

The Application of NLP for Automated IT Support Ticket Classification and Resolution

Student ID: 253001724

Student Name: Anas Abdulkader Alzain

Problem Statement

Problem Domain

IT Service Management (ITSM) and automated ticket processing.

Problem Description

Companies process thousands of IT support tickets each day. 30–40% of the time is spent on classifying, prioritizing, and routing tickets instead of solving the issues themselves (Tai et al., 2024). This manual process is slow and error-prone, often resulting in misrouted tickets, delayed responses, and higher operational costs (Liu et al., 2023; Inspirisys, 2023).

Ticket management is difficult because tickets are written in unstructured natural language, vary in detail and clarity, and often include multiple issues. Correct classification requires domain knowledge, which is challenging to automate (Peru Study, 2024). As IT systems evolve, new categories and services emerge, requiring ongoing updates to manual classification rules (Liu et al., 2023).

Significance of the Problem

Manual ticket classification often leads to poor outcomes, significantly affecting SLA compliance, customer satisfaction, and overall productivity. Misclassification can delay resolution by up to 40% (Ajit et al., 2023), causing user frustration and wasted resources. Financial losses are estimated at \$10,000–\$50,000 per support agent per year due to inefficiencies in manual ticket handling (Peru Study, 2024).

Misrouted tickets may bounce between multiple teams before reaching the correct handler, which further increases delays (Tai et al., 2024). These recurring issues reduce trust in IT support processes and negatively impact organizational efficiency.

Existing Solutions

- Transformer-based Models:
Ticket-BERT (Liu et al., 2023) outperforms Microsoft Cognitive Services and processes more than 7,000 tickets per day using active learning.
Fine-tuned BERT models (Inspirisys, 2023) can achieve up to 89% accuracy with less than 200 ms inference time per ticket.

- **Ensemble Machine Learning:**
Random Forest and XGBoost perform well in predicting resolution times and priorities. Random Forest achieved 94.32% accuracy, and XGBoost obtained a 6.44-minute MAE (Ajit et al., 2023).
 - **Hybrid NLP-LSTM Systems:**
Combining NLP with LSTM improves temporal prediction accuracy, delivering a 65% increase in response efficiency and predicting 78% of recurring incidents in advance.
 - **Lightweight Models for SMEs:**
The Peru Study (2024) developed a small-scale classifier achieving 85% accuracy using 5,000–10,000 samples and operating on standard hardware without GPUs.
-

Research Gaps and Justification

Key challenges remain:

- **High computational cost:**
BERT-based models require significant GPU computing resources and take a long time to train, making them difficult for SMEs to adopt.
- **Poor generalization:**
Significant retraining is required to adapt models to different organizations. Models trained on one company's data do not generalize well to another without retraining (Liu et al., 2023).
- **Ambiguous and multi-issue tickets:**
Between 20–30% of classification failures arise from industry-specific jargon, vague descriptions, and multi-issue tickets (Tai et al., 2024; Inspirisys, 2023).
- **Lack of interpretability:**
IT teams need to understand why a ticket was classified a certain way, especially for high-impact issues. Transformer-based models generally lack transparency and are considered “black boxes.”
- **Accuracy vs. efficiency trade-offs:**
Lightweight models run quickly but sacrifice accuracy, while enterprise-level models such as Ticket-BERT offer higher accuracy at the cost of heavy resource usage.
- **Limited practical guidance:**
Few benchmarking studies compare commonly available models under uniform conditions. Clear recommendations for limited-resource environments remain unavailable.

Because of these gaps, a comparative study of pre-trained models and traditional machine learning approaches using public datasets and standard hardware is necessary to guide real-world adoption.

Research Objectives

Primary Goal

To evaluate and compare AI models for automated IT ticket classification using publicly available datasets, covering both open-source pre-trained models and commercial API solutions to address different resource needs.

Specific Objectives

1. Identify suitable public IT ticket datasets and preprocess them through text cleaning, label encoding, and dataset splitting.
2. Implement classification using one or both of the following:
 - Open-source models: DistilBERT, TinyBERT, RoBERTa-base (fine-tuned or zero-shot)
 - Commercial APIs: GPT-5, Claude, Gemini using structured prompts
3. Evaluate classification performance using accuracy, precision, recall, and F1-score, and record:
 - training and inference times
 - memory usage
 - API cost per request, latency, and reliability
4. Analyze trade-offs between:
 - accuracy vs. speed (open-source models)
 - accuracy vs. cost (API models)
 - setup complexity and maintenance requirements
 - suitability for enterprise, SME, or academic environments
5. Build a simple application that:
 - accepts ticket descriptions

- allows model/API selection
- displays prediction results with confidence scores
- presents performance metrics, cost, and processing time

References

- Ajit, S., Srivastava, P., Sharma, R., & Kumar, A. (2023). Machine learning framework for end-to-end implementation of Incident duration prediction. *arXiv preprint arXiv:2304.11507*. <https://arxiv.org/abs/2304.11507>
- Inspirisys (2023). Automated Customer Ticket Classification using BERT. *Inspirisys Blog*. <https://www.inspirisys.com/blog-details/Customer-Ticket-Classification-using-BERT/136>
- Liu, Z., Wingate, D., & Shou, Z. (2023). Ticket-BERT: Labeling Incident Management Tickets with Language Models. *arXiv preprint arXiv:2307.00108*. <https://arxiv.org/abs/2307.00108>
- Peru Study Authors (2024). A Machine Learning-Based Predictive Model for the management of Incidents in Small and Medium-Sized Enterprises in Peru. *Proceedings of the 2024 3rd Asia Conference on Algorithms, Computing and Machine Learning (CACML '24)*. <https://dl.acm.org/doi/10.1145/3654823.3654913>
- Recent Authors (2025). Enhancing IT incident management with natural language processing and predictive analytics. *International Journal of Science and Research Archive*, 15(3), 224-237. https://www.researchgate.net/publication/392465831_Enhancing_It_incident_management_with_natural_language_processing_and_predictive_analytics
- Tai, T., Haw, S., Ng, K., Al-Tarawneh, M., & Tong, G. (2024). Performance Evaluation on Resolution Time Prediction Using Machine Learning Techniques. *JOIV: International Journal on Informatics Visualization*, 8(2), 583-591. <https://joiv.org/index.php/joiv/article/view/2305>