Modeling a Software Solution for Analyzing Customer Complaints in the Banking Sector

Rozita Oskouei[1], Ramzi Hajiyev[2], and Shahnaz Yagubova[3]

[1]Department of Computer Science and İnformation technology, Islamic Azad University, Mahdishahr, Iran

[2]Department of İnformation technology and Programming, Baku State University, Baku, Azerbaijan

[3]Department of İnformation technology, Western Caspian University, Baku, Azerbaijan

Received: 26.05.2026

Accepted: 01.06.2026

Published: 29.06.2026

https://doi.org/10.54414/MCDL7439

Abstract: The quick advancement in the creation of digital banking systems results in the emergence of more text data, such as complaints, reviews, and service requests, from consumers. The need for systems able to deal with a large amount of data in a sophisticated manner is vital in this situation. Studies show that NLP, together with machine learning, allows financial organizations to gather important information through analyzing texts produced during client-server interactions and thus make better decisions. Automated analysis of complaints made by bank consumers will lead to the identification of service problems and patterns of client discontent. It has been revealed by recent studies that sophisticated classification algorithms and neural network technologies are able to correctly classify complaints made by customers according to predefined categories of services offered by banks, thus minimizing efforts spent on manual analysis and response time. In addition, sentiment analysis techniques make it possible to determine the attitude of customers towards various services provided by banking institutions, which is helpful for effective planning purposes. Topic modeling techniques may also assist in uncovering hidden topics within a set of complaints.

Keywords: Customer complaints, Natural Language Processing (NLP), Machine Learning, Automated Complaint Analysis, Digital Banking

1.Introduction

The banking industry is experiencing accelerated digitalization due to advances in technology, internet penetration, and the development of online financial services. Modern banks gather large volumes of data about their customers via mobile applications, emails, chat conversations, and reviews on feedback forms. Most of the received data comes in the form of unstructured text (complaints, surveys, reviews) that cannot be easily processed manually but contains useful information for enhancing services and making business decisions [1]. Due to the efforts made by financial regulators to improve customer experience and keep banks competitive, there is an increased need for automatically analyzing textual data.

There are currently efficient natural language processing and machine learning solutions that allow for the automatic classification of texts, identification of sentiment polarity in customer feedback in the sentiments expressed by clients, discovery of urgent cases, and detection of common issues [2]. Automatic solutions for dealing with customer complaints can greatly decrease the workload and time required for analyzing textual messages and make decision-making more accurate and efficient [3]. Additionally, intelligent classifiers can be used by banks to properly direct customer inquiries to relevant departments and optimize the workflow [4].

Recent studies have shown that modern deep learning architectures, particularly those based on neural networks, surpass conventional rule-based and statistical methods when applied to financial text analysis [5]. Modern technologies enable algorithms to learn certain language peculiarities and context-based semantics in customer complaints and are well-suited for analyzing large amounts of information, where manual analysis would be impossible [6]. Moreover, sentiment analysis tools enable financial companies to analyze customer perceptions of services and their reaction to services, which helps them to improve customer service and maintain their brand image [7].

However, despite all these technological advancements, most banks operate using semi-automated and decentralized customer complaint management systems that do not allow them to make use of the data they receive. There is a great number of solutions available on the market, but very few provide an integrated and intelligent decision support system, which makes the process ineffective and time-consuming [8]. Thus, there is a growing demand for a solution that integrates data pre-processing and linguistic analysis.

The goal of the current study is to develop a conceptual model for software that can be used for automated analysis of customer complaints submitted by bank customers. This solution will incorporate modern techniques of NLP together with machine learning and the principles of system architecture to design an intelligent system capable of analyzing customer interactions. The purpose of the current research is to enhance intelligent banking support systems by developing a structure that can be easily expanded.

2. Proposed Conceptual Model

Modern banks are involved in communication with an enormous number of customers who provide feedback and complain about different problems in the form of text messages through mobile applications, emails, online forms, and even social networks. Processing and analyzing such volumes of textual data is rather complicated for financial institutions because the data is unstructured. However, effective and timely analysis of customer feedback is important for customer satisfaction. Manual processing of complaints can take too much time and is associated with errors that may lead to dissatisfaction and negative consequences [8].

