

International Journal of Engineering Research and Science & Technology

www.ijerst.org

ISSN : 2319-5991

Vol. 22 No. 2(1) (2026)



ijerst.editor@gmail.com
editor@ijerst.com

Research Paper

Scalable Customer Support Ticket Analysis Using Hybrid NLP and Ensemble-Based Classification Models

Vadi Sravani¹, Dayakar Thalla^{2*}, Vajjeeru Anusree¹, Mahammad Mubin¹, Chinnapaka Eshwar¹,
Palakurthi Meghana¹

¹UG Student, ²Assistant professor, ^{1,2}Department of Computer Science and Engineering (AI&ML)

^{1,2}Vaagdevi Engineering College, Bollikunta, Warangal, 506005, Telangana, India

*Correspondence: Dayakar Thalla (dayasri.thalla@gmail.com)

ABSTRACT

Customer support centers generate large volumes of textual data in the form of support tickets, reflecting customer issues, service priorities, and satisfaction levels. Analyzing this unstructured data is essential for improving service quality and operational efficiency. However, traditional approaches rely on manual processing or basic machine learning techniques using shallow features such as bag-of-words and Term Frequency–Inverse Document Frequency (TF-IDF), which fail to capture contextual meaning and semantic relationships. These limitations result in reduced accuracy, poor scalability, and challenges in handling multiple prediction tasks, ultimately affecting timely customer service. This study presents an advanced analytical framework that integrates natural language processing with machine learning techniques to address these challenges. The textual data is preprocessed using standard NLP techniques, including tokenization, stopword removal, and lemmatization, to clean and normalize input. Transformer-based feature extraction is applied using eXtreme Language Network (XLNet) to generate contextual embeddings that capture deeper semantic information. These embeddings provide a rich numerical representation, enabling improved analysis and prediction. Multiple machine learning models, including Quadratic Discriminant Analysis (QDA), Linear Discriminant Analysis (LDA), Histogram-Based Gradient Boosting (HGB), Stochastic Gradient Descent (SGD), and Nearest Centroid (NC), are trained to perform multi-target classification. The system predicts key attributes such as ticket priority, customer satisfaction rating, and resolution status. Additionally, exploratory data analysis techniques are used to visualize patterns and trends. This integrated approach enhances prediction accuracy, supports scalability, and enables efficient, data-driven customer support management.

Key words: Customer Support Analytics, Natural Language Processing (NLP), Text Classification, XLNet, Contextual Embeddings, Machine Learning, Multi-Target Prediction, Support Ticket Analysis, Gradient Boosting.

1. INTRODUCTION

The field of ticket management and monitoring is undergoing a rapid transformation driven by the emergence of Large Language Models (LLMs) and Generative AI. These advanced technologies enable intelligent and context-aware automation of critical processes such as ticket classification, prioritization, and resolution recommendation by utilizing historical support data and established Standard Operating Procedures (SOPs). As a result, traditionally manual and effort-intensive tasks involved in analyzing and resolving support tickets are significantly streamlined, leading to reduced response times and improved efficiency for support teams. However, the adoption of AI-driven ticket management systems presents several practical challenges. Organizations must address issues such as

transitioning skilled personnel to new intelligent platforms, preserving domain-specific knowledge and intellectual assets, and accurately migrating essential operational data, including Service Level Agreements (SLAs) and expert-defined policies. Overcoming these barriers requires flexible and scalable solutions. To this end, our work introduces a modular, application-independent microservices architecture that supports key functionalities such as ticket clustering, prioritization, resolution assistance, data ingestion, and visualization. This design enables gradual and customizable integration of AI capabilities into existing workflows.

In real-world IT Service Management (ITSM) environments, effective incident handling extends beyond textual data analysis. It requires incorporating spatial information such as system topology and infrastructure relationships, as well as temporal patterns related to incident occurrence and progression. Additionally, managing data consistency and provenance across heterogeneous sources is crucial. Given the continuously evolving nature of IT environments, clustering approaches must also be adaptive, allowing dynamic adjustment of cluster structures as new data emerges. To address these complexities, this research proposes a context-aware framework that combines semantic similarity techniques with spatial and temporal insights for enhanced ticket clustering. By capturing the multifaceted characteristics of IT operations, the proposed approach improves the accuracy of incident grouping and prioritization. This leads to more effective diagnostics, better resource utilization, enhanced system stability, and improved customer satisfaction. Ultimately, the study advances the capabilities of ITSM ticket analysis by integrating LLM-based semantic intelligence with real-world operational context.

