Ticket Classification Tool using Streamlit and GenAI

B Vinodhini[1], M Lavanya[2], Abhinav S[3], Akhilesh P[4], Jeeva K[5] and Kalaiarasan S[6]
{ profbvinodhini1975@gmail.com[1], lavanyamarimuthu@gmail.com[2], abhinavsureshabhi003@gmail.com[3], ap0845546@gmail.com[4], jeevakrishnasamy12@gmail.com[5], kalaiarasan6923@gmail.com[6] }

Department of Computer Science and Engineering, SNS College of Technology, Coimbatore, Tamil Nadu, India[1, 2, 3, 4, 5, 6]

Abstract. This project presents an automated IT support and customer service system that categorizes incoming requests based on priority. It uses Streamlit, a Python framework, to create an interactive app for ticket classification. The system leverages machine learning algorithms to process and categorize ticket content into technical support issues, billing inquiries, or general questions. By automating ticket categorization, it speeds up response times, reduces manual effort, and optimizes resource allocation. Key features include easy ticket upload, real-time predictions, and visual analytics. Streamlit's scalability makes the solution adaptable to various organizational needs. This project demonstrates how machine learning and modern web technology can enhance ticket management efficiency.

Keywords: AI, ML, NLP, TF-IDF, XAI, IT, UI, SVM.

1 Introduction

In today's digital age, organizations face a growing number of support tickets from various channels like emails, chats, and web forms. Efficient handling of these tickets is crucial for customer satisfaction and timely resolution. However, manually classifying tickets is time-consuming, error-prone, and struggles to keep up with increasing volumes. Automated ticket classification, powered by machine learning, addresses these challenges by quickly sorting tickets based on their content. This project utilizes Streamlit, an open-source Python library, to build an easy-to-use app for automatic ticket classification, enabling real-time predictions and insights. The system applies natural language processing (NLP) to categorize tickets into predefined types, such as technical problems, account inquiries, and payment issues. It uses state-of-the-art models, including transformers and traditional machine learning techniques, to ensure accuracy and adaptability across various datasets. The user-friendly interface allows administrators to upload tickets, visualize predictions, and track performance metrics. The solution is scalable and integrates with existing ticket management systems, improving ticket routing, reducing resolution time, and increasing customer satisfaction.

2 Literature Survey

Automating ticket categorization has been applied widely in industries including banking, IT services and customer support. Bonechi et al. [1], who focused on banking, evidencing that AI classification systems allow faster responses and a more efficient customer support. Similarly, Zangari et al. [2] offered an in-depth overview on multi-level classification ticket automation and discussed property of scalable, challenge for productionalize strong automation pipeline.

ICITSM-Part II 2025, April 28-29, Tiruchengode, India

DOI 10.4108/eai.28-4-2025.2357989

In terms of CRM systems, Az et al. [3] investigated support ticket for text classification models focusing on how machine learning can be adapted to pragmatic business needs. Oliveira et al. [4] compared a number of different algorithms for IT incident ticket classification and highlighted the appropriateness of some models to high-dimensional textual data. Ishizuka et al. [5] was used to visualize categorized tickets in software development projects, thus enabling the teams to better understand developed features.

The closest previous work is that of Miao et al. [6] designed the reliable ticket routing for expert networks and Sun et al. [7], where content-aware resolution sequence mining is proposed to enhance routing efficiency. Building up on this work, method courtesy of Gupta et al. [8] offered scheduling policies for service-based ticketing delivery. Xu et al. [9] further advanced the field by using signature-based technique in classification that increased accuracy on massive IT systems.

Recent works have applied more sophisticated methods like graph convolutional networks. Schad et al. [10], which forecast help desk ticket reassignments to minimize inefficiencies of the ticket routing. Bruni et al. [11] optimized hyperparameters for ticket classification in a black-box manner, showing improved performance. Complementarily, Gandla et al. [12] used ML to classify IT tickets in a conference environment, again demonstrating the relevance of traditional supervised learning techniques in academic and business contexts.

