Simulation-Driven Predictive AI for Printer Repair Services: A Proactive Ticket Resolution Approach

Samia Zouaoui [1] *

[1] Computer Science Department, LINFI Laboratory, Biskra University, University of Batna 2, Batna, Algeria
samia.zouaoui@univ-batna2.dz

Abstract. This study presents a predictive, artificial intelligence–driven ticket resolution framework aimed at improving customer service efficiency within the printer maintenance domain, a critical technical support area for both organizational and individual users. Owing to the limited availability of reliable, publicly accessible real-world maintenance data, the proposed system is developed and evaluated using a simulation-based approach. A synthetic dataset comprising 150 maintenance cases and 10 relevant attributes was generated to realistically model common printer service scenarios. Twelve machine learning classification algorithms were implemented and systematically evaluated using two distinct random states (20 and 42) to ensure robustness and reproducibility. Model performance was assessed using accuracy and F-score metrics. The experimental results demonstrate that the Light Gradient Boosting Machine (LightGBM) classifier outperformed the other models, achieving an accuracy of 0.70 and an F-score of 0.7033. These findings confirm the feasibility and effectiveness of simulation-driven predictive modeling for proactive ticket resolution in contexts where real-world data are unavailable or incomplete. The study underscores the potential of artificial intelligence to transition customer support systems from traditional reactive mechanisms toward proactive, intelligent service management solutions in operational maintenance environments

Keywords: Artificial Intelligence, Customer Service, Classification, Prediction, Machine Learning, Printer Maintenance.

1 Introduction

In recent years, artificial intelligence (AI) has become a key enabler in enhancing customer support and service management systems [1] , particularly in technical service environments where timely issue resolution is critical. As organizations increasingly rely on automated and intelligent solutions, AI-driven approaches offer significant potential to improve efficiency, accuracy, and customer satisfaction by transforming traditional reactive support models into proactive service systems [2].

D. Agti et al. (eds.), *Proceedings of the International Conference on Artificial Intelligence Applications in Business Administration in MENA Region (ICAIABA 2026)*, Advances in Economics, Business and Management Research 393,
https://doi.org/10.2991/978-94-6239-711-8_17

The domain of printer maintenance represents a highly important technical service area for both organizations and individual customers [3]. Printers remain essential tools in daily business operations and personal workflows, as many administrative, legal, and operational tasks depend on the timely production of physical documents. When a printer malfunctions, work processes can be severely disrupted. In addition, the lack of accurate information regarding the repair status or expected resolution time often leads to uncertainty, delays, and reduced productivity [4]. This issue is particularly problematic in traditional support environments where customers are left without clear updates or reliable expectations about service completion.

Conventional ticketing systems commonly rely on manual ticket classification, rule-based workflows, and reactive problem handling. These approaches suffer from several limitations, including slow resolution times, high dependency on human intervention, and poor scalability when handling large volumes of service requests [3]. As a result, support teams struggle to prioritize tickets effectively and provide transparent and timely communication to users.

Predictive modeling in the domain of printer maintenance cans greatly improve operational efficiency by enabling proactive ticket resolution, optimizing resource allocation, and enhancing customer satisfaction [2]. Many users struggle to find skilled and reliable repair technicians who respect time commitments, which is critical for minimizing workflow disruptions. When they fail to locate a competent service provider, they often discard their printers, even for relatively simple issues, leading to unnecessary expenses and economic losses, particularly for organizations and administrative institutions that rely heavily on printing equipment. Accurate prediction of maintenance requirements and appropriate resolution actions allows organizations to reduce downtime, minimize costs, and streamline service processes [7]. Developing machine learning models for such tasks requires access to structured datasets that capture the essential characteristics of printer maintenance operations, including common issues, printer types, and resolution outcomes.

To address these challenges, predictive intelligence and proactive service automation have emerged as promising advances in customer support systems [8]. Predicting ticket resolution time and recommending appropriate resolution actions can significantly improve decision-making, improve resource allocation, and increase customer trust by providing clear expectations and faster responses [9]. Such predictive capabilities represent a natural next step in the evolution of intelligent helpdesk systems.

