

# Bank Customer Complaint Classification Using Machine Learning and Deep Learning Models

Sumedha Arya

Aryasumedha06@gmail.com

**Abstract**— Customer complaints are prominent in evaluating service quality and customer satisfaction for banking sector. With the rapid growth of digitalization, banks receive a large volume of unstructured complaints, making them difficult to perform manual analysis due to inefficiency and time-consumption. This study proposes an automated framework for analyzing and classifying bank customer complaints using machine learning (ML) and deep learning (DL) techniques. A large-scale dataset consisting of 162,421 customer complaints was collected from Kaggle. After data preprocessing, several ML models, including Logistic Regression, Decision Tree, Multinomial Naïve Bayes, and Random Forest, were trained using TF-IDF features, along with a DL model as Long Short-Term Memory (LSTM). Experimental results shows that the LSTM model outperforms traditional ML models, achieving a highest classification accuracy of 96.61%. The findings indicate that credit reporting-related complaints are the most dominant, highlighting critical areas for service improvement. The proposed system provides an effective and scalable solution for automated complaint classification and can support banks in service improvement.

**Index Terms**—Bank Customer Complaints, Complaint Classification, Machine Learning, Deep Learning, Natural Language Processing.

## I. INTRODUCTION

In the modern banking ecosystem, customer satisfaction is important for long-term success. With digitalization, customers increasingly express their concerns and gratitude through complaints. These complaints provide valuable insights into service failures, operational inefficiencies, and customer expectations. However, the increase in complaints has made manual analysis impractical, highlighting the need for automated and intelligent complaint management systems.

Previous studies have emphasized the importance of mechanism for effective customer complaints resolution in the banking sector. With advancements in machine learning (ML), Deep Learning (DL) and natural language processing (NLP), automated complaint classification has gained significant attention. Traditional ML approaches have been widely used for text classification tasks. More recently, deep learning models, and transformer-based architectures, have demonstrated superior performance by capturing contextual and sequential information in textual data. Several studies have also explored sentiment analysis and topic modeling to better understand customer emotions and complaint severity.

Despite these advancements, there remains a need for a comprehensive framework that can effectively recognize and classify bank customer complaints. This study addresses an end-to-end complaint classification system that categorizes complaints by banking product type based on narrative content. By using both ML and DL techniques, the research aims to identify the most effective approach for large-scale complaint classification.

The main contributions of this study are:

1. systematic preprocessing and balancing of a large real-world banking complaint dataset,
2. comparative evaluation of multiple ML and DL models for complaint classification, and
3. deployment of the best-performing model using a Streamlit-based web application for real-time usage.

The paper is further divided into following sections; review of the literature, research methodology, results analysis, conclusion and future work.

## II. LITERATURE REVIEW

In this section, we highlight the various techniques used for addressing customer complaint in banking sector. The entire review focus on latest work done by authors to enhance the redressal system.

The authors [1] in their study examined the efficiency of the Banking Ombudsman Scheme (BOS), introduced by the Reserve Bank of India in 1995. In a trend analysis of 10-year data, the study highlighted that BOS has proved to be more effective for urban and metropolitan customers. Specially, retail clients, with foreign banks have shown excellence in customer satisfaction. However, in limitations, customers from rural and semi-urban areas, have registered unsatisfaction due to limited awareness and access to complaint filing mechanisms. The review highlights the need for a robust grievance redressal system.

In another study [2], authors explored the importance of efficient complaints management system in retail banking within a Customer Relationship Management (CRM) framework. The effective complaint handling management system is important for maintaining customer satisfaction and loyalty. The unresolved complaints can lead to unhealthy relationships or damage to the reputation. By analyzing real complaints from four retail banks, the study develops a database to classify complaints by topic and importance. The proposed recommendations based on the results highlights improved staff training, faster complaint resolution, systematic complaint classification, and robust follow-up mechanisms.

The authors [3] explored customer complaining attitudes in the Indian banking sector. They focussed on the impact of service failures on customer satisfaction. Authors in their study found that there is a need for effective complaint handling strategies. According to them, the bank customers are divided into four groups through cluster analysis. These are non-complainers, switchers, prompt complainers, and positive thinkers. This is based on their complaining behaviours and their perceptions of the complaint handling system. The research highlights the shift towards relationship-focused marketing, service recovery goal in retaining customers, and the classification of service failures. The proposed recommendation highlights a need for banks to adopt proactive complaint management system.