Other similar works generalise to a variety of classification problems. Borg et al. [13] researched email classification employing word embeddings for enhanced customer support responses. Tuveri et al. [14] presented Beep4Me, ticket validation for transport system as an example of real-world applications other than IT service for ticets automatization. Truss and Böhm [15] investigated the use of AutoML for classification of customer support tickets, with reduced need for manual feature engineering. Jain [16] proposed the purported first AI-clustered ticket management system with a focus on enhanced support workflows.

The introduction of generative AI has propelled this field even further. Akaike [17] showed enough generative AI-based categorization activities in enterprise case studies that this approach may apply to complex support data for analysis. “CategorizeAI: An open-source platform for multi-model NLP ticket classification with streamlit” by the Streamlit Community19which demonstrated an end-to-end solution for multi-model NLP ticket classification, has emphasized that Streamlit is perfect on rapid prototyping and deployment. Finally, Liu et al. [19] used large language models to automate consumer ticket escalation in support systems, showing the potential of LLMs for real-time ticket prioritization and routing.

3 Methodology

The project starts with a Load Data Center module, where administrators can upload departmental documents in PDF format. These documents, e.g., HR policies or IT troubleshooting procedures, are converted to text and divided into manageable bits using a text-splitting method. Each bit is embedded in high-dimensional vectors with the all-MiniLM-L6-v2 model. These embeddings are subsequently pushed into Pinecone, a vector database, where they're stored under certain indexes. This constitutes the knowledge base that the system draws upon at the time of user interaction. It facilitates quick and smart document retrieval at the time of user query by semantic similarity.

Users input their support tickets using an intuitive Streamlit web interface. This interface receives user queries or grievances in plain text. The system processes the text immediately after submission through different modules such as classification, sentiment recognition, and response generation. Streamlit session state is utilized to handle ticket information between departments dynamically.

Every ticket is stored and showcased under its corresponding category according to real-time prediction. This helps in maintaining an interactive as well as efficient means for both users as well as admins to engage with the system. The process of department classification starts by converting the user's ticket into a numerical vector with the help of Sentence Transformer embeddings. These vectors are then fed into a pre-trained Support Vector Machine (SVM) model that makes predictions for the appropriate department: HR, IT, or Transport. This model is trained and stored using the Model Training module where administrators import a CSV file with labeled ticket data. The SVM model is trained to classify based on semantic patterns in the embedded ticket text. This guarantees uniform and effective classification of user problems. It helps in classifying the tickets based on its categories which help to improve the ticket resolving time and improvises the effectiveness of the ticket classification.

To aid users in real-time, the system also incorporates GPT-3.5 Turbo via the LangChain framework. When a ticket is entered, the text is embedded and utilized to fetch appropriate context from Pinecone. These retrieved documents and the original ticket are passed on to GPT-3.5 Turbo through a QA chain. The model produces a helpful and context-sensitive reply based on the retrieved knowledge. This enables the system to generate dynamic, smart responses automatically without any human interaction. Users can get informative responses specific to their question and its department.

Apart from the classification, the system also carries out the sentiment analysis with the help of a GPT-3.5 Turbo. The ticket material is passed into the LLM by a personalized prompt that tells the LLM to classify if the sentiment is negative or positive. In contrast to standard sentiment models, it does this based on the LLM's linguistic intelligence to translate emotional tone into emotional meaning. What comes back is a higher degree of subtle, accurate sentiment tagging. The response is logged together with ticket details. It makes the support team understand better the mindset and expectations of the user.

4 Architecture of Aglecare

Automatic Ticket Classification Tool using AI is an AI-driven ticket categorization platform that aims to optimize IT support processes with automated, real-time ticket categorization. At the center of the system is a web application built in Streamlit that offers a simple user interface for ticket upload and analysis. Ticket data in textual form can be entered directly into the application by the users, with the application being rendered with Streamlit's own components and added CSS to support a clean and responsive user experience on any device. The Python-based backend manages data ingestion, preprocessing, and model inference.

