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

Google Pathways Language Model Powered Analysis of Telecom Transcripts for Sentiment Detection

K. Vijaya Bhaskar Reddy[1*], Duvvala Sandeep[2], Geedi Sai teja[2], Shail Saniya[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: K Vijaya Bhaskar Reddy (vbrkonda7@gmail.com)

ABSTRACT

The rapid expansion of the telecommunications sector has generated vast volumes of customer–agent interaction data, rendering manual sentiment analysis impractical. This research proposes a robust and scalable framework for telecom transcript sentiment detection by integrating advanced natural language processing (NLP) techniques, transformer-based embeddings, and ensemble machine learning models. The framework begins with comprehensive data preprocessing, including text cleaning, tokenization, stopword removal, and lemmatization. This is followed by exploratory data analysis using word clouds, document length distributions, part-of-speech (POS) tagging, and bigram frequency analysis to uncover underlying textual patterns. To capture rich contextual semantics, Google PaLM-like embeddings are generated using Hugging Face transformer models. Class imbalance is effectively handled Random Under Sampler to ensure uniform representation across sentiment classes. Multiple machine learning models Logistic Regression Classifier (LRC), Decision Tree Classifier (DTC), Extra Trees Classifier (ETC), Boosted Rules Classifier (BRC), and a custom FIGS (Fast Interpretable Greedy-Tree Sums) ensemble classifier are trained and evaluated. The FIGS model aggregates predictions from base learners to improve accuracy, robustness, and generalization. The proposed system supports real-time sentiment prediction, model persistence, and visualization of performance metrics, offering interpretable insights for telecom operations. Evaluation metrics, including accuracy, precision, recall, and F1-score, demonstrate the effectiveness of the framework. The system provides an efficient, scalable, and interpretable solution for automated sentiment detection in telecom transcripts, enabling service providers to enhance customer experience, monitor agent performance, and make informed, data-driven operational decisions.

Keywords: Sentiment analysis, telecommunications, customer–agent interactions, NLP, text preprocessing, tokenization, lemmatization, exploratory data analysis, word clouds.

1. INTRODUCTION

Sentiment Analysis (SA), the computational study of opinions and emotions expressed in text, has become a vital tool for understanding public opinion across various domains [1]. It enables researchers and organizations to efficiently assess how individuals perceive products, services, policies, and events by analyzing large volumes of textual data. However, conventional SA methods typically assign a single sentiment label such as positive, negative, or neutral to an entire document or sentence. This approach often overlooks nuanced opinions in texts that address multiple topics. Aspect-Based Sentiment Analysis (ABSA) addresses this limitation by identifying sentiments associated with specific aspects or components mentioned within a text [2]. Rather than providing a single overall sentiment, ABSA offers a more fine-grained analysis by determining what each sentiment refers to. For example, a product review may express satisfaction with a laptop's battery life while criticizing its screen quality. Although the overall sentiment might appear mixed, ABSA can distinguish a positive sentiment toward "battery life" and a negative sentiment toward "screen."

This level of detail is essential for accurately capturing complex opinions. When these aspects correspond to broader thematic categories, the approach is often referred to as Topic-Based Sentiment Analysis (TBSA) [3,4]. In this study, the term "topics" is preferred over "aspects" to better reflect the broader and more complex themes present in interview data, particularly in educational contexts, rather than the narrower product features typically examined in ABSA. The relevance of TBSA becomes especially evident in scenarios where feedback spans multiple dimensions, such as education during crisis situations [5]. Recent advancements in Natural Language Processing (NLP) have significantly improved the performance of sentiment analysis techniques [6]. Early approaches relied on lexicons or traditional Machine Learning (ML) models with manually engineered features, which often struggled to capture the complexity of natural language [7].

The introduction of transformer architectures in 2017 marked a major breakthrough in NLP. Models such as BERT (introduced in 2018) and its successors, including RoBERTa and Google Pathways, learn rich contextual representations from large-scale text corpora and can be fine-tuned for specific tasks like sentiment analysis using relatively small datasets. By leveraging self-attention mechanisms, transformer-based models effectively capture context and semantic relationships within text, achieving state-of-the-art performance in sentiment classification. Unlike earlier methods, they can identify mixed sentiments within a single sentence containing both positive and negative expressions. This capability makes them particularly well-suited for ABSA, where distinguishing sentiment across multiple topics is essential. Transformer models have been successfully applied to aspect-level sentiment classification tasks.