2. RELATED WORK

2.1 Incident Management Systems

Incident management systems are crucial in IT service management (ITSM), aiming to restore normal operations swiftly and minimize business disruptions. Traditional systems, as outlined by the ITIL framework [1], emphasize structured processes and manual ticket handling. However, these systems often struggle with high volumes of incidents and the complexity of manual categorization and resolution, leading to inefficiencies and delays [2]. Ticket categorization and routing are critical for ensuring swift resolution, customer satisfaction, high productivity, and adherence to Service-Level Agreements (SLAs), which often require issues to be resolved within a set timeframe [3]. However, improper routing can lead to inefficient assignments and increased costs for both customers and service providers [4]. Ticket Automation (TA) systems aim to minimize the steps from ticket submission to resolution. A key function of these systems is the accurate classification of incoming tickets, essential for quickly addressing customer requests. This need has grown with the increasing volume of support tickets, particularly in IT companies, highlighting the importance of efficient automated resolution systems [5].

2.2 NLP techniques in Incident Management

Demonstrated an effective method to enhance classification accuracy by incorporating information from the labels' hierarchical structure into the classifier [6]. They focus on classifying tickets using a pre-trained BERT language model, highlighting how this approach can refine categorization processes [7]. Their study further reveals that the choice of ticket embedding strategy significantly influences classification metrics, underlining the importance of embedding selection in achieving optimal classification performance [8]. The classification of incidents in ticketing systems typically involves assigning a priority to assess the urgency and a type to gauge the importance of the incident, such as distinguishing between an information request and an incident report [9].

2.3 Clustering Algorithms for Ticket Grouping

Clustering algorithms are vital for grouping similar incident tickets, facilitating more efficient resolution processes. Jain et al. [11] provided a comprehensive review of clustering techniques, emphasizing the importance of similarity measures. Huang et al. [10] proposed an adaptive clustering method for IT service management that considers temporal patterns and incident similarities, showcasing the benefits of dynamic clustering approaches in improving incident handling efficiency [12].

2.4 Hierarchical Clustering and Embedding Techniques

Hierarchical clustering is particularly relevant for incidents involving networked devices, where understanding the network topology is crucial. Murtagh and Contreras [13] reviewed hierarchical clustering algorithms, detailing their applications and advantages. Embedding techniques, such as those provided by Hugging Face BGE models [14], have enhanced the capability to capture semantic and hierarchical relationships in text data. Introduced BERT, a transformer-based model that significantly improved NLP tasks, paving the way for subsequent embedding models used in various applications, including incident management [15].

3. PROPOSED SYSTEM

The proposed methodology presents a structured and intelligent approach for analyzing customer support data using advanced natural language processing and machine learning techniques. It follows a systematic pipeline that begins with dataset acquisition, preprocessing, and textual feature extraction, followed by multi-model classification and predictive analysis. Transformer-based feature extraction is utilized to capture the contextual and semantic meaning of textual data, enabling more accurate representation compared to traditional methods. Multiple machine learning models are employed to perform classification tasks, allowing comparative evaluation and improved decision-making. The methodology also integrates visualization and evaluation mechanisms to assess model performance, as shown in figure 2. while a graphical user interface ensures seamless interaction, real-time analysis, and user-friendly operation.

