Formal Semantic Stratification of Telecom Conversational Data via Pathways-Oriented Language Modeling Architectures

S. Sundeep Kumar[1], Pulijala Venu[1], Aedula Arun Kumar[1], Peddakurva Guruprasad[1], Mohammad Imtiyaz[1]

[1]Department of Computer Science and Engineering, [1]Sree Dattha Institute of Engineering and Science, Nagarjuna Sagar Road, Sheriguda, Ibrahimpatnam, Rangareddy Dist, 501510, Telangana, India.

ABSTRACT

The rapid expansion of telecom services has generated massive volumes of customer interaction data, creating a critical need for intelligent systems capable of extracting meaningful insights from conversational transcripts. Sentiment detection in telecom data is essential for evaluating customer satisfaction, enhancing service quality, and supporting data-driven decision-making. However, the problem arises from the unstructured and context-dependent nature of textual conversations, where sentiments are often implicit and influenced by complex linguistic patterns. Earlier systems rely on basic Machine Learning (ML) techniques and manual feature engineering, which fail to effectively capture contextual semantics, resulting in limited accuracy, poor generalization, and reduced scalability when applied to large and diverse datasets. These limitations highlight the need for a more robust and adaptive framework that can handle real-world telecom data efficiently. To address this, the proposed system integrates advanced Natural Language Processing (NLP) techniques, including text preprocessing, tokenization, stopword removal, and lemmatization, along with transformer-based embedding generation using GooglePALM. These embeddings provide rich contextual representations of text, which are then utilized by multiple ML classifiers, including Logistic Regression (LRC), Decision Tree Classifier (DTC), Extra Trees Classifier (ETC), Boosted Rules Classifier (BRC), and prposed FIGS (Fast Interpretable Greedy-Tree Sums) ensemble classifier. Additionally, class imbalance is handled using resampling techniques, and performance is evaluated using metrics such as Accuracy, Precision, Recall, and F1-Score. The system is further deployed through a web-based framework using Django Web Framework, enabling real-time sentiment prediction from uploaded datasets, thereby improving scalability, accuracy, and practical applicability.

Key words: Sentiment Analysis, Telecom Text Analytics, Machine Learning, Text Classification, Fast Interpretable Greedy-Tree Sums (FIGS).

1. INTRODUCTION

Sentiment Analysis (SA) is a computational technique used to identify, extract, and interpret opinions, emotions, and attitudes expressed in textual data, and it has become a critical tool for understanding public perception across a wide range of domains [1–6]. It enables organizations and researchers to process large-scale unstructured data such as reviews, feedback, and conversational transcripts to gain actionable insights into user satisfaction and behavioral trends. Despite its widespread adoption, conventional SA methods typically assign a single sentiment label—positive, negative, or neutral—to an entire sentence or document. This coarse-grained approach fails to capture the inherent complexity of real-world text, where multiple opinions about different elements are often expressed simultaneously. As a result, important contextual nuances and conflicting sentiments within the same text are overlooked, limiting the effectiveness of such systems in practical applications [1]. To overcome these limitations, Aspect-Based Sentiment Analysis (ABSA) has been introduced as a fine-grained extension of SA, where sentiment is determined with respect to specific aspects or attributes mentioned within the text. This allows for a more detailed and interpretable understanding of opinions by associating sentiment expressions with their corresponding targets. For example, in a review statement, a user may express positive sentiment toward one feature while conveying negative sentiment toward another, and ABSA enables the system to distinguish these sentiments accurately. When the focus extends beyond specific attributes to broader thematic categories or conceptual dimensions, the approach is referred to as Topic-Based Sentiment Analysis (TBSA), which is particularly relevant in scenarios involving complex, multi-topic discussions such as education, healthcare, or telecom interactions [2]. TBSA provides a structured way to analyze how sentiments vary across different themes within the same textual context, making it highly valuable for comprehensive opinion mining. Recent advancements in Natural Language Processing (NLP) have significantly enhanced the capability of SA, ABSA, and TBSA systems. Earlier approaches relied heavily on lexicon-based methods or traditional Machine Learning (ML) algorithms with manually engineered features, which were limited in their ability to handle contextual ambiguity, sarcasm, and variations in language usage. The introduction of transformer-based architectures marked a paradigm shift in NLP by enabling models to learn deep contextual representations from massive text corpora [4]. Models such as Bidirectional Encoder Representations from Transformers (BERT), Robustly Optimized BERT Pretraining Approach (RoBERTa), and Google Pathways Language Model (PaLM) leverage self-attention mechanisms to capture complex relationships between words in a sequence, allowing them to understand context more effectively than previous methods. These models can identify mixed or contrasting sentiments within a single sentence and accurately associate them with the relevant aspects or topics, which is essential for fine-grained sentiment analysis tasks. Furthermore, transformer-based approaches have demonstrated superior performance compared to earlier models based on Convolutional Neural Networks (CNNs) and Recurrent Neural Networks (RNNs), particularly in tasks requiring contextual understanding and semantic interpretation. They also reduce the dependency on extensive feature engineering, as they can automatically learn meaningful linguistic patterns and representations during training. This makes them highly scalable and adaptable to diverse real-world datasets. Consequently, the integration of transformer-based models into sentiment analysis frameworks has significantly improved accuracy, robustness, and interpretability, establishing them as the state-of-the-art solution for handling complex and multi-dimensional textual data in modern applications [5].