While there is a variety of automated tools for banks, no single solution has been created that would allow for pre-processing the data, comprehending the text, and categorizing it correctly. Semi-automated solutions cannot identify the category of complaints accurately, their priority level, and overall sentiment because of the complexity of such data processing [3]. This problem increases due to the size of datasets and their complexity in terms of semantics because of the nature of textual data [5].

In conclusion, the principal objective is the development of a smart and robust system that automatically analyses, classifies complaints, determines the emotion of a customer, and gives valuable information for managers. It would be wise if such a system were designed with the aim of minimizing workloads, accelerating operations, improving the accuracy of complaint analysis, and supporting data-driven management decisions in the banking industry [2].

In this regard, the proposed research problem is solved by presenting a theoretical model of software that performs automatic analysis of inquiries from bank customers. The developed model uses state-of-the-art NLP techniques, classification algorithms based on machine learning, and visualization elements to develop a smart approach to managing complaints and making decisions [2], [8], [13].

Customer complaint analysis in banks has been extensively investigated in the recent past, proving the significance of text processing technology in enhancing service delivery and operations in the banking sector. Earlier approaches to customer complaint categorization were mainly rule-based or keyword-matching, which were limited both in terms of effectiveness and scalability [1], [2]. The importance of using automated text processing techniques to categorize complaints in the banking industry and the significant benefits of doing so using basic machine learning models were outlined by Pallavi et al [3], [8].

Natural language processing techniques have become crucial for complaint analysis since then. Kaur [9], pointed out that the implementation of sentiment analysis of e-banking customers' complaints will enable financial institutions to uncover emerging customer dissatisfaction trends and operational issues, and modify their business approach accordingly. Furthermore, Pangaribuan et al.[5], introduced a novel text categorization method utilizing LSTM, which proves significantly more accurate than traditional machine learning approaches and demonstrates the potential of deep learning in text processing [5].

It has been proven in many studies that the importance of sentiment analysis in understanding customer complaints should be stressed. Bhuiyan et al. [7] showed that sentiment analysis enables organizations to assess customer satisfaction and make informed decisions to reduce service-related problems. Furthermore, automated complaint analysis systems classify and prioritize customer complaints, thereby improving decision-making and customer service efficiency [8].

The use of data mining and clustering methods for analyzing the structure of datasets has been widely proposed. In particular, Nikitha et al. [1] demonstrated that data mining techniques applied to banking customer complaints can help improve service quality and identify recurring problem areas. Furthermore, Mutesi et al. [8] emphasized that automated complaint management systems should integrate complaint classification, sentiment analysis, and visualization to support more effective decision-making [8].

LSTM and other neural network models prove promising tools for classifying sophisticated banking texts. According to Baskaran [6], these models yield better results compared to conventional rule-based and statistical approaches, particularly in large environments comprising various types of complaints. Moreover, Mitta [11] demonstrated the value of AI-based sentiment analysis for capturing subtle nuances in customer opinions and supporting more accurate decision-making [11].

Furthermore, the topic modeling approach and feature extraction have also been tested on complaint classification using thematic approaches. In particular, Bastani et al. [12] employed Latent Dirichlet Allocation (LDA) to identify recurring complaint themes, while García-Méndez et al. [4] successfully applied a Support Vector Machine (SVM) approach to transaction-related text classification.

In summary, it is evident from the literature that the integration of NLP, ML, and DL techniques facilitates the automatic and precise analysis of complaints received by banks on a massive scale. However, most of the approaches used today still do not incorporate a common platform that incorporates such steps as data preprocessing, feature extraction, classification, sentiment analysis, and visualization [10]. This deficiency prompted the development of the present study and its proposal for a conceptual software model of complaint management.

In particular, the methodology employed in this paper seeks to create a conceptual software model that can facilitate the analysis of inquiries received by financial institutions based on the use of modern NLP technologies together with ML and DL for effective classification and visualization of complaints. The methodology comprises the following five steps: data collection, data preprocessing, feature extraction and classification, and data visualization and decision support.

1. Data Gathering

The system gathers data about customer complaints and feedback through various digital channels such as mobile banking applications, email, web-based forms, and social media sites [1]. The model uses historical complaint datasets to be trained, while incoming datasets will continuously undergo processing for classification purposes. Quality data and a complete dataset are needed for the successful operation of the model [8].