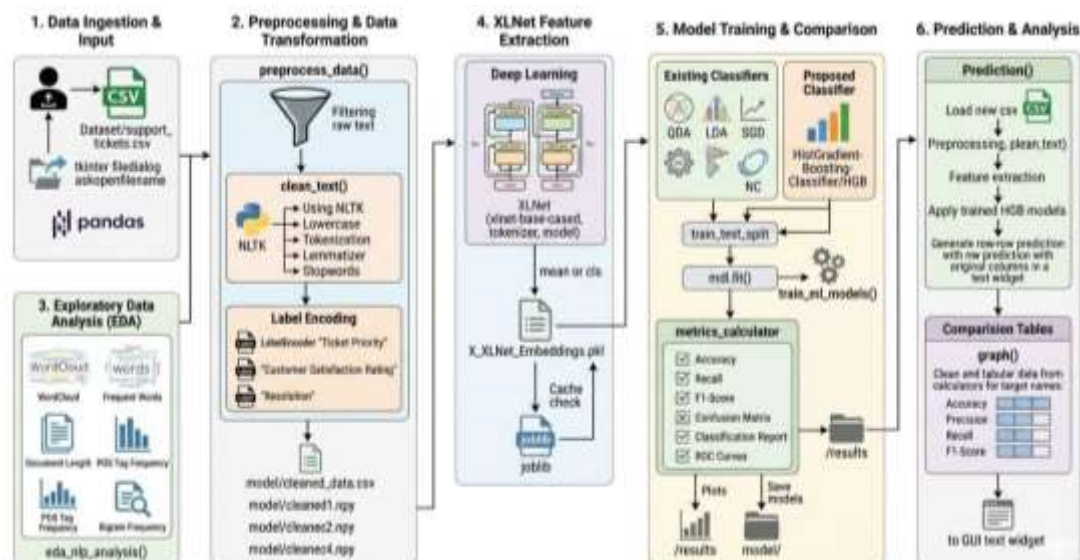


Figure 2: Proposed system architecture.

User Interface (Tkinter GUI)

- The user interacts with the system through a dedicated graphical interface built using Tkinter, providing a centralized dashboard for support analytics.

- It provides a comprehensive suite of operational modules, including dataset upload, NLP preprocessing, EDA visualization, feature extraction, and model training.
- All user actions from triggering a prediction to comparing results—are captured through intuitive button clicks and translated into backend function calls.
- The interface is designed to bridge the gap between complex machine learning workflows and the end-user's need for actionable insights.

Core Application Layer (Main GUI Controller)

- This layer acts as the primary "command center," controlling the entire analytical workflow and synchronizing all sub-modules.
- It manages the flow of user inputs and orchestrates the sequence of preprocessing, feature extraction, and model inference.
- The controller ensures smooth data communication between the underlying CSV datasets, the trained models, and the final output displays.
- Its modular design allows for the isolation of specific tasks, ensuring the system remains stable during intensive computational processes.

Dataset Input (CSV File)

- The system accepts customer support telemetry in CSV format as the foundational input for the analysis.
- The dataset comprises a rich mix of textual and categorical attributes, such as raw ticket descriptions, historical priority levels, and resolution logs.
- This raw data serves as the primary source for identifying patterns in customer sentiment and support efficiency.
- Once ingested, the data is automatically routed to the preprocessing module for linguistic refinement.

Data Preprocessing (NLP Processing)

- Raw support text undergoes a multi-stage cleaning and normalization process using established NLP techniques.
- The pipeline includes standard operations such as case standardization (lowercasing), tokenization, stopword removal, and POS-based lemmatization.
- These steps strip away linguistic noise, transforming unstructured text into a clean, structured format optimized for deep learning.
- This module ensures that the subsequent feature extraction stage focuses only on the most semantically relevant words.

Feature Extraction (XLNet Embeddings)

- The refined text is converted into high-dimensional numerical feature vectors using XLNet-based transformer embeddings.
- Tokenized inputs are passed through the pretrained XLNet model to generate context-aware representations that capture the subtle nuances of support language.
- Mean pooling is applied to the transformer outputs to obtain fixed-length vectors that encapsulate the core semantic information of each ticket.
- This approach allows the system to understand the "meaning" behind a customer's query rather than just matching keywords.

Feature Storage (Model Cache)

- To optimize performance, extracted XLNet features are persisted using the joblib library.
- This caching mechanism eliminates the need to re-run the intensive transformer model during repeated experiments or model tuning.
- By storing these "frozen" feature vectors, the system significantly reduces the computational overhead of the training cycle.
- This ensures that the framework remains responsive even when working with large volumes of support tickets.