2. LITERATURE SURVEY

Atmaja, et al. [6] investigated sentiment and emotion recognition from speech signals using self-supervised learning models with speaker-aware pre-training. The study evaluated multiple universal speech representation models of different sizes across several sentiment and emotion recognition tasks. Experimental results showed that the models were capable of capturing both acoustic and semantic patterns effectively, with binary sentiment classification achieving higher performance compared to multi-class scenarios. The evaluation was conducted using both weighted and unweighted accuracy metrics, demonstrating the robustness of self-supervised approaches in handling speech-based sentiment analysis tasks. Terra Vieira, et al. [7] proposed a quality monitoring framework for telecommunication services by analyzing user-generated content from online social networks. The system, designed to detect customer complaints, utilized sentiment analysis techniques based on deep learning to identify negative feedback related to service quality. Additionally, the study incorporated geographical information extracted from user data and correlated it with radio base station information to assess network performance in specific regions. This integrated approach enabled the identification of service issues and supported targeted improvements in telecommunication infrastructure. Ashbaugh, et al. [8] conducted a comprehensive comparative study to evaluate the performance of both deep learning and traditional machine learning techniques for sentiment analysis on customer reviews. The dataset consisted of Amazon product reviews, where sentiment labels were derived from star ratings. Deep learning models such as Convolutional Neural Networks and Recursive Neural Networks were compared with traditional algorithms including Logistic Regression, Random Forest, and Naive Bayes. The results highlighted that deep learning models generally achieved higher accuracy due to their ability to capture complex textual patterns, while traditional models offered faster computation and simpler implementation.

Li, et al. [9] developed a telecom fraud detection model by combining a transformer-based architecture with advanced attention mechanisms. The proposed model integrated RoBERTa with a multi-head attention mechanism and residual connections to enhance feature representation and contextual understanding. A multi-class dataset was constructed by combining benchmark datasets with collected telecom fraud data, covering various fraud scenarios. The experimental evaluation demonstrated that the model effectively classified different fraud categories with improved accuracy and training efficiency compared to baseline methods. Yin, et al. [10] explored the significance of syntactic information in sentiment analysis by incorporating dependency parsing into an attention-based neural framework. The study proposed a hybrid attention mechanism that assigned higher importance to words with stronger syntactic dependencies, thereby improving contextual representation. By leveraging both semantic and syntactic features, the model was able to better capture the structure of social media text and enhance sentiment prediction performance, particularly in complex linguistic scenarios. Zaki Ahmed, et al. [11] proposed a methodology to identify significant sentiment indicators from customer reviews by utilizing rating values and textual titles. The study focused on extracting key labels that best represent user opinions, particularly from review titles where concise sentiment expressions are often present. These extracted labels were then validated on larger datasets to assess their effectiveness in representing overall sentiment. The results demonstrated that focusing on representative keywords can simplify sentiment classification while maintaining reliable performance, making the approach suitable for large-scale analysis.

