
ISSN: 3068-272X www.ijdim.com Original Research Paper

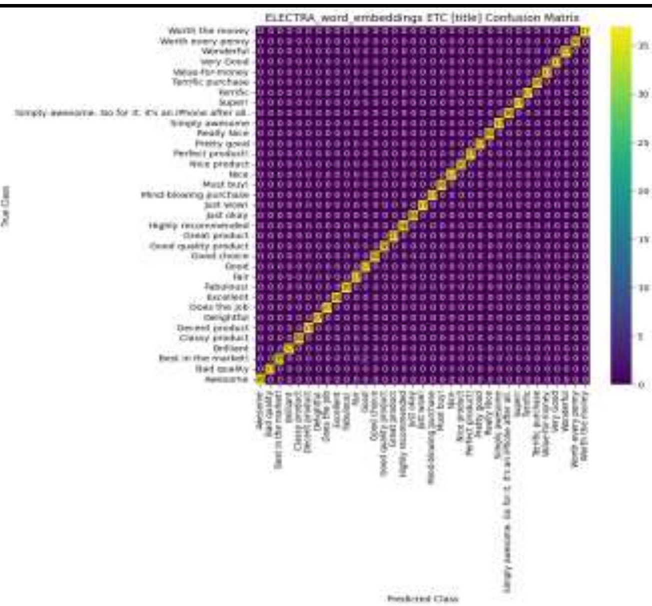


Figure. 3: Confusion matrix obtained using ELECTRA _WE_ ETC Classifier for title.

Figure 3 presents a comparative analysis of confusion matrices for review title classification using ELECTRA word embeddings across showcases the confusion matrix for the proposed ETC Classifier, demonstrating superior performance with near-perfect diagonal alignment across all title classes, minimal off-diagonal noise, and exceptionally high prediction counts on correct labels (e.g., "Worth the money," "Simply awesome"), confirming its robustness and accuracy in multi-class sentiment title prediction.

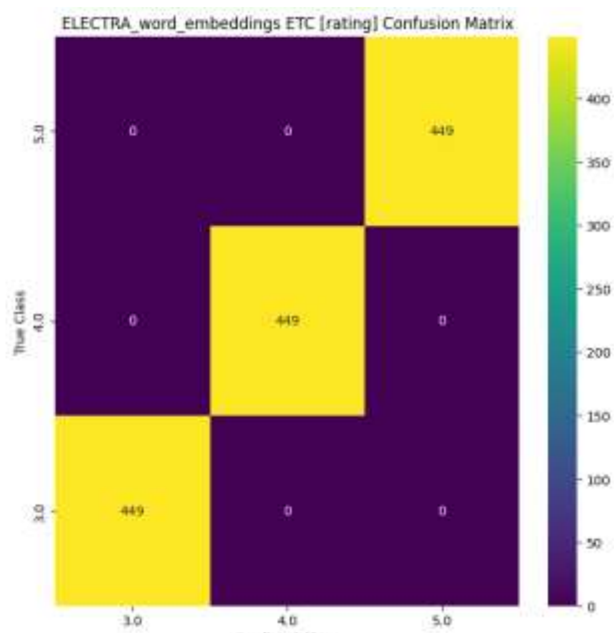


Figure. 4: Confusion matrix obtained using ELECTRA _WE_ETC for rating.

International Journal of
DATA SCIENCE AND IOT MANAGEMENT SYSTEM
Peer Reviewed, Referred & Indexed Journal
ISSN: 3068-272X www.ijdim.com Original Research Paper

Figure 4 compares confusion matrices for rating prediction (classes 3.0, 4.0, 5.0) using ELECTRA word embeddings across displays the proposed ETC Classifier's confusion matrix, achieving near-perfect diagonal dominance with 449 correct predictions each for classes 3.0, 4.0, and 5.0, and zero misclassifications across all off-diagonal cells, confirming its exceptional ability to accurately predict star ratings despite severe class imbalance.