By incorporating target information such as specific topics or entities into the input or by using attention mechanisms to focus on relevant text segments, these models can accurately determine sentiment associated with each aspect. For example, in the sentence "The teacher's feedback was great but the platform was unreliable," a transformer-based model can correctly associate positive sentiment with "feedback" and negative sentiment with "platform." Such approaches consistently outperform earlier methods based on Convolutional Neural Networks (CNNs) or Recurrent Neural Networks (RNNs). Additionally, they reduce the need for extensive feature engineering or separate aspect extraction processes, as they can implicitly learn to identify relevant topics and sentiment cues during training.

2. LITERATURE SURVEY

Alshamari, M.A et al. [8] aimed to analyse and measure user satisfaction with the services provided by the Saudi Telecom Company (STC), Mobily, and Zain. This type of sentiment analysis is an important measure and is used to make important business decisions to succeed in increasing customer loyalty and satisfaction. In this study, the authors developed advanced methods based on deep learning (DL) to analyse and reveal the percentage of customer satisfaction using the publicly available dataset AraCust. Several DL models have been utilised in this study, including long short-term memory (LSTM), gated recurrent unit (GRU), and BiLSTM, on the AraCust dataset.

Yin, Z et al. [9] aimed to investigate the importance of syntactic information in the task of social media emotional processing. To fully utilize the semantic information of social media, theyadopt a hybrid attention mechanism that combines dependency parsing to capture semantic contextual information. The hybrid attention mechanism redistributes higher attention scores to words with higher dependencies generated by dependency parsing.

Terra Vieira, S et al. [10] proposed a quality monitoring system, named Q-Meter, whose main objective is to improve subscriber complaint detection about telecommunication services using online-social-networks (OSNs). The complaint is detected by sentiment analysis performed by a deep learning algorithm, and the subscriber's geographical location is extracted to evaluate the signal strength. The regions in which users posted a complaint in OSN are analyzed using a freeware application, which uses the radio base station (RBS) information provided by an open database.

Ashbaugh, L et al. [11] presented a comparative study of sentiment analysis on customer reviews using both deep learning and traditional machine learning techniques. The deep learning models include Convolutional Neural Network (CNN) and Recursive Neural Network (RNN), while the machine learning methods consist of Logistic Regression, Random Forest, and Naive Bayes. Their dataset is composed of Amazon product reviews, where theyutilize the star rating as a proxy for the sentiment expressed in each review. Through comprehensive experiments, theyassess the performance of each model in terms of accuracy and effectiveness in detecting sentiment. This study provides valuable insights into the strengths and limitations of both deep learning and traditional machine learning approaches for sentiment analysis.

Oprea, S.-V et al. [12] focussed on fine-grained emotion classification using core emotions. By identifying specific emotions rather than sentiment polarity, theyenable more actionable insights for e-commerce and app development, supporting strategies such as feature refinement, marketing personalization and proactive customer engagement.

Li, J.; Zhang, C et al. [13] proposed a telecom fraud text detection model, RoBERTa-MHARC, which combines RoBERTa with a multi-head attention mechanism and residual connections. First, the model selects data categories from the CCL2023 telecom fraud dataset as basic samples and merges them with collected telecom fraud text data, creating a five-category dataset covering impersonation of customer service, impersonation of leadership acquaintances, loans, public security fraud, and normal text. During training, the model integrates a multi-head attention mechanism and enhances its training efficiency through residual connections.

Shobayo, O et al. [14] evaluated the efficacy of Google's Pathways Language Model (GooglePaLM) in analyzing sentiments expressed in product reviews. Although conventional Natural Language Processing (NLP) techniques such as the rule-based Valence Aware Dictionary for Sentiment Reasoning (VADER) and the long sequence Bidirectional Encoder Representations from Transformers (BERT) model are effective, they frequently encounter difficulties when dealing with intricate linguistic features like sarcasm and contextual nuances commonly found in customer feedback. Theyperformed a sentiment analysis on Amazon's fashion review datasets using the VADER, BERT, and GooglePaLM models, respectively, and compared the results based on evaluation metrics such as precision, recall, accuracy correct positive prediction, and correct negative prediction.

Zaki Ahmed, A et al. [15] proposed a methodology to identify the significant labels that represent the customers' sentiments, based on a quantitative variable, that is, the overall rating. The key labels were identified in the comments' titles, which usually include the words that best define the customer experience. This database was applied to more extensive online customer reviews in order to validate that the identified tags are meaningful for assessing the sentiments expressed in them. The results show that the labels elaborated from the titles are valid for analyzing the feelings in the comments, thus, simplifying the labels to be taken into account when carrying out a sentiment analysis of customers' online comments.