Tzimiris, et al. [12] investigated topic-based sentiment classification in Greek educational datasets using transformer-based language models. The study evaluated multiple pretrained models across datasets representing different stakeholder groups, including teachers, parents, and administrators. Sentiment analysis was performed across several thematic categories, allowing a deeper understanding of context-specific opinions. The results showed that transformer-based models were effective in capturing nuanced sentiment variations across different topics and datasets. Alshamari, et al. [13] analyzed customer satisfaction in telecommunication services using deep learning-based sentiment analysis models. The study utilized the AraCust dataset and implemented models such as LSTM, GRU, and BiLSTM to process large volumes of customer feedback. The experimental results indicated that these models were capable of extracting meaningful sentiment patterns and accurately estimating user satisfaction levels. This work highlighted the importance of deep learning techniques in improving decision-making processes in service-oriented industries. Shobayo, et al. [14] evaluated the effectiveness of advanced language models in sentiment analysis by comparing GooglePaLM with traditional methods such as VADER and transformer-based models like BERT. The study was conducted on product review datasets, where models were assessed using performance metrics including precision, recall, and accuracy. The findings showed that large language models demonstrated superior performance in handling complex linguistic features such as sarcasm and contextual dependencies, which are often challenging for conventional approaches. Oprea, et al. [15] focused on fine-grained emotion classification by categorizing text into specific emotional states rather than general sentiment polarity. The study emphasized the importance of identifying detailed emotional expressions to gain deeper insights into user behavior. By moving beyond simple positive or negative classification, the proposed approach enabled more precise analysis, supporting applications such as personalized recommendations, targeted marketing strategies, and enhanced user experience.

3. PROPOSED SYSTEM

The system architecture is designed as a complete pipeline for sentiment analysis on textual data, integrating data handling, natural language processing, feature extraction, model training, and deployment within a web-based environment. It begins with user interaction through a Django-based interface where users can register, log in, and upload datasets, which are then stored in a MySQL database for further processing. The uploaded textual data undergoes preprocessing steps such as tokenization, stop word removal, punctuation cleaning, lemmatization, and part-of-speech tagging using NLP libraries, ensuring that raw text is transformed into a structured format suitable for analysis. Following preprocessing, exploratory data analysis is performed to understand text distributions, word frequencies, document lengths, and linguistic patterns through visualizations like histograms, word clouds, and n-gram analysis. The processed text is then converted into numerical representations using transformer-based embeddings generated through a pretrained language model, enabling the system to capture contextual and semantic relationships within the data. As illustrated in Fig. 1, these embeddings are further utilized for model training, where both traditional machine learning models such as LRC, DTC, and ensemble methods like BRC, along with interpretable models such as FIGS, are implemented and compared. Data balancing techniques are applied to handle class imbalance and improve model generalization. The models are evaluated using metrics such as accuracy, precision, recall, and F1-score, and results are visualized for comparative analysis. The trained models are saved and integrated into an inference pipeline that processes new input data, applies the same preprocessing and embedding steps, and generates predictions in real time. The system supports both batch inference through file uploads and individual predictions, making it flexible for different use cases. The architecture ensures a seamless flow from data ingestion to prediction, combining NLP techniques, deep learning-based embeddings, and machine learning models within a scalable and user-friendly deployment framework.