Existing Models (QDA, LDA, NC, SGD)

- The system establishes a performance baseline by providing the XLNet features to a diverse group of traditional machine learning models.
- QDA and LDA are used to explore quadratic and linear decision boundaries for class separation.
- NC provides a distance-based perspective, while SGD offers a scalable linear approach for large datasets.
- These models generate independent predictions for ticket priority and satisfaction, serving as a benchmark for the proposed hybrid.

Proposed Model (HGB Classifier)

- The HGB classifier serves as the primary engine for high-accuracy prediction.
- It builds an ensemble of decision trees iteratively, specifically optimized for the high-dimensional vectors produced by XLNet.
- The HGB architecture is engineered for speed and memory efficiency, handling large-scale support data with significantly better accuracy than traditional learners.
- This model effectively captures the complex, non-linear relationships between ticket text and support outcomes.

Model Training and Evaluation

- The dataset is split into training and testing subsets to facilitate a rigorous validation of the HGB and baseline models.
- The system computes a full suite of performance metrics, including Accuracy, Precision, Recall, and F1-score, for each predicted target.
- Confusion matrices and ROC curves are generated dynamically to provide deep insights into the model's ability to distinguish between different priority and satisfaction levels.
- This evaluative data is critical for ensuring the system provides reliable guidance for support desk decision-making.

Prediction Module

- The inference engine applies the learned knowledge to new, incoming support tickets in real-time.
- New data is automatically routed through the exact same NLP preprocessing and XLNet extraction pipeline to maintain feature consistency.
- The trained HGB model generates simultaneous predictions for all three target variables.
- These predicted outputs are then mapped back to their original human-readable labels (e.g., "High Priority") for display in the GUI.

Visualization and Reporting

- The framework translates abstract model data into intuitive graphical outputs, such as Word Clouds and frequency plots of common support issues.
- Comparative performance tables are generated to show the relative effectiveness of each model across different support metrics.
- These visualizations allow administrators to quickly grasp data trends and trust the model's underlying logic.
- High-fidelity charts are used to track how satisfaction levels vary across different ticket categories.

System Output & Adaptive Workflow

- The final output provides actionable intelligence, including predicted ticket priority, satisfaction ratings, and the likely resolution status.
- The system's Adaptive Workflow allows users to re-run any stage of the pipeline from preprocessing to training with new parameters or datasets.
- This modularity ensures the system can be continuously refined as customer language patterns and support priorities evolve over time.
- The result is a complete analytical lifecycle that turns raw support data into a streamlined, automated operational strategy.

4. RESULTS ANALYSIS

The results demonstrate the effectiveness of the proposed hybrid framework for customer support ticket analysis using XLNet embeddings and multiple machine learning models. Exploratory data analysis reveals meaningful patterns in ticket text, such as frequent keywords, common bigrams, and typical document lengths, highlighting the nature of customer queries. Comparative evaluation shows that traditional models like QDA and NC perform poorly, while LDA and SGD achieve moderate accuracy. The proposed Histogram-Based Gradient Boosting (HGB) classifier consistently outperforms all baseline models across all prediction tasks, including ticket priority, customer satisfaction, and resolution. Confusion matrices and ROC curves indicate strong classification performance with high precision and minimal misclassification.

Figure 3 represents the confusion matrix of the HGB classifier for predicting ticket priority levels across four classes: Critical, High, Low, and Medium. It shows that most predictions are concentrated along the diagonal, indicating a high level of correct classification. The model performs particularly well in identifying High and Low priority tickets with strong accuracy. Critical and Medium classes also show a significant number of correct predictions, though with some misclassification into neighboring categories. The distribution of values demonstrates that the model effectively distinguishes between different priority levels with minimal overlap.

