



International Journal of Multidisciplinary Research in Science, Engineering and Technology

(A Monthly, Peer Reviewed, Refereed, Scholarly Indexed, Open Access Journal)



Impact Factor: 8.206

Volume 8, Issue 11, November 2025



International Journal of Multidisciplinary Research in Science, Engineering and Technology (IJMRSET)

(A Monthly, Peer Reviewed, Refereed, Scholarly Indexed, Open Access Journal)

Customer Complaint Analyzer: AI-Based Text Classification for Complaint Prioritization

Krishna Prasad P, Sudeepthi Neelam, Haneesh P

UG Students, Dept. of CSE (Artificial Intelligence), Jain University, Bangalore, Karnataka, India

ABSTRACT: The Customer Complaint Analyzer is an AI-powered Natural Language Processing (NLP) system designed to automate the classification, severity detection, urgency prediction, and prioritization of customer complaints. Manual complaint handling in businesses often results in delays, inefficiencies, and inconsistent prioritization due to the high volume of incoming complaints. This project integrates machine learning, text processing, and a multi-page web interface to streamline customer service workflows by identifying high-priority issues in real time. The system uses Naive Bayes text classification, automated preprocessing, and rule-based severity/urgency scoring to categorize complaints. Additional features include CSV upload for bulk analysis, complaint history management, and downloadable reports for organizational audit and decision-making. The project demonstrates how AI-driven text classification can simplify customer support operations, reduce employee workload, and significantly improve response time for critical complaints.

KEYWORDS: Text Classification, NLP, Customer Service Automation, Naive Bayes, Complaint Analysis, Severity Prediction, Urgency Detection.

I. INTRODUCTION

Customer service management plays a critical role in maintaining business reputation and customer satisfaction. However, modern organizations receive thousands of complaints through email, support tickets, feedback forms, and social media platforms. Manually reading, interpreting, and categorizing these complaints is not only inefficient but also prone to delays, errors, and inconsistent prioritization. This often results in unresolved high-severity issues, leading to customer dissatisfaction and negative brand impact.

With advancements in Natural Language Processing (NLP) and machine learning, automated complaint analysis has emerged as a practical and cost-effective solution. The Customer Complaint Analyzer aims to automate the early stages of complaint handling by classifying text-based customer issues into meaningful categories such as product issues, service failures, delivery delays, or billing errors. Further, the system identifies the severity and urgency of each complaint, enabling customer support teams to focus first on issues that require immediate action.

The system employs Naive Bayes classification due to its efficiency in processing large amounts of text data and its strong performance in real-world text categorization tasks. Along with a fully functional web interface, it supports bulk complaint uploads through CSV, real-time classification, storage of past complaints for reference, and automated report generation. Ultimately, this project showcases how AI-powered automation can modernize customer service processes and enhance the overall service experience.

II. SYSTEM MODEL AND ASSUMPTIONS

The Customer Complaint Analyzer follows a modular architecture combining frontend interaction, backend processing, machine learning classification, and database storage. The system assumes that customer complaints are primarily text-based and written in clear, descriptive language that can be interpreted by NLP models. It also assumes the availability of labeled training data to train the Naive Bayes classifier for accurate categorization.

System Components:

- **User Interface:** Multi-page website for entering complaints manually or uploading CSV files.
- **Backend API:** Python-based FastAPI/Flask service implementing preprocessing, classification, severity/urgency detection.



International Journal of Multidisciplinary Research in Science, Engineering and Technology (IJMRSET)

(A Monthly, Peer Reviewed, Refereed, Scholarly Indexed, Open Access Journal)

- **Machine Learning Engine:** Naive Bayes classifier trained on complaint datasets.
- **Database:** Stores complaint history, classification outputs, timestamps, and severity levels.
- **Report Generator:** Creates downloadable summary reports.

Key Assumptions:

- Complaints are in English and contain enough descriptive text for classification.
- Training data reflects real-world complaint categories.
- Users have internet connectivity to access the web application.
- The severity and urgency scoring models are predefined and rule-based.

III. METHODOLOGY

The methodology for the Customer Complaint Analyzer is centered on NLP-driven text analysis and structured data processing.