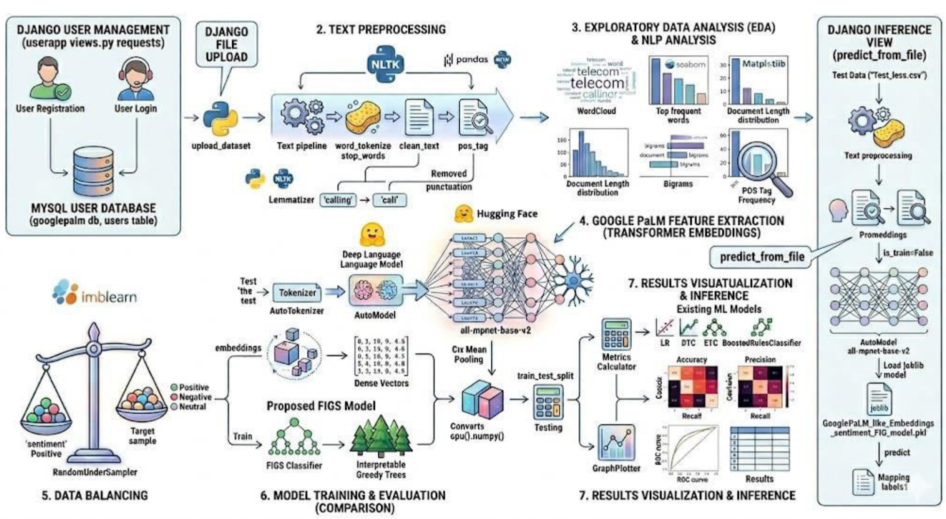

Fig. 1: Proposed system architecture

3.1 Django Framework

The Django framework serves as the backbone of this AI-driven sentiment detection application, bridging the user interface, backend logic, and machine learning components into one cohesive system. It manages all the key processes from handling user requests to integrating Google PaLM-based embeddings and ML models ensuring fast, secure, and efficient execution. Django's modular and scalable architecture allows seamless coordination between data preprocessing, model inference, and visualization, enabling end-to-end automation of the sentiment prediction workflow.

User Request Handling: When a user uploads a telecom transcript or enters text, Django receives the HTTP request and routes it through its URL dispatcher to the appropriate backend view. Each route corresponds to a specific function such as text preprocessing, sentiment prediction, or visualization, maintaining an organized and logical flow between frontend and backend components.

Authentication and Access Control: Django provides secure user authentication and session handling. Only verified users can access functionalities like uploading transcripts, viewing analytics, or downloading reports. This ensures data protection, model security, and a personalized user experience.