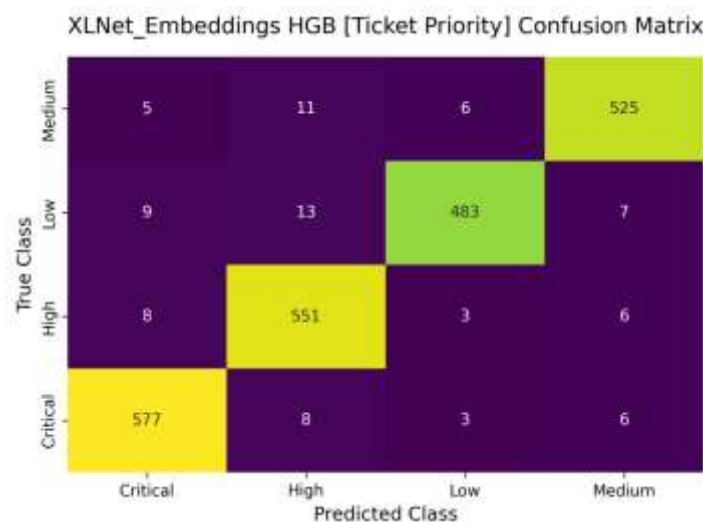


Figure 3: Confusion Matrix obtained using HGB Classifier for column “Ticket Priority”.



Figure 4: Confusion Matrix obtained using HGB Classifier for column “Customer Satisfaction”.

Figure 4 HGB Classifier (Proposed Model) The HGB classifier produces an exceptionally strong confusion matrix with dominant diagonal elements across all five rating levels (e.g., 420/429 for rating 1, 421/440 for rating 2, 442/462 for rating 3, 412/426 for rating 4, and 443/458 for rating 5).

Very few off-diagonal misclassifications occur, and most errors are limited to adjacent classes, demonstrating outstanding discriminative power and establishing HGB as the clearly superior model for predicting customer satisfaction ratings using XLNet embeddings.

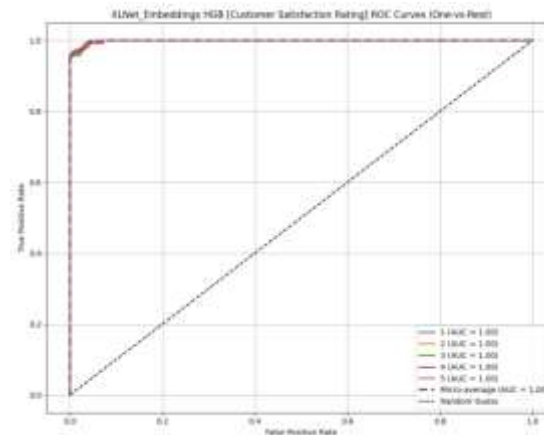


Figure 5: Confusion Matrix obtained using HGB Classifier for column “Customer Satisfaction”.

Figure 5 illustrates the ROC curves of the HGB classifier for predicting customer satisfaction levels using a one-vs-rest approach. It shows that all class curves are positioned very close to the top-left corner, indicating excellent true positive rates with minimal false positives. The Area Under the Curve (AUC) values for all classes are nearly equal to 1.00, reflecting outstanding classification performance. The micro-average ROC curve also demonstrates near-perfect predictive capability across all satisfaction categories. The model consistently distinguishes between different satisfaction levels with very high accuracy.

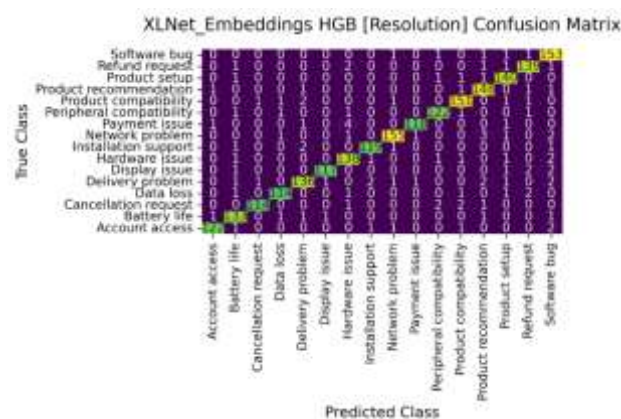


Figure 6: Confusion Matrix obtained using HGB Classifier for column Resolution.