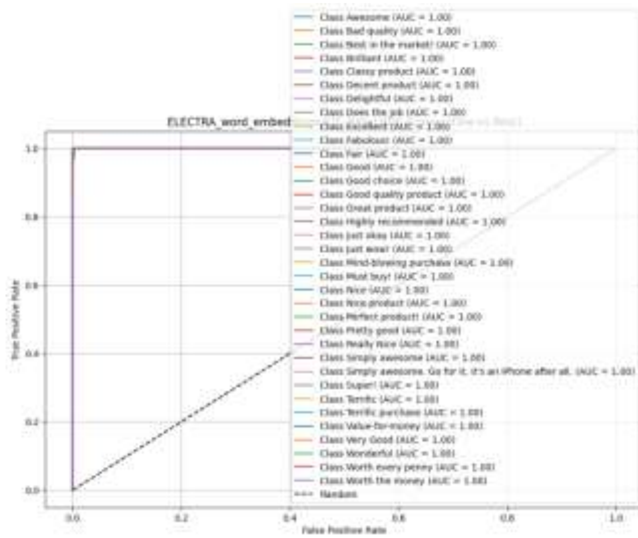


Figure. 5: ROC Curve obtained using ELECTRA _WE_ETC for title

Figure 5 presents the Receiver Operating Characteristic (ROC) curves with corresponding Area Under the Curve (AUC) scores for multi-class review title classification using ELECTRA word embeddings ETC Classifier, achieving perfect AUC = 1.00 across all title classes, with every curve tightly aligned to the top-left corner, indicating flawless true positive detection at zero false positives and confirming the model's outstanding capability to accurately distinguish even subtle variations in customer sentiment titles.

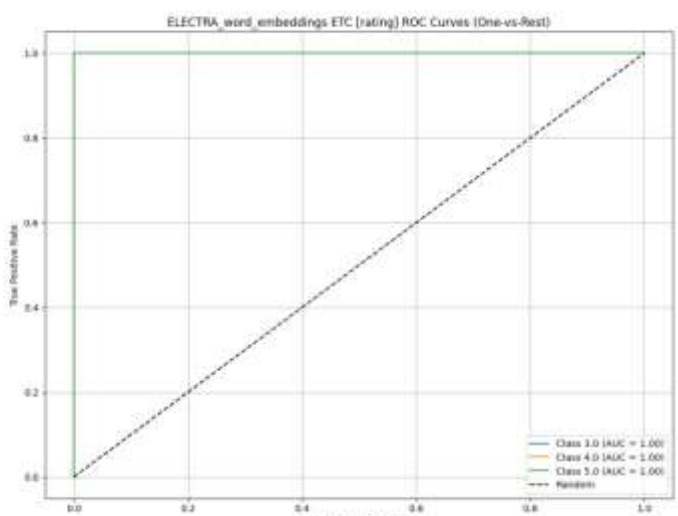
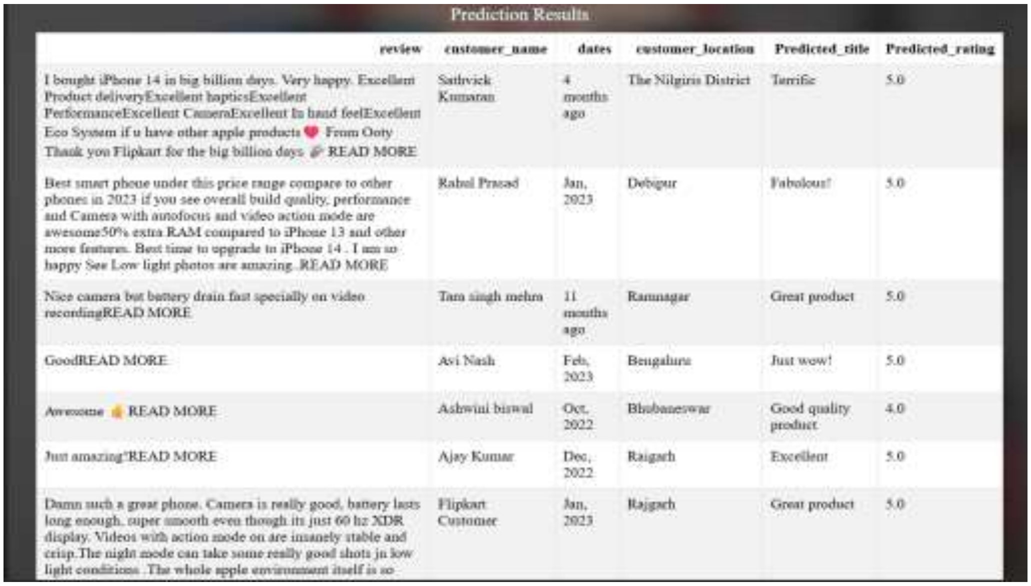


Figure. 6: ROC Curve obtained using ELECTRA _WE _ETC for rating.