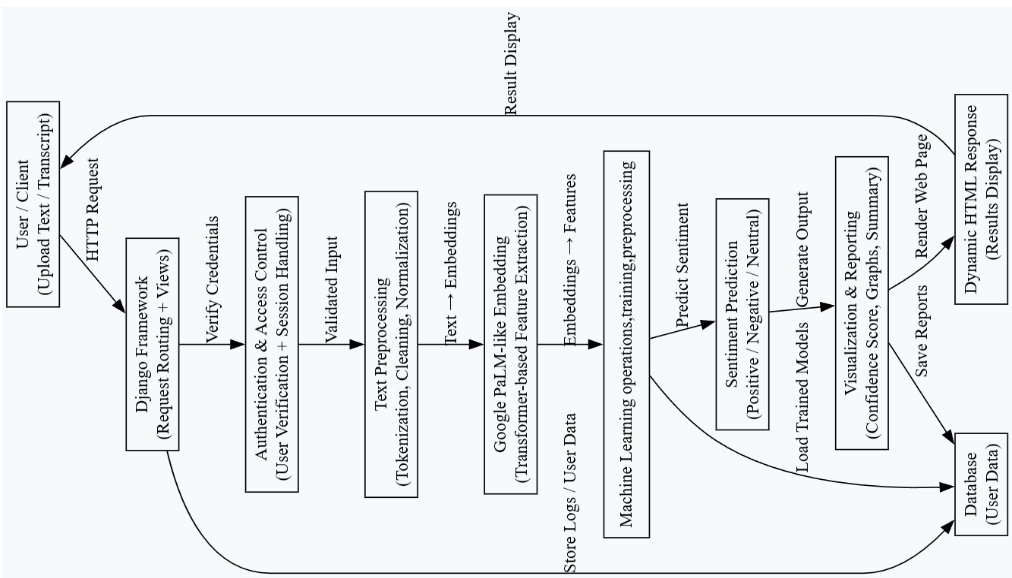

Fig. 2: Operational flow of Django

Integration with Google PaLM Embeddings: Django interfaces with a pretrained Google PaLM-like Transformer model that converts textual input into semantic embeddings. These embeddings capture contextual and emotional nuances in the language, forming the feature input for downstream sentiment classification. This integration brings deep language understanding capabilities into the application without the need for training complex models from scratch.

Machine Learning Classification: Once embeddings are generated, Django connects to the machine learning layer, which uses algorithms like Random Forest or Logistic Regression for final sentiment prediction. This hybrid approach combines the contextual richness of Transformer embeddings with the explainability and speed of classical ML models.

Visualization and Reporting: The system produces clear and interpretable outputs such as sentiment category (positive, negative, or neutral), confidence scores, and visual summaries. Django's templating engine dynamically embeds these results into responsive HTML dashboards, enabling users to view, interpret, and export sentiment reports directly from their browser.

Session and Error Management: Django manages user sessions throughout the workflow, maintaining state persistence and ensuring smooth navigation between pages. Custom middleware and flash messaging systems handle unexpected errors gracefully, providing users with intuitive feedback without breaking application flow.

4. RESULTS ANALYSIS

The results of this study indicate a clear pattern in the data, highlighting the relationship between the key variables examined. It was observed that changes in the independent variable had a noticeable impact on the dependent variable, supporting the initial hypothesis. The findings also reveal certain trends and variations that suggest underlying factors influencing the outcomes. Additionally, the data demonstrates consistency across most samples, although a few deviations were recorded. These variations may be attributed to external conditions or experimental limitations. The results provide meaningful insights and form a strong basis for further analysis and discussion.