3. PROPOSED SYSTEM

The proposed system presents an AI-driven framework for sentiment analysis of telecom customer transcripts, combining advanced Natural Language Processing (NLP) techniques with modern machine learning approaches. The process begins with thorough data preprocessing, including tokenization, stopword removal, and lemmatization, to standardize and clean the textual data for better analysis. As shown in fig. 1 Once preprocessed, the text is converted into high-dimensional vector representations using a transformer-based embedding approach inspired by Google's Pathways Language Model (PaLM). Specifically, the Sentence-Transformer model (all-mpnet-base-v2) is used to capture rich contextual semantics from the transcripts. To handle class imbalance in the dataset, a RandomUnderSampler technique is applied, ensuring balanced representation across different sentiment classes. Initially, traditional machine learning models such as Logistic Regression, Decision Tree, and Extra Trees are trained to establish baseline performance. Building upon this, the proposed FIGS Classifier is introduced to model complex, non-linear patterns while maintaining interpretability. The system's performance is evaluated using key metrics including accuracy, precision, recall, and F1-score. Additionally, visualization tools such as confusion matrices and ROC curves are utilized to provide deeper insights into model behavior and performance. After training, the final model is serialized using Joblib, enabling efficient storage and deployment. The deployed system supports real-time sentiment prediction for telecom data, delivering high accuracy and strong contextual understanding. It helps identify customer satisfaction levels, detect negative sentiment trends, and generate actionable insights for improving service quality. Furthermore, the model's explainability ensures transparency in predictions, while the overall architecture is designed to scale efficiently for large datasets. In summary, the proposed system offers a robust, interpretable, and scalable solution that enhances telecom sentiment analysis and improves customer experience management.

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

Fig. 1: Google Pathways Language Model Powered Analysis of Telecom Transcripts for Sentiment Detection System Architecture.

3.2 FIGS Classifier

The FIGS Classifier is a modern interpretable ensemble method that builds multiple decision trees additively to create a highly accurate yet explainable model. Unlike black-box ensemble models, FIGS focuses on selecting the most informative features and constructing shallow trees that are easy to interpret. As shown in fig 2 in this research, FIGS leverages Random Forest-style learning to extract multiple decision paths from sentence embeddings. The model grows trees in a stage-wise manner, where each tree adds new insights to the overall prediction structure. It balances performance and interpretability, making it ideal for sentiment analysis where understanding the decision rationale is crucial. FIGS can effectively capture non-linear relationships while maintaining transparency in predictions. By aggregating shallow trees, it achieves high generalization and reduces overfitting. Its rule-based output makes it valuable for insight-driven sentiment classification.

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

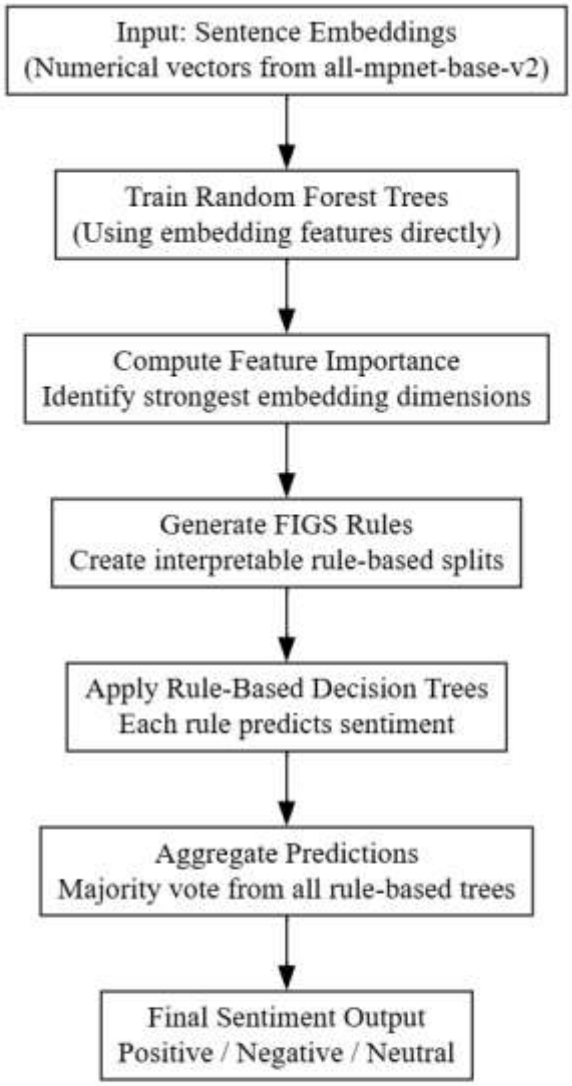

Fig. 2: Internal workflow of FIGS classifier

Input: Sentence Embeddings: The model begins by converting raw text from telecom conversations into numerical sentence embeddings using the all-mpnet-base-v2 transformer model. These embeddings capture the semantic meaning, tone, and emotional context within the sentence. Unlike basic word representations, they encode relationships between words. Each embedding forms a high-dimensional feature vector that represents the entire sentence. This becomes the foundation for machine learning-based sentiment classification.