Figure 6 presents the One-vs-Rest ROC curves for rating classification (3.0, 4.0, 5.0) using ELECTRA word embeddings across displays the ROC curves for the proposed ETC Classifier, achieving perfect AUC = 1.00 for all three classes (3.0, 4.0, and 5.0), with each curve tightly hugging the top-left corner, confirming flawless class separation and exceptional model performance across the entire rating spectrum.


Prediction Results

review	customer_name	dates	customer_location	Predicted_title	Predicted_rating
I bought iPhone 14 in big billion days. Very happy. Excellent Product deliveryExcellent hapticsExcellent PerformanceExcellent CameraExcellent In hand feelExcellent Eco System if u have other apple products ❤ From Ooty Thank you Flipkart for the big billion days READ MORE	Sathvick Kumaran	4 months ago	The Nilgiris District	Terrific	5.0
Best smart phone under this price range compare to other phones in 2023 if you see overall build quality, performance and Camera with autofocus and video action mode are awesome50% extra RAM compared to iPhone 13 and other more features. Best time to upgrade to iPhone 14 . I am so happy See Low light photos are amazing..READ MORE	Rahul Prasad	Jan, 2023	Debipur	Fabulous!	5.0
Nice camera but battery drain fast specially on video recordingREAD MORE	Tara singh mehra	11 months ago	Ramnagar	Great product	5.0
GoodREAD MORE	Avi Nash	Feb, 2023	Bengaluru	Just wow!	5.0
Awesome 👍 READ MORE	Ashwini biswal	Oct, 2022	Bhubaneswar	Good quality product	4.0
Just amazing'READ MORE	Ajay Kumar	Dec, 2022	Raigarh	Excellent	5.0
Damn such a great phone. Camera is really good, battery lasts long enough, super smooth even though its just 60 hz XDR display. Videos with action mode on are insanely stable and crisp.The night mode can take some really good shots in low light conditions .The whole apple environment itself is so	Flipkart Customer	Jan, 2023	Rajgarh	Great product	5.0

Figure. 7: Predictions page screen

Figure 7 represents the Predictions Page, displaying the results generated after processing the uploaded test dataset. The output table contains multiple columns such as review**,** customer_name, dates, customer_location, Predicted_title, and Predicted_rating, presenting both original data and model-generated outputs. Each row corresponds to an individual entry from the test dataset and its respective predictions. This page allows users to view how the system interprets and evaluates the review content.

Table 1 presents a performance comparison of three classifiers ABC, TTC, and the proposed ETC using ELECTRA word embeddings for review title classification. The table reports four key metrics: Accuracy, Precision, Recall, and F1-Score (all in percentage). The results reveal a stark contrast in model effectiveness: while ABC and TTC achieve low performance (Accuracy: ~23.8% and ~24.4%, respectively), with particularly weak precision and recall, the proposed ETC Classifier delivers near-perfect results (Accuracy: 99.844%, Precision: 99.853%, Recall: 99.846%, F1-Score: 99.845%), demonstrating exceptional capability in accurately predicting diverse sentiment-laden review titles despite high class complexity and potential imbalance.

Table 1: Performance comparison of ELECTRA_WE classifier models for title column.

Algorithm	**Accuracy (%)**	**Precision (%)**	**Recall (%)**	**F1-Score (%)**
ABC	23.810	33.509	23.775	23.013

TTC	24.356	26.064	24.331	22.617
ETC	99.844	99.853	99.846	99.845

Table 2 presents a performance comparison of three classifiers such as ABC, TTC, and the proposed ETC using ELECTRA word embeddings for rating prediction (3.0, 4.0, 5.0). The table evaluates Accuracy, Precision, Recall, and F1-Score (all in percentage). Results show that while ABC and TTC deliver moderate performance (Accuracy: 79.510% and 78.471%, respectively), with balanced but limited discriminative power, the proposed ETC Classifier achieves perfect scores across all metrics (100.000%), confirming its superior ability to accurately predict star ratings with zero misclassification, even in the presence of class imbalance and semantic complexity.