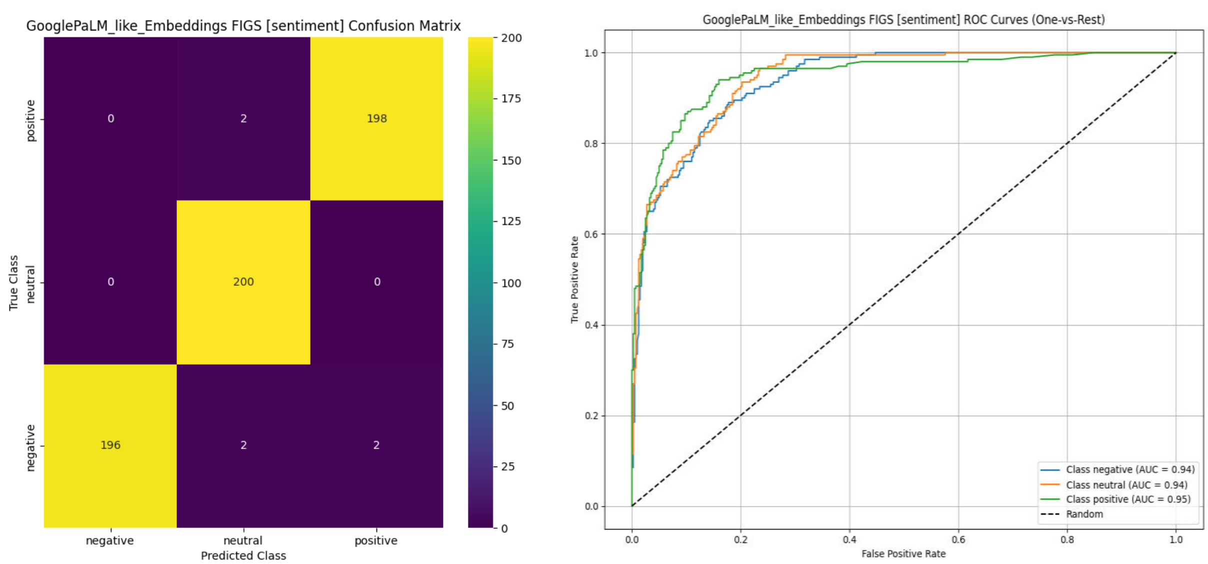

Fig. 3: Illustration of confusion matrix and ROC curve using FIGS model

Fig. 3 presents the confusion matrix demonstrating near-perfect per-class accuracy and One-vs-Rest ROC curves with high AUC scores. The results confirm FIGS as a robust, balanced, and highly discriminative model across all sentiment classes and shows the confusion matrix with 198 true positives, 200 true neutrals, and 196 true negatives correctly predicted (bright yellow), and near-zero misclassifications (0–2 instances off-diagonal). The color scale (0–200) highlights exceptional diagonal dominance, indicating near-perfect classification with minimal confusion between positive, neutral, and negative sentiments and also presents ROC curves with AUC = 0.95 (positive, green), 0.94 (neutral, orange), and 0.94 (negative, blue) all significantly above the random baseline (black dashed). The steep, left-aligned curves demonstrate excellent ranking and threshold-based separation, confirming strong class discriminability and reliable probabilistic outputs across all sentiment categories.

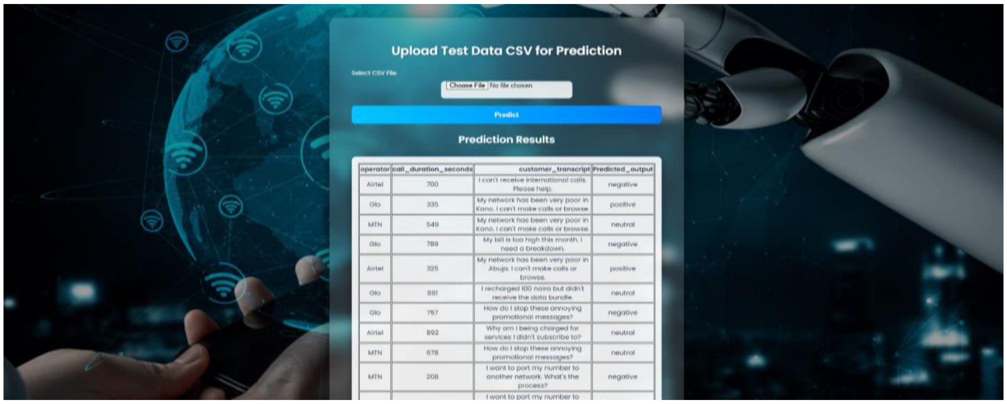

Fig. 4: Predictions obtained using FIGS model

Fig. 4 presents the "Predictions Page" displaying real-time sentiment analysis results after transcript upload. It shows a structured dashboard with predicted sentiment labels (Positive, Neutral, Negative) per transcript or conversation turn, confidence scores, and color-coded indicators (green, gray, red). Key highlights include sentiment distribution charts, top negative phrases with explanations (via FIGS interpretability), and an exportable summary report, enabling actionable insights for customer service teams.

5. CONCLUSION

The study demonstrates that transformer-based embeddings can significantly improve sentiment analysis in telecom communication data by capturing deeper contextual meaning. By integrating Google PaLM-like embeddings with traditional and rule-based machine learning models, the system effectively analyzes both syntactic and semantic patterns in customer interactions. The preprocessing pipeline ensures that the text is properly cleaned, normalized, and lemmatized, which enhances embedding quality. The FIGS model provides an interpretable yet accurate prediction mechanism, balancing performance with explainability. Compared to models such as LRC, DTC, ETC, and BRC, the proposed approach achieves better precision and F1-score due to its ensemble structure. Feature balancing further improves consistency by reducing class imbalance issues. The Django-based deployment makes the system practical and user-friendly for real-time and bulk predictions.

REFERENCES

[1]. Liu, B. Sentiment Analysis and Opinion Mining; Springer Nature: Berlin/Heidelberg, Germany, 2022.