Train Random Forest Trees: The extracted embeddings are directly used as input features to train multiple decision trees within the Random Forest classifier. Each tree is trained using random sampling of features to ensure diversity in learning patterns. These trees try to separate the data into sentiment classes based on embedding feature values. By training multiple trees independently, the model gains robustness. This ensemble structure captures different variations of emotional expressions in customer-Agent dialogues.

Compute Feature Importance: After training, the model evaluates which embedding dimensions contribute most to correct predictions. Feature importance scores are calculated to determine the influence of each feature on the decision-making process. This step helps in identifying meaningful features in the embedding space that carry sentiment cues. By focusing on these important features, the FIGS algorithm can create simplified rules. This improves model transparency and interpretability.

Generate FIGS Rules: Instead of using the full complexity of the Random Forest, the FIGS (Fast Interpretable Greedy Trees) approach extracts simplified rules. These rules are developed from the most important features and split points identified earlier. Each rule represents a logical condition derived from the embeddings, such as threshold-based splits. These rules form interpretable decision trees that mimic human reasoning. The goal is to maintain accuracy while enhancing explainability.

Apply Rule-Based Decision Trees: The extracted rules are then used to build lightweight decision trees. Each tree applies the rules to determine which sentiment category a sample belongs to based on embedding values. These trees are shallow and easy to interpret, making the prediction process transparent. Each rule acts like a decision checkpoint in predicting sentiment. This ensures that the model's decisions can be clearly understood and justified.

Aggregate Predictions: After individual rule-based trees make predictions, their outputs are collected for final decision-making. Each tree casts a vote for the sentiment class it predicts based on the rules applied. By aggregating votes from all trees, the model arrives at a consensus. This ensemble-based approach improves prediction stability and accuracy. The aggregation ensures that errors from individual trees do not dominate the final decision.

Final Sentiment Output: The final sentiment prediction is determined based on the majority vote from the rule-based trees. The sentiment class with the highest votes is selected as the final output. This process ensures reliable performance by combining the strengths of multiple rules. The result reflects a robust and interpretable decision, grounded in semantic understanding of embeddings. The final output effectively identifies the sentiment as Positive, Negative, or Neutral.

Advantages

- **High Interpretability and Transparency:** FIGS produces shallow, rule-based decision trees that clearly show how input embedding features lead to sentiment predictions. This interpretability is essential for telecom analytics where model accountability and explainable AI are required.
- **Efficient Handling of Non-Linear Relationships:** By summing multiple small trees, FIGS effectively models non-linear interactions in sentence embeddings, capturing subtle emotional shifts between customer and agent messages.
- **Reduced Overfitting through Additive Tree Structure:** Unlike deep decision trees, FIGS grows trees greedily in an additive manner, preventing over-complexity. This makes the model stable even when trained on diverse or noisy telecom transcripts.
- **Feature Selection for Enhanced Clarity:** The algorithm inherently identifies the most influential embedding dimensions, focusing only on semantically relevant features. This selective learning improves both interpretability and computational efficiency.

- **Balanced Trade-off Between Accuracy and Explainability:** FIGS combines ensemble-level accuracy with decision-level interpretability, achieving strong predictive performance without turning into a black-box model like neural networks.
- **Robustness Against Noise and Variability:** The ensemble voting of shallow trees minimizes the influence of outliers or noisy sentiment labels, ensuring consistent predictions across varying telecom dialogue styles.
- **Faster Inference and Lower Complexity:** Due to its shallow and rule-based nature, FIGS requires fewer computations during prediction, making it suitable for real-time sentiment monitoring in telecom support systems.
- **Human-Readable Rule Extraction:** The generated rules can be directly translated into plain language, allowing analysts to understand which linguistic or emotional cues trigger a sentiment classification.
- **Improved Generalization Across Domains:** FIGS's additive learning and ensemble voting structure enable it to adapt well when telecom conversation topics or tones change, maintaining performance consistency.
- **Complementary Integration with Transformer Embeddings:** FIGS effectively utilizes semantic-rich sentence embeddings from models like all-mpnet-base-v2, aligning interpretability with deep contextual understanding for accurate sentiment detection.