However, one of the main challenges in this research area is the lack of publicly available, real-world datasets related to printer maintenance and technical support tickets [4]. To overcome this limitation, this work proposes the creation of a simulated data set that reflects realistic printer maintenance scenarios and support workflows. For this data set, twelve machine learning classification algorithms (XGBoost, LightGBM, MultinomialNB, Neural Network (MLP), Logistic Regression, Support Vector Machine, K-Nearest Neighbors, Decision Tree, Ran-dom Forest, Extra Trees, Gradient Boosting, and Gaussian Naive Bayes) were applied and tested with two different random states (20 and 42) to guaranty both robustness and reproducibility. Model performance was assessed through accuracy and F-score, allowing a comprehensive comparison of their effectiveness. Through this approach, the study aims to demonstrate the feasibility and effectiveness of predictive AI techniques to improve service efficiency and customer satisfaction within the printer maintenance domain.

2 Related Work

Recent research has demonstrated significant progress in the application of artificial intelligence and machine learning to automate and improve the help desk and support ticket systems [5]. Automatic ticket classification is a common focus, as it directly affects routing accuracy and resolution speed [6]. Truss and Boehm investigate the use of automated machine learning (AutoML) to classify customer support tickets, showing that AutoML-generated models can achieve strong performance without requiring specialized expertise, thereby reducing classification errors and supporting faster resolution workflows [1]. This aligns with efforts in improving technical support tasks where accurate ticket labeling is crucial for effective response.

Deep learning has also been applied to ticket assignment tasks. The TaDaa model uses transformer-based representations to assign incoming tickets both to appropriate groups and to specific resolvers, demonstrating high accuracy in routing support issues, which has implications for reducing average resolution time in complex service domains [10]. Other machine learning approaches for ticket routing have explored ensemble and hybrid techniques, showing that ad-vanced classifiers can outperform traditional models in handling unstructured ticket data [11].

Beyond classification and assignment, the structure and complexity of helpdesk workflows have been studied using graph-based models. Schad et al. employ graph convolutional networks to model ticket resolution paths and predict reassignment needs, capturing interactions between tickets and analysts to support more informed handling of difficult cases [12]. Such techniques provide deeper in-sights into ticket dynamics that are valuable for technical support environments where issues may involve multiple steps or escalations.

Classic machine learning pipelines remain relevant in helpdesk research. Al-Hawari and Barham propose a ticket classification system that uses text preprocessing and feature extraction to associate tickets with the correct service area. Their findings show substantial improvements in resolution efficiency when tick-ets are automatically categorized and routed from the outset [13]. Additionally, studies have evaluated machine learning models specifically for predicting resolution times, highlighting the potential of predictive analytics to provide realistic time estimates and improve planning and customer satisfaction [14].

Finally, systematic overviews of ticket automation research emphasize the growing importance of machine learning and natural language processing in streamlining support workflows. These works show that contextualized language models and hierarchical classification strategies can significantly enhance ticket processing accuracy, which is fundamental for efficient technical service management [15].

3 Methodology

3.1 Proposed Architecture

The proposed architecture provides a structured and systematic framework for implementing the predictive printer maintenance system (see figure 1). It is de-signed to integrate data collection, preprocessing, model training, and prediction modules into a cohesive pipeline, ensuring efficient and accurate maintenance outcome predictions.

At the core of the architecture is the simulated dataset, which captures relevant features of maintenance operations, including printer specifications, is-sue categories, ticket priorities, and resolution outcomes. The data flows into a preprocessing layer, where it is cleaned, normalized, and prepared for machine learning. Multiple classification algorithms are then trained and evaluated (XGBoost, LightGBM, MultinomialNB, Neural Network (MLP), Logistic Regression, Support Vector Machine, K-Nearest Neighbors, Decision Tree, Random Forest, Extra Trees, Gradient Boosting, and Gaussian Naive Bayes) , allowing selection of the most effective model for predicting ticket resolution success and recommending appropriate maintenance actions. The architecture also includes a user interface and reporting module, enabling visualization of predictions and insights, and facilitating proactive decision-making for maintenance teams.

3.2 Construction of a Simulated Dataset

Motivation and Importance:
The simulated dataset was designed to replicate realistic service scenarios and capture the inherent variability of operational conditions, providing a solid foundation for training and evaluating machine learning algorithms. By systematically including the most relevant features of maintenance operations, the dataset supports reproducible experimentation and thorough assessment of classification models aimed at predicting ticket outcomes and resolution strategies.

Lack of Existing Official Datasets:

Despite the clear advantages of predictive analytics in printer maintenance, no publicly available or officially curated datasets exist for this domain, creating a significant challenge for model development and validation. To address this limitation, a synthetic dataset was generated to simulate realistic maintenance scenarios. This dataset provides a structured and comprehensive foundation for evaluating multiple machine learning classification algorithms, capturing sufficient variability and coverage to train and assess predictive models. It serves as a practical proxy for real-world datasets and supports reproducible research in proactive printer maintenance analytics.