The authors [4] investigate consumer complaints in the banking sector using data mining techniques. Focusing on datasets related to credit reporting, mortgages, debt collection, consumer loans, and banking accounts, the study employs hybrid clustering and predictive modeling to identify patterns in complaints across various banks, products, and regions. The analysis reveals that the most common issues are their geographic distribution, specific products or services involved, enabling banks to improve in these areas. By addressing these issues, banks can improve customer satisfaction, which can lead to revenue growth. In future, the study can be extended to customers complaint data analysis for additional products and services.

The authors [5] examined the nature and extent of customer complaints in Indian banks, focusing on public and private sectors along with and foreign banks during 2006–2008. The study finds that public sector banks receive the highest number of complaints, related to deposits, credit cards, and housing loans. This overall negatively impacts the customer satisfaction and bank performance. In comparison to them, private sector and foreign banks better resolve complaints on time. The study focusses on urgent need for public sector banks to establish dedicated customer care centers to resolve customer complaints on time. It also recommends that the Reserve Bank of India should make it compulsory for quick customer complaints resolution for the survival of public sector banks in a competitive market.

In their study, authors [6] explored customer complaints in the banking sector using natural language processing (NLP) and data mining techniques. They focussed on sentiment analysis for user comments. The research analyses consumer finance complaints related to credit reporting, mortgages, debt collection, consumer loans, and banking accounts. Machine learning techniques such as t-SNE and Latent Dirichlet Allocation (LDA) were used to identify patterns in complaint data across 50 banks. The study highlights the importance of digital transformation and competition in fintech and tech giants based on customers feedback to enhance service quality. By using sentiment analysis, the proposed model helps banks to gain insights into customer opinions, prioritize service issues, and improve strategic decision-making, for enhancing customer satisfaction.

Authors [7] in their work, conducted a literature review on sentiment-aware complaint identification. They used FINCOP financial dataset, which includes annotated complaints with sentiments, emotions, and severity labels. The study explored NLP techniques, such as text classification and sentiment analysis, to develop models for customer sentiment, emotions and complaints analysis and classification. Using embeddings such as DistilBERT, GloVe, FastText, and RoBERTa with classifiers including logistic regression, Bi-LSTM, LSTM, and Random Forest, the research achieved high accuracies. These are 95% for complaint classification and 86% for severity classification using DistilBERT and GloVe, respectively. Data augmentation has been done to improve the performance of model by giving diverse dataset for training. The review highlights the strength of advanced embedding models with their generalization across various tasks and datasets, suggesting future exploration of LLMs to improve customers complaint identification systems.

The authors [8] reviewed applications of machine learning models for the automatic railway complaints classification focussing on sentiment analysis and customer feedback on Twitter. The study proposed algorithms such as Random Forest, Support Vector Machine, Logistic Regression, and Naïve Bayes to classify tweets into positive, neutral, or negative sentiments with urgent complaints prioritization. In research methodology, steps such as data collection, its preprocessing, feature extraction, and training multiple models for tweets sentiment classification, using a vectorization pipeline is performed. The review highlights the role of machine learning based sentiment analysis in managing large-scale customer feedback.

A development of an automated Customer Complaint Classification System is reviewed by the authors [9] in their study using NLP and ML techniques to handle customer complaints in businesses. The study addresses the limitations of manual complaint processing by proposing a ML based classifier model. Linear SVM and Logistic Regression models demonstrated high accuracy and computational efficiency, making them suitable for real-time deployment. The review focusses on the system's ability to reduce manual workload, improve response times, and customer satisfaction. They also identified the limitations in handling evolving complaint patterns. Future directions include the integration of deep learning ensemble models, with real-time deployment to improve classification accuracy for new complaints.