Figure 6 HGB Classifier (Proposed Model) The HGB classifier produces an exceptionally strong and clean confusion matrix with dominant diagonal elements across nearly all resolution categories (e.g., high correct counts for Software bug, Product setup, Product recommendation compatibility, Hardware issue, Network problem support, Payment issue, Delivery problem, Cancellation request, etc.). Off-diagonal values are minimal and mostly limited to closely related categories, confirming the proposed HGB model's outstanding performance and clear superiority in accurately predicting resolution types from XLNet embeddings.

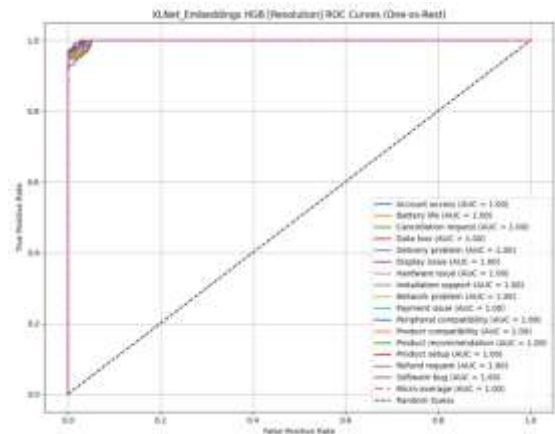


Figure 7: Confusion Matrix obtained using HGB Classifier for column “Resolution”.

Figure 7 presents the ROC curves of the HGB classifier for predicting resolution categories using a one-vs-rest strategy. It shows that the curves for all resolution classes are closely aligned near the top-left corner, indicating very high true positive rates and extremely low false positive rates. The Area Under the Curve (AUC) values for all categories are approximately 1.00, reflecting near-perfect classification performance. The micro-average ROC curve also confirms consistent and reliable predictions across all resolution types. The model effectively differentiates between multiple resolution classes with minimal overlap.

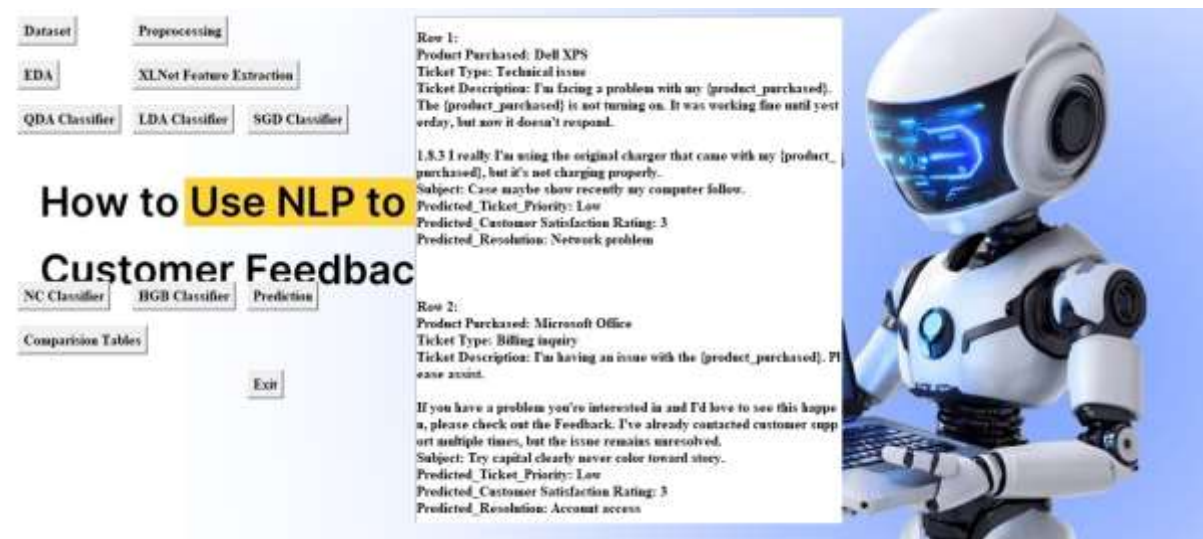


Figure 8: Real time predictions of customer satisfaction in support tickets.