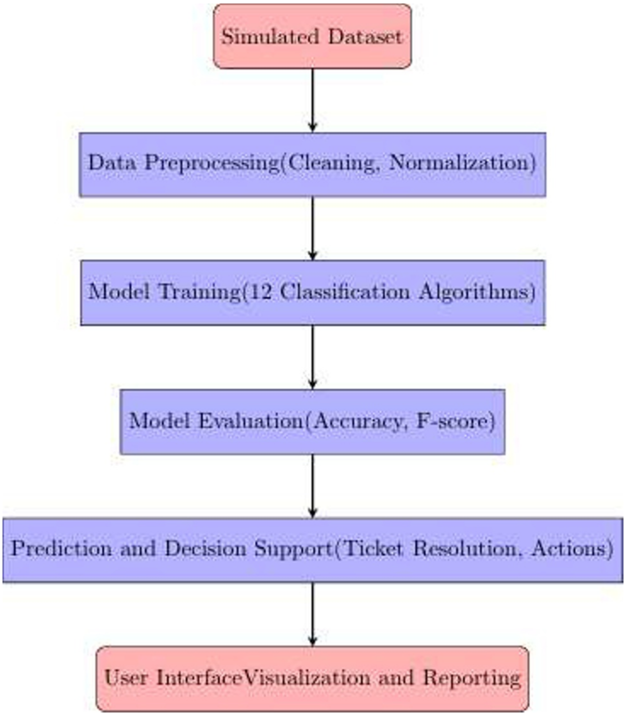

Fig. 1. Proposed Architecture for Predictive Printer Maintenance System

Description of the Simulated Dataset:

To develop predictive models for printer maintenance, a synthetic dataset was generated due to the absence of publicly available or official datasets in this domain. The dataset comprises 150 maintenance cases described by 10 key attributes (as shown in the table 1), including ticket identifiers, printer characteristics, issue categories, priority levels, technician expertise, resolution time, customer type, resolution actions, and operation success (see fig. 2).

3.3 Model Training

We considered twelve standard classifiers: XGBoost, LightGBM, MultinomialNB, Neural Network (MLP), Logistic Regression, Support Vector Machine, K-Nearest Neighbors, Decision Tree, Random Forest, Extra Trees, Gradient Boosting, and Gaussian Naive Bayes. The dataset was split once into a stratified 80/20 split (training/testing), with the test set kept fixed. On the training set, experiments were conducted using two different random states (20 and 42) to ensure robustness and reproducibility of the results

	ticket_id	print.cket_id	printer_brand	printer_type	issue_category	issue_description	priority	technician_level	resolution_time_minutes	Customer_type	resolution_action	operation_success
0	TCKT-0001	:KT-0001	Ricoh	Thermal	Paper Jam	Printer produces blank pages	Low	Intermediate	139	Individual	Escalated to senior technician	Yes
1	TCKT-0002	:KT-0002	Canon	Laser	Hardware Failure	Scanner not working	Low	Intermediate	240	Individual	Replaced toner cartridge	Yes
2	TCKT-0003	:KT-0003	Xerox	Multifunction	Paper Jam	Paper stuck inside the printer	Critical	Expert	15	Administration	Reset printer settings	Yes
3	TCKT-0004	:KT-0004	Xerox	Thermal	Print Quality	Wi-Fi connection drops frequently	Low	Senior	54	Individual	Cleared paper jam	Yes
4	TCKT-0005	:KT-0005	Xerox	Multifunction	Connectivity	Printer produces blank pages	High	Intermediate	131	Individual	Configured network settings	Yes
...	...	...	...	...	...	...	...	...	...	...	...	...
145	TCKT-0147	:KT-0147	Canon	Inkjet	Ink Level	Incorrect ink level reading	Low	Junior	20	Individual	Reset ink level monitor	Yes
146	TCKT-0148	:KT-0148	Epson	Dot Matrix	Print Quality	Uneven character spacing	Medium	Junior	35	Warehouse	Adjusted platen alignment	No
147	TCKT-0149	:KT-0149	HP	Laser	Security Issue	Unauthorized access error	High	Senior	80	Administration	Updated security settings	Yes
148	TCKT-0150	:KT-0150	Brother	Laser	Network Issue	Printer drops connection	Medium	Senior	55	Small Business	Updated network firmware	No
149	TCKT-0151	:KT-0151	Canon	Laser	Paper Jam	Repeated jams in tray 2	High	Senior	45	Administration	Replaced tray rollers	No

150 rows × 11 columns × 11 columns

Fig. 2. Illustrative Screenshot of the Simulated Dataset.