The development of a machine learning-based system to automate the classification of customer complaints for PT PLN, Indonesia's largest electricity company, is reviewed by the authors [10] to improve the customer-centric services. The study addresses the limitations of manual complaint classification by Contact Center operators. They utilized a dataset of complaints from January to December 2023 and compared various machine learning models. The SVM achieved a highest accuracy at 93%, outperforming Random Forest 92%, Logistic Regression 91%, and Naïve Bayes 90%, due to its ability to find optimal hyperplanes for customer complaints classification. The review highlights system's ability to improve operational efficiency by automating customer complaints using machine learning models. In future, authors suggested to explore deep learning models with multi-label classification to handle complex complaints, particularly from social media platforms.

### III. RESEARCH METHODOLOGY

In this section, we detailed methodology used in this research to analyze bank customer complaints. Various ML and DL models are used to classify them by product type based on their narrative content. This research methodology comprises of various steps such as data collection, preprocessing, exploratory data analysis (EDA), text processing, visualization, and ML and DL classification for customer complaints in banking.

1. **Data Collection:** The dataset is sourced from Kaggle. It contains 162,421 records with three columns: an unnamed index, product (credit\_card, retail\_banking, credit\_reporting, mortgages\_and\_loans, debt\_collection), and narrative.

#### 2. Data Preprocessing

To ensure data quality and suitability for analysis, the following preprocessing steps are applied:

- **Removing Unnecessary Columns:** The unnamed index column (Unnamed: 0) is dropped.
- **Handling Duplicates:** Duplicate records are identified, reveals 37,742 duplicates and removed resulting in 124,676 unique records.
- **Handling Missing Values:** Missing values are checked, identifying three missing entries in the narrative column, which are removed.
- **Text Preprocessing:** The narrative column is cleaned using a custom function (clean\_text). Following steps were used:
  - Convert text to lowercase.
  - Remove URLs, punctuation, digits, and extra whitespace using regular expressions.
  - Remove stop words.
  - Apply lemmatization to normalize words.
  - The cleaned text is stored in the narrative column.
- **Data Balancing:** To address class imbalance, upsampling is performed. Each product category is resampled to match the size of the largest class (56,303 records), resulting in a balanced dataset of 281,515 records.

#### 3. Exploratory Data Analysis (EDA)

The EDA phase focuses on understanding the distribution of complaints and key terms in narratives.

- **Frequency Analysis:** The distribution of complaints across product categories is analyzed using bar and pie plots.

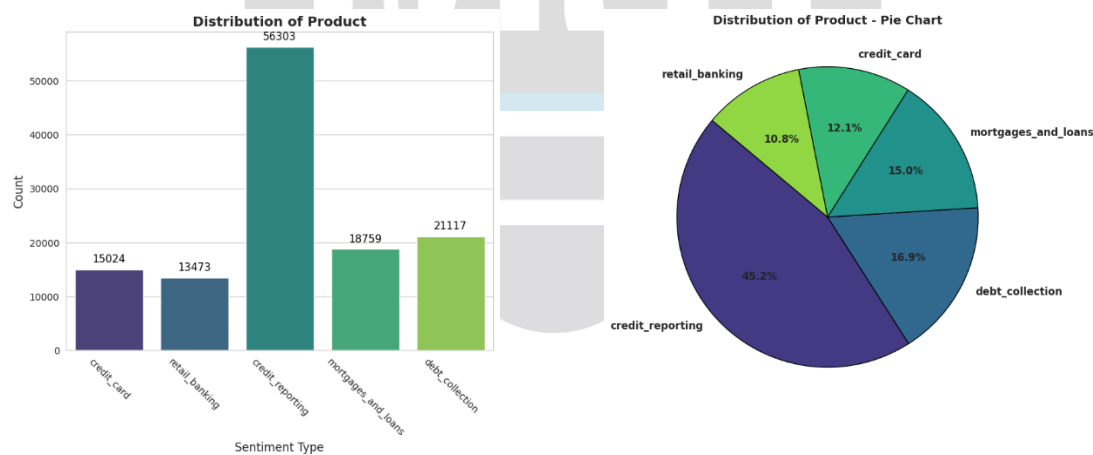


Fig. 1: Count Plot and Pie Plot for Product Distribution



#### IV. COMPARATIVE RESULTS ANALYSIS

In this section, we demonstrated the results of machine learning and deep learning models in classifying bank customer complaints based on narrative text. After preprocessing and balancing, the dataset comprised 281,515 records across five product categories, ensuring that model performance was not biased toward majority classes.