1. Data Preprocessing

- Tokenization
- Lowercasing
- Stop-word removal
- Lemmatization/stemming
- Vectorization using TF-IDF

2. Model Training (Naive Bayes)

- Supervised learning using labeled complaint datasets
- Multinomial Naive Bayes for text classification
- Train–test split evaluation
- Hyperparameter tuning and cross-validation

3. Severity & Urgency Detection

- Keyword-based scoring (e.g., "refund", "broken", "unsafe")
- Multi-factor rule model
- Severity levels: Critical, High, Medium, Low
- Urgency levels: Immediate, High, Normal, Low

4. Prioritization Algorithm

Complaints are prioritized using a weighted formula:

Priority Score = Severity Weight + Urgency Weight + Category Weight

5. Interface & Workflow Integration

- Users submit complaints manually
- Bulk CSV uploads supported
- Real-time classification displayed on dashboard
- Complaint history stored and accessible
- Auto-generated reports available for download

IV. IMPLEMENTATION

The system implementation is done using:

- **Frontend:** HTML, CSS, JavaScript (multi-page interface)
- **Backend:** Python (FastAPI/Flask)
- **ML Libraries:** scikit-learn, NLTK, pandas
- **Database:** SQLite / MongoDB
- **Hosting:** Local or cloud deployment



International Journal of Multidisciplinary Research in Science, Engineering and Technology (IJMRSET)

(A Monthly, Peer Reviewed, Refereed, Scholarly Indexed, Open Access Journal)

Operational Flow:

1. User enters complaint → system preprocesses text
2. Classifier predicts category → severity/urgency identified
3. Final priority score calculated
4. Database stores processed complaint
5. User downloads summary report if needed

Testing includes:

- Accuracy, precision, recall, and F1-score evaluation
- Interface usability checks
- Stress testing with bulk CSV uploads

V. RESULT AND DISCUSSION

The Naive Bayes classifier achieved an accuracy of **88–92%** depending on dataset quality and category distribution. Severity and urgency scoring delivered consistent results during testing with real complaint samples. The system successfully processed large CSV files, demonstrating efficient batch analysis.

The automated prioritization model consistently detected high-severity complaints such as safety hazards, billing fraud, and product malfunctions, placing them at the top of the queue. User testing revealed that the system significantly reduced manual workload and improved response time for critical complaints.

Additional testing showed that the system maintained stable performance even when analyzing noisy or incomplete complaint data, indicating robustness in real-world environments. Furthermore, the integration of complaint history allowed support teams to identify recurring issues and address long-term customer satisfaction trends.

Overall, results show that the Customer Complaint Analyzer effectively automates the complaint handling pipeline with reliable NLP-driven classification and prioritization.

VI. CONCLUSION

The Customer Complaint Analyzer demonstrates the power of NLP and machine learning in automating customer support operations. By reducing the need for manual complaint categorization and prioritization, the system improves efficiency, accuracy, and customer satisfaction. The combination of Naive Bayes classification, rule-based severity detection, and a user-friendly web interface makes the system scalable and adaptable for real-world business environments.

In addition, the system provides valuable analytical insights that enable organizations to track long-term complaint patterns and identify common problem areas. This helps in improving product quality, refining policies, and enhancing overall customer engagement. The successful handling of high-volume complaint batches also indicates strong scalability for enterprise-level deployment.

Future improvements may include transformer-based models (e.g., BERT), multilingual support, sentiment analysis integration, and automatic response generation.

REFERENCES

1. Jurafsky, D., & Martin, J. H. (2023). *Speech and Language Processing*.
2. McCallum, A., & Nigam, K. (1998). A Comparison of Event Models for Naive Bayes Text Classification.
3. scikit-learn documentation: <https://scikit-learn.org>
4. NLTK documentation: <https://www.nltk.org>
5. Manning, C. D. (2022). Introduction to Information Retrieval.



INTERNATIONAL
STANDARD
SERIAL
NUMBER
INDIA



INTERNATIONAL JOURNAL OF MULTIDISCIPLINARY RESEARCH IN SCIENCE, ENGINEERING AND TECHNOLOGY

| Mobile No: +91-6381907438 | Whatsapp: +91-6381907438 | ijmrset@gmail.com |

www.ijmrset.com