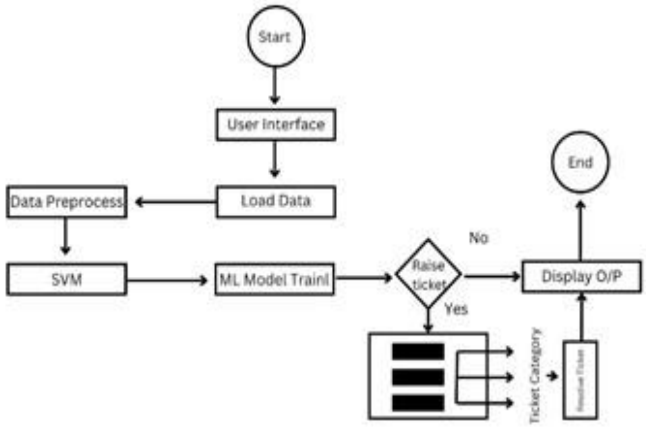

Fig. 1. System Architecture of the AI-Based Ticket Classification Tool.

Tickets are processed through natural language preprocessing operations like tokenization, lemmatization, and stop-word filtering. As feature extraction, TF-IDF and Sentence Transformer (MiniLM-L6-v2) embeddings are used. The features are then fed into a Support Vector Machine (SVM) classifier, which outputs the predicted ticket category with corresponding confidence scores. The platform also leverages Pinecone for semantic search, enabling retrieval of similar historical tickets to assist support agents in making decisions. Deployment is done through Streamlit Cloud or Render, with CI/CD and version control controlled through GitHub. This modular AI-boosted architecture guarantees Automatic ticket classification tool to provide accurate, scalable, and efficient ticket classification, enabling IT teams to enable faster and more Intelligent customer support. It helps to reduce the risk in time management in IT sectors. Fig. 1 shows the System Architecture of the AI-Based Ticket Classification Tool

5 Result and Discussion

The project Ticket classification tool using GenAI showed very accurate output when given training on a labeled support query dataset. Through the conversion of ticket text to dense vector representations through the all-MiniLM-L6-v2 model, the SVM classifier effectively classified HR, IT, and Transport-related issues. The model always yielded very high accuracy when tested on a test set, confirming the robustness of the classification reasoning. Moreover, as embeddings acquired semantic value rather than single words, worded problems put indirectly were even correctly labeled. This ensured that user tickets are directed to appropriate departments with little misclassification. In terms of operations, this reduces the load of human support teams and lets users acquire correct help from competent domain experts. The use of a vectorized representation also lays the foundation for the addition of more advanced classifiers in the future. Streamlit's session state successfully stored and displayed these classified tickets in the dynamic department tabs.

One of the project's greatest contributions was the integration of the GPT-3.5 Turbo via Lang Chain for the real-time question answering. After the ticket was classified, the system searched

Pinecone for embedded documents from uploaded department PDFs that were applicable. These documents were passed to GPT-3.5 Turbo as context to generate a human-like answer to the user's query. The responses were not just accurate but also contextually rich and tailored. This made the system intelligent, interactive, and capable of supporting users without human intervention. In the majority of cases, the responses provided by the model were extremely close to what a domain specialist would answer, which is a testimony to the potency of blending LLM with a well-curated knowledge base. The capacity of GPT-3.5 Turbo to learn to accommodate various ticket wordings, or unclear questions added to its worth. It also eliminated the need for pre-formatted intent templates or rigid rules.

The GPT-3.5 Turbo sentiment analysis module added a valuable extra level of insight by determining the emotional tone of each ticket through specially crafted prompts where the LLM was asked to mark sentiment as positive or negative. The model performed well in distinguishing between the gratitude and frustration, even where the words used were minimal or indirect. With the introduction of a second prompt for priority tagging, the system would be able to tag the tickets as High or Normal priority. This logic was helpful in helping teams to understand not just what the user was needing, but also how urgently they needed it. Perhaps the greatest advantage of using GPT-3.5 Turbo in both tasks was the elimination of separate training pipelines for priority and sentiment models. The final item, the Pending Tickets dashboard, grouped together all the enlightened system outputs into a tidy, usable form. All the tickets were tabulated against their classified department and accompanied with sentiment and priority flags. Admins were automatically able to view the critical tickets and respond appropriately. The use of Streamlit tabs ensured ease of navigation, and live updating facilitated reactive support workflows. The project successfully demonstrates the ability to integrate multiple AI technology into a working platform — embeddings, LLMs, vector databases, and ML into a cohesive whole. Notably, all of the modules interacted well with each other without bottlenecks or incompatibilities.