Among traditional machine learning models, Logistic Regression performed the best, achieving an accuracy of 87.27%, indicating that linear models with TF-IDF features are highly effective for structured text classification tasks. Multinomial Naïve Bayes also showed strong performance with 83.05% accuracy, benefiting from its probabilistic nature and suitability for text data. In contrast, Decision Tree and Random Forest models exhibited comparatively lower accuracies, 74.45% and 69.85%, respectively, suggesting that tree-based models struggle to capture high-dimensional sparse text representations effectively.

The LSTM-based deep learning model significantly outperformed all machine learning models, achieving a validation accuracy of 96.61% with a low validation loss of 0.1755. This superior performance can be attributed to LSTM’s ability to capture sequential dependencies and contextual information in complaint narratives. The confusion matrices further confirm that the LSTM model produces fewer misclassifications across all product categories, demonstrating strong generalization capability.

Training and validation curves indicate stable convergence, with no significant overfitting observed during the 10 training epochs. This confirms the robustness of the proposed LSTM architecture for large-scale complaint classification tasks.

The Visualization of Model Performance are given as:

- **Confusion Matrices:** Heatmaps are created to visualize correct and incorrect predictions for each model.

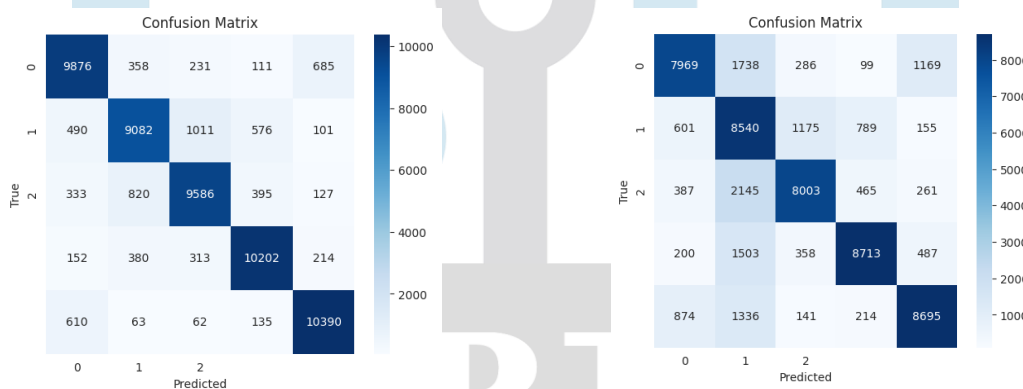


Fig. 4: Confusion Matrix of Logistic Regression and Decision Tree