Figure 8 illustrates the GUI-based implementation of the proposed hybrid machine learning system for support ticket analysis. The left panel shows various workflow buttons including Dataset Upload, Preprocessing, EDA, XLNet Feature Extraction, Classifier Training (QDA, LDA, SGD, NC, HGB), Prediction, and Comparison Tables. The right panel displays real-time ticket entries with predicted values for Ticket Priority, Customer Satisfaction Rating, and Resolution. This interface demonstrates how raw support ticket data is transformed through NLP preprocessing, feature extraction, and classification to produce actionable insights for customer service management.

Table 1: Overall Performance Comparison of Classification models for column “Customer Satisfaction”.

Mode	Accuracy	Precision	Recall	F-Score
------	----------	-----------	--------	---------

I	(%)	(%)	(%)	(%)
QDA	23.41	20.69	22.90	19.02
NC	23.41	20.64	22.90	19.02
SGD	49.75	54.93	49.36	48.97
LDA	56.01	56.03	55.98	56.00
HGB	96.26	96.28	96.27	96.27

Table 1 presents the overall performance comparison of various classification models for predicting Customer Satisfaction using XLNet embeddings. Among all models, HGB clearly outperforms the others, achieving the highest accuracy of 96.26% along with similarly high precision, recall, and F-score values, indicating robust and reliable predictions. Linear models like LDA and SGD Classifier show moderate performance, with accuracies around 50–56%, while QDA and Nearest Centroid perform poorly, achieving only about 23% accuracy. This demonstrates the effectiveness of ensemble tree-based methods over simpler classifiers for multi-class customer satisfaction prediction.

Table 2: Overall Performance Comparison of Classification models for column “Resolution”.

Mode I	Accuracy (%)	Precision (%)	Recall (%)	F-Score (%)
QDA	2.84	0.63	2.57	1.00
NC	7.83	11.44	7.71	5.10
SGD	50.47	69.22	49.83	44.78
LDA	63.17	63.50	63.32	63.16
HGB	95.41	95.60	95.40	95.48

Table 2 presents the performance comparison of classification models for predicting Resolution using XLNet embeddings. The HGB model again outperforms all others, achieving the highest accuracy of 95.41% along with strong precision, recall, and F-score, indicating highly reliable predictions. LDA and SGD Classifier show moderate performance, while simpler models like QDA and Nearest Centroid perform very poorly, with accuracies below 8%. This demonstrates the effectiveness of ensemble tree-based models for complex multi-class resolution prediction in support tickets.

5. Conclusion

The research successfully demonstrates an intelligent, automated solution for predicting multiple critical attributes of customer support tickets like priority, satisfaction rating, and resolution by combining the deep contextual language understanding capability of XLNet embeddings with the efficient and robust classical machine learning techniques, particularly the HGB classifier. This hybrid approach effectively leverages the strengths of both transformer-based feature extraction and

classical algorithms, resulting in high prediction accuracy and operational scalability. It significantly reduces manual intervention, accelerates ticket triage processes, and enhances overall customer service quality. The exploratory analyses and visualization components provide valuable insights into ticket data characteristics, further supporting continuous improvement in customer support workflows.