Attribute	**Description**
ticket_id	Unique identifier for each maintenance ticket.
printer_brand	Brand of the printer (e.g., HP, Canon, Epson). Type
printer_type is-	of printer (e.g., laser, inkjet).
sue category	General category of the reported issue (e.g., hardware, software, paper jam).
issue_description	Detailed description of the issue reported by the customer.
priority techni-	Priority level of the ticket (e.g., low, medium, high). Skill
cian level _	or expertise level required for the resolution (e.g., junior, senior, expert).
resolution_time_minutes	Time required to resolve the ticket, measured in minutes.
Customer_type	Type of customer (e.g., individual, business, institutional).
resolution_action	Action taken to resolve the issue (e.g., repair, replacement, remote guidance).
operation_success	Binary indicator of successful resolution (1 = yes, 0 = no).

Table 1. Attributes of the Simulated Printer Maintenance Dataset

4 Results

The overall performance comparison of the twelve individual classifiers is illustrated in Figure 3, which reports the full test-set metrics for all models. For random state 20, LightGBM achieved the highest test accuracy of 0.70, followed closely by Gradient Boosting and Random Forest, both with accuracies of 0.67. For random state 42, LightGBM, MultinomialNB, Neural Network (MLP), and Gaussian Naive Bayes all achieved the same accuracy of 0.60, sharing the same rank. These results indicate that ensemble methods like LightGBM and tree-based models consistently perform well, while for the second random state, several models exhibit similar predictive performance on the simulated dataset. For the random state 20, LightGBM achieved the highest test F-score of 0.703, followed by Gradient Boosting with 0.680 (see the figure 4). For random state 42, Random Forest obtained the highest test F-score of 0.660, with Extra Trees ranking second. These results highlight the robust predictive performance of ensemble methods, particularly LightGBM and tree-based models, in capturing the patterns of the simulated printer maintenance dataset.

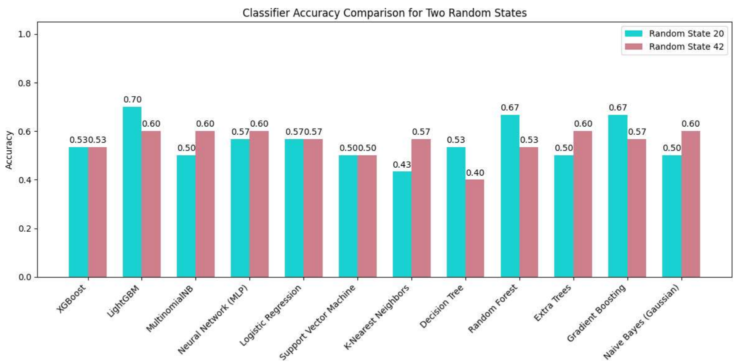

Fig. 3. Comparison of Test Accuracy Across Twelve Classifiers

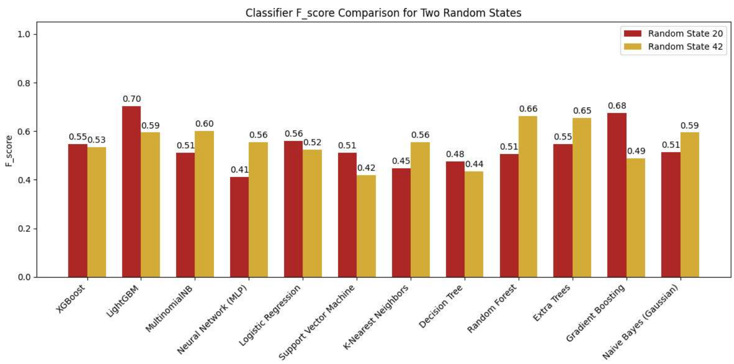

Fig. 4. Comparison of Test F-score Across Twelve Classifiers

5 Conclusion & Future Work

This study demonstrates the feasibility of using simulated datasets to evaluate predictive machine learning models in the domain of printer maintenance. By generating a synthetic dataset with 150 cases and 10 relevant attributes, we were able to systematically test twelve standard classifiers under two different

random states. LightGBM consistently achieved the highest F-score, hig lighting the effectiveness of ensemble tree-based methods to predict ticket outcomes and recommending resolution actions. The results confirm that simulation-driven predictive modeling can provide valuable insight and support proactive maintenance decisions in the absence of real-world data.