6 Conclusion

Streamlit’s intelligent ticket categorization capabilities offer a robust solution capable of maximizing the customer service organization through the use of natural language processing and machine learning. This new system breaks down new tickets into different areas - billing questions, account troubles, technical problems, etc. It’s easy-to-learned Streamlit interface makes it simple for teams to onboard and manage ticket data with a few clicks. With real-time classification, the platform accelerates response times and productivity, automatically prioritizing and directing tickets to the appropriate support representatives. Engaging dashboards & visualizations Streamlit's rich dashboards and visualizations facilitate data-driven decision-making by providing visibility into ticket trends, workloads, and classification accuracies. What's more, the trust metrics provided by the machine learning model help support agents judge the credibility of a classification or learn how to cope with ambiguous tickets. This automated process in turn enables agents to concentrate basically on kind of more challenging tickets. Fig. 2 shows the Data Preprocessing Interface in the Streamlit-Based Ticket Classification Tool

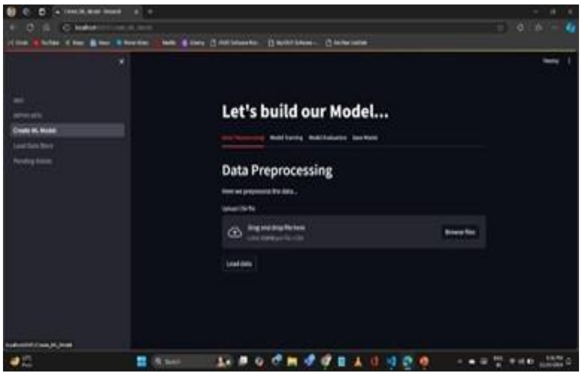

Fig. 2. Data Preprocessing Interface in the Streamlit-Based Ticket Classification Tool.

References

[1] Bonechi, S., Palma, G., Caronna, M., Ugolini, M., Massaro, A., & Rizzo, A. (2024). Enhancing customer support in banking: Leveraging AI for efficient ticket classification. *Procedia Computer Science*, *246*, 128–137. https://doi.org/10.1016/j.procs.2024.09.235

[2] Zangari, A., Marcuzzo, M., Schiavinato, M., Gasparetto, A., & Albarelli, A. (2023). Ticket automation: An insight into current research with applications to multi-level classification scenarios. *Expert Systems with Applications*, *225*, 119984. https://doi.org/10.1016/j.eswa.2023.119984

[3] Az, Ş., Takazoğlu, U., Karakaş, A. (2021). Text Classification Models for CRM Support Tickets. In: Hemanth, J., Yigit, T., Patrut, B., Angelopoulou, A. (eds) Trends in Data Engineering Methods for Intelligent Systems. ICAIAME 2020. Lecture Notes on Data Engineering and Communications Technologies, vol 76. Springer, Cham. https://doi.org/10.1007/978-3-030-79357-9_8

[4] Oliveira, D. F., Nogueira, A. S., & Brito, M. A. (2022). Performance Comparison of Machine Learning Algorithms in Classifying Information Technologies Incident Tickets. *AI*, *3*(3), 601-622. https://doi.org/10.3390/ai3030035

[5] Ishizuka, R., Washizaki, H., Tsuda, N., Fukazawa, Y., Ouji, S., Saito, S., & Iimura, Y. (2022). Categorization and Visualization of Issue Tickets to Support Understanding of Implemented Features in Software Development Projects. Applied Sciences, 12(7), 3222. https://doi.org/10.3390/app12073222