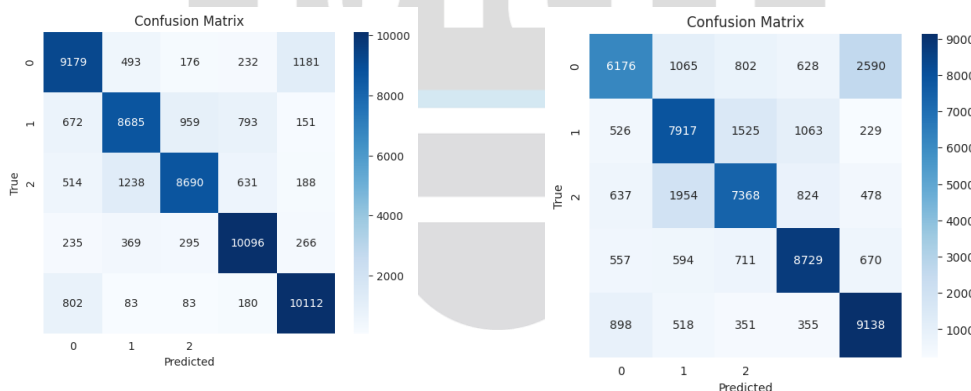


Fig. 5: Confusion Matrix of Multinomial and Random Forest

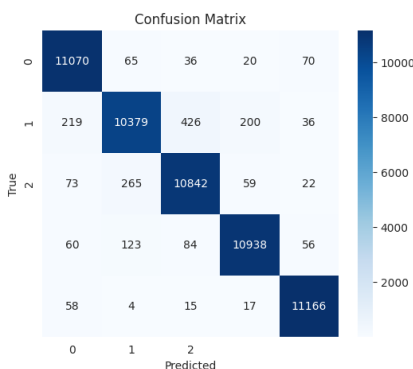


Fig. 6: Confusion Matrix of LSTM

- **LSTM Training Plots:** Training and validation accuracy/loss are plotted to assess the LSTM model's learning behavior over 10 epochs, showing steady improvement in accuracy and convergence of loss.

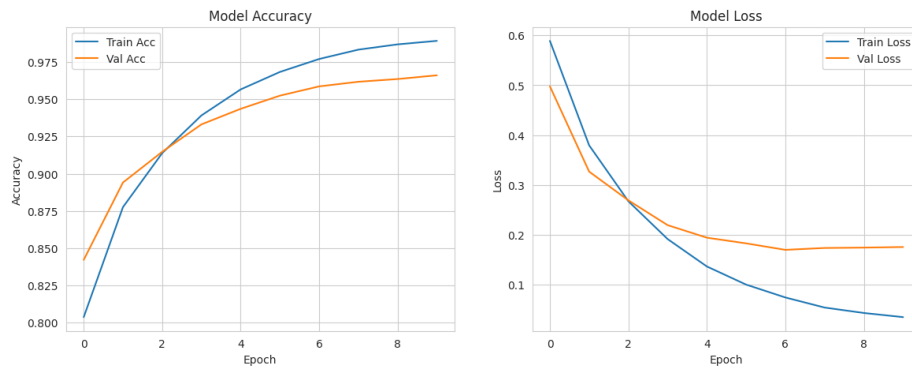


Fig. 7: LSTM training plots for accuracy and loss

Overall, the results clearly indicate that deep learning models, particularly LSTM, are more effective than traditional machine learning approaches for narrative-based customer complaint classification in the banking domain.

We have deployed the lstm model on streamlit, the web application for Bank Customer Complaint Classifier is shown below.

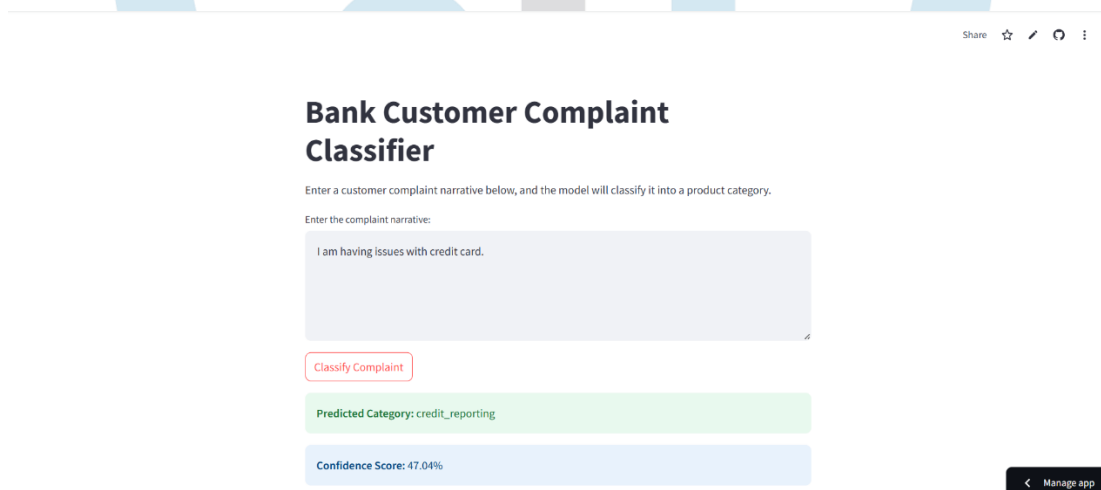


Fig. 8: Bank Customer Complaint Web Application

## V. CONCLUSION AND FUTURE WORK

This study presented a comprehensive framework for analyzing and classifying bank customer complaints using ML and DL techniques. Multiple classification models were evaluated to assess their effectiveness in categorizing complaints by product type. The experimental results show that while traditional ML models such as Logistic Regression and Naïve Bayes perform reasonably well, DL model - LSTM neural network achieve significantly higher classification accuracy, 96.61%. The findings shows that credit reporting-related issues dominate customer complaints, emphasizing the need for banks to prioritize improvements in this area.