The next step is to develop a real-world dataset by collecting operational maintenance data from printers across different institutions. This will enable

the validation of the models in practical scenarios and improve the generalizability of the predictive system. Additionally, future work will explore the integration of dynamic data collection mechanisms and continuously updating the models to enhance prediction accuracy and support real-time decision-making in printer maintenance operations.

References

1. Truss, M., Boehm, S.: AI-based Classification of Customer Support Tickets: State of the Art and Implementation with AutoML. arXiv preprint arXiv:2406.01789 (2024). doi:10.48550/arXiv.2406.01789.
2. Wenninger, A., Rau, D., Ro¨glinger, M.: Improving Customer Satisfaction in Proactive Service Design. Electronic Markets 32, 1399–1418 (2022). doi:10.1007/s12525-022-00565-9
3. FUJIFILM Business Innovation: Prevent Costly Printer Breakdowns: The Impor-tance of Printer Maintenance (2023). https://www.fujifilm.com/fbca/en/solutions/ insights/article/prevent-costly-printer-breakdowns
4. Joviˇci´c, E., Primorac, D., Cupi´c, M., Jovic, A.: Publicly Available Datasets for Predictive Maintenance in the Energy Sector: A Review. IEEE Access 10, 1–20 (2023). doi:10.1109/ACCESS.2023.3295113
5. Truss, M., Boehm, S.: AI-based Classification of Customer Support Tickets: State of the Art and Implementation with AutoML. arXiv preprint arXiv:2406.01789 (2024). doi:10.48550/arXiv.2406.01789
6. Molino, P., Zheng, H., Wang, Y.-C.: COTA: Improving the Speed and Accu-racy of Customer Support through Ranking and Deep Networks. arXiv preprint arXiv:1807.01337 (2018). doi:10.48550/arXiv.1807.01337
7. Aziz, A., Sowmyashree, P.: AI-Powered Predictive Maintenance System for In-dustrial Equipment: Reducing Downtime and Costs Through Sensor-Based Machine Learning. Journal of Scientific Research and Technology 3(8), 95–105 (2025). https://www.jsrtjournal.com/index.php/JSRT/article/view/387
8. Adeniyi, D., Ikeh, C., Duraisamy, A., Peterson, C.: Predictive Service and Proactive Support: How AI Transforms Reactive Companies into Experi-ence Leaders. Journal of Service Innovation and Technology (2025).

https://www.researchgate.net/publication/397745711_Predictive_Service_and_Proactive_Support_How_AI_Transforms_Reactive_Companies_into_Experience_Leaders
9. Arqawi, S.M., Hijazi, S.: Predicting Incident Resolution Time in Digital Trans-formation Systems and Its Impact on Decision-Making Using Machine Learn-ing. Lex Localis - Journal of Local Self-Government 23(S6), 1047–1055 (2025). doi:10.52152/801900
10. Feng, L., Senapati, J., Liu, B.: TaDaa: Real-Time Ticket Assignment Deep Learn-ing Auto Advisor for Customer Support Systems. arXiv preprint arXiv:2207.11187 (2022). doi:10.48550/arXiv.2207.11187.
11. Seanbudh, S., Siriborvornratanakul, T.: Using Machine Learning to Improve Ticket Classification for IT Service Management System. Science and Engineering Connect 44(4), 667–680 (2021)
12. Schad, J., Sambasivan, N., Woodward, C.: Predicting Help Desk Ticket Reassignments with Graph Convolutional Networks. Machine Learning with Applications 7, 100204 (2022)
13. Al-Hawari, F., Barham, H.: A Machine Learning Based Help Desk System for IT Service Management. Journal of King Saud University - Computer and Information Sciences 33(6), 702–714 (2021)
14. Tai, T.E., Ahmad, R., Mustapha, A.: Performance Evaluation of Resolution Time Prediction Techniques Using Machine Learning. International Journal on Informatics Visualization 8(2), 583–591 (2024)
15. Wijaya, A.S., Oktavia, T.: Machine Learning Approaches for Helpdesk Ticketing System: A Systematic Literature Review. Journal of Theoretical and Applied Information Technology 102(5), 1321–1335 (2024)