[6] Miao, G., Moser, L.E., Yan, X., Tao, S., Chen, Y., Anerousis, N. (2012). Reliable Ticket Routing in Expert Networks. In: Dai, H., Liu, J., Smirnov, E. (eds) Reliable Knowledge Discovery. Springer, Boston, MA. https://doi.org/10.1007/978-1-4614-1903-7_7

[7] Sun, P., Tao, S., Yan, X., Anerousis, N., Chen, Y. (2010). Content-Aware Resolution Sequence Mining for Ticket Routing. In: Hull, R., Mendling, J., Tai, S. (eds) Business Process Management. BPM 2010. Lecture Notes in Computer Science, vol 6336. Springer, Berlin, Heidelberg. https://doi.org/10.1007/978-3-642-15618-2_18

[8] Gupta, H.S., Sengupta, B. (2012). Scheduling Service Tickets in Shared Delivery. In: Liu, C., Ludwig, H., Toumani, F., Yu, Q. (eds) Service-Oriented Computing. ICSOC 2012. Lecture Notes in Computer Science, vol 7636. Springer, Berlin, Heidelberg. https://doi.org/10.1007/978-3-642-34321-6_6

[9] Xu, J., Zhang, H., Zhou, W., He, R., & Li, T. (2018). Signature based trouble ticket classification. Future Generation Computer Systems, 78(Part 1), 41–58. https://doi.org/10.1016/j.future.2017.07.054

[10] Schad, J., Sambasivan, R., & Woodward, C. (2022). Predicting help desk ticket reassignments

with graph convolutional networks. Machine Learning with Applications, 7, 100237. https://doi.org/10.1016/j.mlwa.2021.100237
[11] Bruni, R., Bianchi, G., & Papa, P. (2023). Hyperparameter Black-Box Optimization to Improve the Automatic Classification of Support Tickets. Algorithms, 16(1), 46. https://doi.org/10.3390/a16010046
[12] Gandla, P.K.K., Verma, R.K., Panigrahi, C.R., Pati, B. (2024). Ticket Classification Using Machine Learning. In: Pati, B., Panigrahi, C.R., Mohapatra, P., Li, KC. (eds) Proceedings of the 7th International Conference on Advance Computing and Intelligent Engineering. ICACIE 2022. Lecture Notes in Networks and Systems, vol 1. Springer, Singapore. https://doi.org/10.1007/978-981-99-5015-7_41
[13] Borg, A., Boldt, M., Rosander, O., & Svensson, P. (2021). E-mail classification with machine learning and word embeddings for improved customer support. Neural Computing & Applications, 33, 1881–1902. https://doi.org/10.1007/s00521-020-05058-4
[14] Tuveri, G., Garau, M., Sottile, E., Pintor, L., Atzori, L., & Meloni, I. (2022). Beep4Me: Automatic Ticket Validation to Support Fare Clearing and Service Planning. *Sensors*, *22*(4), 1543. https://doi.org/10.3390/s22041543
[15] Truss, M., & Böhm, S. (2024). AI-based classification of customer support tickets: State of the art and implementation with AutoML. *arXiv preprint arXiv:2406.01789*. https://arxiv.org/abs/2406.01789
[16] Jain, A. (2024). AI enhanced ticket management system for optimized clustering. *ACM Digital Library*. https://doi.org/10.1145/3703412.3703433
[17] Upasna Doshi (2023). Generative AI powered support ticket categorization. *Akaike Case Studies*. https://www.akaike.ai/case- studies/generative- ai- powered- support- ticket- categorization
[18] Streamlit Inc. (2023). CategorizeAI: Multi-model NLP text classification platform. Streamlit Discussion Forum. https://discuss.streamlit.io/t/categorizeai-multi-model-nlp-text-classification-platform/93815
[19] Liu, Y., Zhang, R., Chen, H., & Wang, T. (2025). TickIt: Leveraging large language models for automated ticket escalation. *arXiv preprint arXiv:2504.08475*. https://arxiv.org/abs/2504.08475