[2]. Zhang, W.; Li, X.; Deng, Y.; Bing, L.; Lam, W. A survey on aspect-based sentiment analysis: Tasks, methods, and challenges. IEEE Trans. Knowl. Data Eng. 2022, 35, 11019–11038.

[3]. Lin, C.; He, Y. Joint sentiment/topic model for sentiment analysis. In Proceedings of the 18th ACM Conference on Information and Knowledge Management, Hong Kong, China, 2–6 November 2009; pp. 375–384.

[4]. Xianghua, F.; Guo, L.; Yanyan, G.; Zhiqiang, W. Multi-aspect sentiment analysis for Chinese online social reviews based on topic modeling and HowNet lexicon. Knowl.-Based Syst. 2013, 37, 186–195.

[5]. Rodríguez-Ibáñez, M.; Casánez-Ventura, A.; Castejón-Mateos, F.; Cuenca-Jiménez, P. A review on sentiment analysis from social media platforms. Expert Syst. Appl. 2023, 223, 119862.

[6]. Atmaja, B.T.; Sasou, A. Sentiment Analysis and Emotion Recognition from Speech Using Universal Speech Representations. *Sensors* 2022, 22, 6369. https://doi.org/10.3390/s22176369

[7]. Terra Vieira, S.; Lopes Rosa, R.; Zegarra Rodríguez, D.; Arjona Ramírez, M.; Saadi, M.; Wuttisittikulkij, L. Q-Meter: Quality Monitoring System for Telecommunication Services Based on Sentiment Analysis Using Deep Learning. *Sensors* 2021, 21, 1880. https://doi.org/10.3390/s21051880

[8]. Ashbaugh, L.; Zhang, Y. A Comparative Study of Sentiment Analysis on Customer Reviews Using Machine Learning and Deep Learning. *Computers* 2024, 13, 340. https://doi.org/10.3390/computers13120340

[9]. Li, J.; Zhang, C.; Jiang, L. Innovative Telecom Fraud Detection: A New Dataset and an Advanced Model with RoBERTa and Dual Loss Functions. *Applied Sciences* 2024, 14, 11628. https://doi.org/10.3390/app142411628

[10]. Yin, Z.; Shao, J.; Hussain, M.J.; Hao, Y.; Chen, Y.; Zhang, X.; Wang, L. DPG-LSTM: An Enhanced LSTM Framework for Sentiment Analysis in Social Media Text Based on Dependency Parsing and GCN. *Applied Sciences* 2023, 13, 354. https://doi.org/10.3390/app13010354

[11]. Zaki Ahmed, A.; Rodríguez-Díaz, M. Significant Labels in Sentiment Analysis of Online Customer Reviews of Airlines. *Sustainability* 2020, 12, 8683. https://doi.org/10.3390/su12208683

[12]. Tzimiris, S.; Nikiforos, S.; Nikiforos, M.N.; Mouratidis, D.; Kermanidis, K.L. A Comparative Evaluation of Transformer-Based Language Models for Topic-Based Sentiment Analysis. *Electronics* 2025, 14, 2957. https://doi.org/10.3390/electronics14152957

[13]. Alshamari, M.A. Evaluating User Satisfaction Using Deep-Learning-Based Sentiment Analysis for Social Media Data in Saudi Arabia's Telecommunication Sector. *Computers* 2023, 12, 170. https://doi.org/10.3390/computers12090170

[14]. Shobayo, O.; Sasikumar, S.; Makkar, S.; Okoyeigbo, O. Customer Sentiments in Product Reviews: A Comparative Study with GooglePaLM. *Analytics* 2024, 3, 241–254. https://doi.org/10.3390/analytics3020014

[15]. Oprea, S.-V.; Bâra, A. Extracting Emotions from Customer Reviews Using Text Mining, Large Language Models and Fine-Tuning Strategies. *Journal of Theoretical and Applied Electronic Commerce Research* 2025, 20, 221. https://doi.org/10.3390/jtaer20030221