Table 2: Performance comparison of ELECTRA_WE classifier models for Rating column.

Algorithm	Accuracy (%)	Precision (%)	Recall (%)	F1-Score (%)
ABC	79.510	79.390	79.510	79.442
TTC	78.471	78.346	78.471	78.327
ETC	100.000	100.000	100.000	100.000

5. Conclusion

The proposed NLP-based framework successfully automates the extraction of customer sentiment and the prediction of review ratings, delivering a high-precision, scalable alternative to traditional manual or rule-based methods. By utilizing the ELECTRA model for contextual feature extraction, the system captures nuanced linguistic variations, while the K-Means SMOTE algorithm ensures robustness by effectively balancing the dataset. The ETC further enhances the pipeline, providing interpretable and accurate predictions for both titles and ratings. Demonstrating superior performance and lower computational costs compared to existing ABC and TTC methods, the framework is complemented by an intuitive for seamless data visualization and processing. Ultimately, this holistic approach offers organizations a reliable, real-time solution for monitoring brand perception and driving product improvements across diverse e-commerce platforms.

References

[1] Allouch, M.; Azaria, A.; Azoulay, R. Conversational Agents: Goals, Technologies, Vision and Challenges. Sensors 2021, 21, 8448. https://doi.org/10.3390/s21248448

[2] Pota, M.; Ventura, M.; Catelli, R.; Esposito, M. An Effective BERT-Based Pipeline for Twitter Sentiment Analysis: A Case Study in Italian. Sensors 2021, 21, 133. https://doi.org/10.3390/s21010133

[3] García-Madurga, M.-Á.; Gil-Lacruz, A.-I.; Saz-Gil, I.; Gil-Lacruz, M. The Role of Artificial Intelligence in Improving Workplace Well-Being: A Systematic Review. Businesses 2024, 4, 389-410. https://doi.org/10.3390/businesses4030024

[4] Chen, X.; Ibrahim, Z. A Comprehensive Study of Emotional Responses in AI-Enhanced Interactive Installation Art. Sustainability 2023, 15, 15830. https://doi.org/10.3390/su152215830

[5] Alabdan, R. Phishing Attacks Survey: Types, Vectors, and Technical Approaches. Future Internet 2020, 12, 168. https://doi.org/10.3390/fi12100168

[6] Leo, M.; Carcagnì, P.; Mazzeo, P.L.; Spagnolo, P.; Cazzato, D.; Distante, C. Analysis of Facial Information for Healthcare Applications: A Survey on Computer Vision-Based Approaches. Information 2020, 11, 128. https://doi.org/10.3390/info11030128

[7] Borna, S.; Maniaci, M.J.; Haider, C.R.; Gomez-Cabello, C.A.; Pressman, S.M.; Haider, S.A.; Demaerschalk, B.M.; Cowart, J.B.; Forte, A.J. Artificial Intelligence Support for Informal Patient Caregivers: A Systematic Review. Bioengineering 2024, 11, 483. https://doi.org/10.3390/bioengineering11050483

[8] Pu, Q.; Huang, F.; Li, F.; Wei, J.; Jiang, S. Integrating Emotional Features for Stance Detection Aimed at Social Network Security: A Multi-Task Learning Approach. Electronics 2025, 14, 186. https://doi.org/10.3390/electronics14010186

[9] Li, S.; Zhu, B.; Zhang, Y.; Liu, F.; Yu, Z. A Two-Stage Nonlinear User Satisfaction Decision Model Based on Online Review Mining: Considering Non-Compensatory and Compensatory Stages. J. Theor. Appl. Electron. Commer. Res. 2024, 19, 272-296. https://doi.org/10.3390/jtaer19010015

[10] Silva, M.C.; Eugénio, P.; Faria, D.; Pesquita, C. Ontologies and Knowledge Graphs in Oncology Research. Cancers 2022, 14, 1906. https://doi.org/10.3390/cancers14081906