REFERENCES

- [1] Aamodt, A., & Plaza, E. (1994).” Case-based reasoning: Foundational issues, methodological variations, and system approaches.” *AI Communications*, 7(1), pp. 39-59.
- [2] Alessandro Zangari, Matteo Marcuzzo, Michele Schiavinato, Andrea Gasparetto, Andrea Albarelli, Ticket automation: An insight into current research with applications to multi-level classification scenarios, *Expert Systems with Applications*, Volume 225, 2023, 119984, ISSN 0957-4174.
- [3] Ali Zaidi, S.S., Fraz, M.M., Shahzad, M. and Khan, S., 2022. A multiapproach generalized framework for automated solution suggestion of support tickets. *International Journal of Intelligent Systems*, 37(6), pp.3654-3681.
- [4] Brown, T., Mann, B., Ryder, N., Subbiah, M., Kaplan, J., Dhariwal, P., Neelakantan, A., et al. (2020).” Language Models are Few-Shot Learners.” *arXiv preprint arXiv: 2005.14165*.
- [5] Cannon, D., & Wheeldon, D. (2007). *ITIL Service Operation*. The Stationery Office.
- [6] Chinthapatla, Y. (May, 2024) “AI Integration with ServiceNow and CMDB: Revolutionizing Industries and Society”. *Journal of Science & Technology*, vol. 5, no. 3, pp. 1-13.
- [7] Devlin, J., Chang, M. W., Lee, K., & Toutanova, K. (2019).” BERT: Pre-training of Deep Bidirectional Transformers for Language Understanding.” *Proceedings of the 2019 Conference of the North American Chapter of the Association for Computational Linguistics: Human Language Technologies (NAACL-HLT)*, pp. 4171-4186.
- [8] Explainable AI Framework for Policy-Compliant Anomaly Detection in Data Pipelines. (2025). *International Journal of Communication Networks and Information Security*, 16(4). <https://doi.org/10.48047/ijcnis.16.4.2111>
- [9] Gupta, H.S. and Sengupta, B., 2012. Scheduling service tickets in shared delivery. In *Service-Oriented Computing: 10th International Conference, ICSOC 2012, Shanghai, China, November 12-15, 2012. Proceedings 10* (pp. 79-95). Springer Berlin Heidelberg.
- [10] Gaddam, S. *Integrating Analytics into the Development Process: Bridging the Gap between Data Insights and Design Execution*.
- [11] Jain, A. K., Murty, M. N., & Flynn, P. J. (1999).” Data Clustering: A Review.” *ACM Computing Surveys*, 31(3), pp. 264-323.
- [12] Sai Maneesh Kumar Prodduturi. (2025). EFFICIENT DEBUGGING METHODS AND TOOLS FOR IOS APPLICATIONS USING XCODE. *International Journal of Data Science and IoT Management System*, 4(4), 1–6. <https://doi.org/10.64751/ijdim.2025.v4.n4.pp1-6>
- [13] Murtagh, F., & Contreras, P. (2012).” Algorithms for hierarchical clustering: an overview.” *Wiley Interdisciplinary Reviews: Data Mining and Knowledge Discovery*, 2(1), pp. 86-97.

- [14] Paramesh, S.P. and Shreedhara, K.S., 2019. Automated IT service desk systems using machine learning techniques. In *Data Analytics and Learning: Proceedings of DAL 2018* (pp. 331-346). Springer Singapore.
- [15] Radford, A., Wu, J., Child, R., Luan, D., Amodei, D., & Sutskever, I. (2019). "Language Models are Unsupervised Multitask Learners." *OpenAI Blog*, 1(8).
- [16] Reddy, S. K. R. (2024). Designing Blockchain Architecture to Transform Loyalty Rewards into Cryptocurrency Investments.
- [17] Poojari, R. (2025). A Comparative Analysis of Fine-Tuning Versus Retrieval-Augmented Approaches for Enhancing Healthcare-Centric Large Language Models.
- [18] Kalae, U. K. (2021). Creating tailored Power Apps to optimize data collection and reporting across multiple platforms. *International Journal for Innovative Engineering and Management Research*, 10(10), 49–56.
- [19] Kalae, U. K. (2020). Developing scalable Power BI dashboards for enhanced data analysis and strategic business decision-making. *International Journal of Enhanced Research in Science, Technology & Engineering*, 9(3), 8–15.
- [20] Vasagam, M., Kumar, A., & Garg, A. (2026). Learning Execution Plan Embeddings for Multi-Dimensional Query Resource Prediction. *IEEE Access*.
- [21] Patel, S., & Patyrykin, K. (2025). Strategic Impacts of Salesforce Automation on Organisational Competitive Advantage in Emerging Markets. *Journal of Posthumanism*, 5(12), 357–372. <https://doi.org/10.63332/joph.v5i12.3782>
- [22] Saai Reddy Purmani, S. (2023). The Transformation of IT Leadership in Business Organizations: Shifting from Technical Supervision to Strategic Empowerment. *JOURNAL OF ADVANCE AND FUTURE RESEARCH*, 1(5). <https://doi.org/10.56975/jaafr.v1i5.503885>