In future, we can enhance this research based on following points:

1. Sentiment Analysis Integration can be done to identify customer emotions such as frustration, urgency, or dissatisfaction, enabling better complaint prioritization.
2. Topic Modeling for Interpretability can be performed using Latent Dirichlet Allocation (LDA) or BERTopic to uncover hidden sub-topics within complaint categories for deeper insights.
3. Advanced Deep Learning Models such as BERT, RoBERTa, or domain-specific LLMs can be explored to further improve classification accuracy and contextual understanding.
4. Multi-label Complaint Classification can be done for various products simultaneously.
5. Explainable AI (XAI) techniques such as SHAP or attention visualization can be integrated to help banks understand model decisions and improve transparency.

By addressing these directions, the proposed system can evolve into a more intelligent, interpretable, and customer-centric complaint management solution for modern banking environments.

**REFERENCES**

- [1] A. Singh and A. Singh, "Complaints and Their Redressal: A Study on the Indian Banking System," *Indian Journal of Finance*, vol. 15, no. 3, pp. 23–38, Mar. 2021, doi: 10.17010/ijf/2021/v15i3/158127.
- [2] W. Hakiri, "For an Efficient Complaints Management System for Banks: A Conceptual Framework and an Exploratory Study," *Journal of Marketing Research & Case Studies*, vol. 2012, Article ID 624789, pp. 1–11, 2012, doi: 10.5171/2012.624789.
- [3] M. H. Siddiqui and S. N. Tripathi, "An analytical study of complaining attitudes: With reference to the banking sector," *Journal of Targeting, Measurement and Analysis for Marketing*, vol. 18, no. 2, pp. 119–137, May 2010, doi: 10.1057/jt.2010.2.
- [4] S. Chugani, G. K., and S. Ramasubbareddy, "Data Analysis of Consumer Complaints in Banking Industry using Hybrid Clustering," in *Proc. 2nd Int. Conf. Computing Methodologies and Communication (ICCMC)*, Vellore, India, 2018, pp. xxx–xxx, doi: 10.1109/ICCMC.2018.8487532.
- [5] R. K. Uppal, "Customer complaints in banks: Nature, extent and strategies to mitigation," *Journal of Economics and International Finance*, vol. 2, no. 10, pp. 212–220, Oct. 2010.
- [6] C. Chandana, N. Neelashree, G. N. Nikitha, J. Nisargapriya, and J. Vishwesh, "Bank customer complaints analysis using natural language processing and data mining," *International Journal of Progressive Research in Science and Engineering*, vol. 1, no. 3, Jun. 2020.
- [7] Athira, A. M. Adith, and D. Gupta, "Effective complaint detection in financial services through complaint, severity, emotion and sentiment analysis," *Procedia Computer Science*, vol. 258, pp. 2220–2231, 2025, doi: 10.1016/j.procs.2025.04.47.
- [8] T. S. Roy, G. Vasukidevi, T. N. Malleswari, S. Ushasukhanya, and N. Namratha, "Automatic classification of railway complaints using machine learning," *E3S Web of Conferences*, vol. 477, Art. no. 00085, 2024, doi: 10.1051/e3sconf/202447700085.
- [9] K. Akshitha, K. Susmitha, K. L. Praneetha, A. Jaswanth, Z. Patel, and B. A. Vaghela, "Customer complaint classification system: An automated approach using machine learning," *International Research Journal of Modernization in Engineering Technology and Science*, vol. 7, no. 3, Mar. 2025.
- [10] A. Firmansyah, "Classification of customer complaints with case study contact center PLN 123," unpublished manuscript, Master's Program in Informatics, School of Electrical Engineering and Informatics, Institut Teknologi Bandung, Bandung, Indonesia, 2024.