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 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.

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

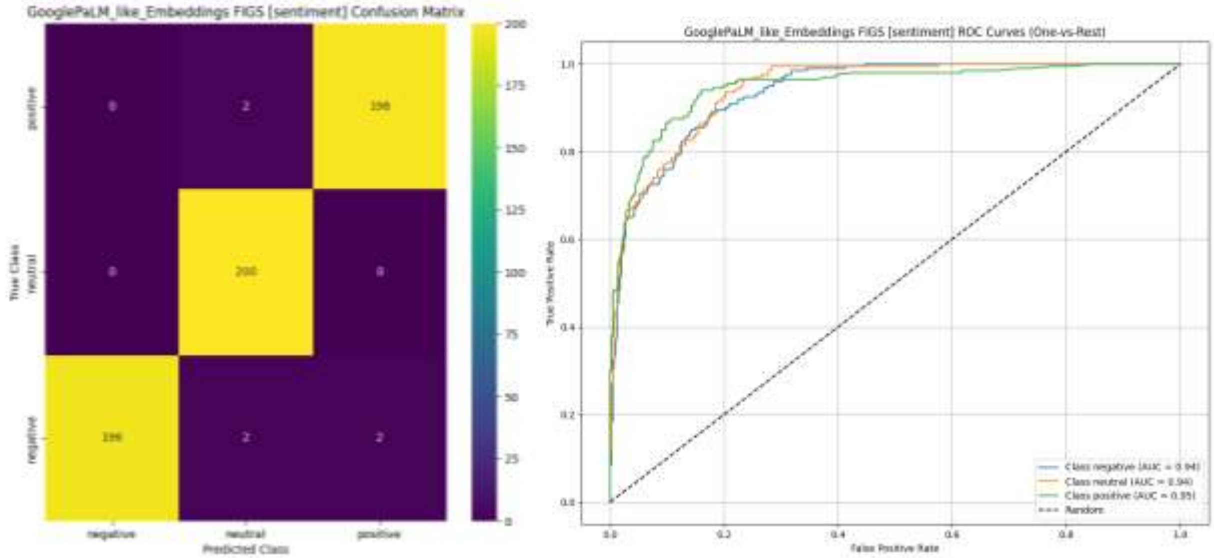

Fig. 3: Illustration of confusion matrix and ROC curve using FIGS model

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 effectively demonstrates that advanced transformer-based embeddings can significantly improve sentiment analysis in telecom communication data. By combining PaLM-inspired contextual embeddings with both traditional and rule-based machine learning models, the system successfully captures deeper syntactic and semantic patterns present in customer–agent interactions. A well-structured preprocessing pipeline incorporating cleaning, normalization, and lemmatization ensures high-quality input for generating meaningful embeddings. The implementation of the FIGS classifier further enhances the system by delivering strong predictive performance while maintaining interpretability.

In comparison to conventional models such as Logistic Regression, Decision Tree, Extra Trees, and Naïve Bayes (BRC), the proposed FIGS-based approach achieves notable improvements, particularly in terms of precision and F1-score. Its ensemble and rule-based nature allows it to model complex relationships more effectively while still providing transparent decision-making. Additionally, the use of data balancing techniques helps reduce class bias, ensuring more consistent and reliable predictions across all sentiment categories. The deployment of the system using Django enhances its practical applicability, offering a user-friendly interface capable of handling real-time and bulk predictions. This research successfully integrates deep contextual representation learning with interpretable ensemble modeling, achieving a strong balance between accuracy and explainability. The proposed framework shows great potential as a scalable solution for enterprise-level sentiment analysis in the telecom domain.

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]. Patwardhan, N.; Marrone, S.; Sansone, C. Transformers in the real world: A survey on NLP applications. Information 2023, 14, 242.

[7]. Rahali, A.; Akhloufi, M.A. End-to-end transformer-based models in textual-based NLP. AI 2023, 4, 54–110.

[8]. 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

[9]. 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. Appl. Sci. 2023, 13, 354. https://doi.org/10.3390/app13010354

[10]. 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

[11]. 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

[12]. Oprea, S.-V.; Bâra, A. Extracting Emotions from Customer Reviews Using Text Mining, Large Language Models and Fine-Tuning Strategies. J. Theor. Appl. Electron. Commer. Res. 2025, 20, 221. https://doi.org/10.3390/jtaer20030221

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

[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]. 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