2. Data Preprocessing

There are different ways to preprocess the raw data that will be fed into the machine learning model to classify complaints. Such methods are listed below:

1. Text normalization: lowering case and eliminating punctuation marks, digits, or other unnecessary characters [2].
2. Tokenization: breaking text into tokens or relevant chunks [9].
3. Stopword elimination: taking out words not applicable to the classifications [9].
4. Stemming and lemmatization: bringing down the words into their base or root form [11].

These procedures are used to filter noise in the data as well as increase the effectiveness of extracting features and training the model.

3. Feature Extraction and Classification

Having preprocessed the data, features can be extracted via both classical techniques and new advancements, which include:

Bag of Words (BoW) and TF-IDF: conventional ways of converting text data into numbers by taking into account the frequency and significance of words [1].
Word vector representations (Word2Vec, FastText): embedding words as dense vectors indicating their meaning.
Contextual representation via neural networks (LSTM, Bi-LSTM): considering the sequential nature of complaints as well as their meanings [5], [6].
Techniques for classification involve multiple models, both machine and deep learning-based, such as:
SVM algorithms are a strong classifier [13].
Ensembles of random forests and gradient boosting trees [8].
LSTM and Bi-LSTM models capture textual sequences [5], [6]. Machine learning and deep learning models are trained using datasets consisting of complaints and then evaluated in terms of metrics like precision, accuracy, recall, and F1 score [2], [3].

4. Sentiment Analysis

Alongside classification, sentiment analysis is done to evaluate the emotional polarity of customer messages: positive, neutral, or negative. This helps banks to understand trends of customer dissatisfaction and handle complaints promptly [7], [11]. The use of advanced NLP techniques allows combining semantic and emotional aspects in order to improve the classification of complaints.

5. Visualization and Decision Support

Ultimately, all analyzed data is presented visually through dashboards and reporting solutions. Visual analytics may involve the following indicators: number of complaints per category, changes in sentiments over time, and urgency [8], [10]. These insights will help managers of banks make decisions based on data and manage complaints effectively.

Overall, this methodology offers a complete theoretical framework for complaint classification automation in banking companies. Combining data processing, feature extraction, classification, sentiment analysis, and visualization, this approach makes complaint analysis scalable, accurate, and effective [2], [12].
The conceptual software model designed for analyzing automated inquiries of customers of banks is built around a modular architecture that consists of several modules, such as data acquisition, data processing,

feature extraction, classification, sentiment analysis, and visualization. The proposed system contains five major modules that contribute to efficient complaints management and decision-making.

1. Data Collection Module

This module gathers customer complaints received through multiple channels, such as mobile banking applications, emails, online forms, and social networking sites [1]. Historical databases are used as training data, and live complaint feeds are constantly monitored for prompt processing. This module guarantees the quality of collected data by eliminating duplicates and preparing the collected data for pre-processing [8].

2. Data Preprocessing Module

The data pre-processing module converts unstructured text data into structured data ready for machine learning and NLP techniques. Tokenization, elimination of stopwords, stemming, lemmatization, and normalization are some of the processes included in data pre-processing [2], [9], [11]. This module ensures that text data is prepared for feature extraction and classification.

3. Feature Extraction and Classification Module

The basic functionality of this module is to represent text data numerically and categorize complaints based on predefined categories:

Techniques used for feature extraction: BoW (Bag-of-Words), TF-IDF, Word2Vec, FastText, contextual embedding for use in deep learning models [1], [14].
Algorithms for classification: SVM (Support Vector Machine), Random Forest, Gradient Boosting, and LSTM/Bi-LSTM networks [5], [6].
The output from classification includes the category of the complaint, priority level, and related features.

4. Sentiment Analysis Module

This module combines with the classification process to evaluate the emotional nature of the complaint made by the customer. The system determines whether a complaint contains positive, neutral, or negative emotions [7], [11]. Assessment of emotional states is then integrated with classification outcomes.

5. Module for Visualization and Reporting

This module generates dashboards and reports that give the following data:

1. Complaint classification;
2. Trends of sentiment by time;
3. Priority analysis;
4. Department workload overview [8],[10].

The visualization of complaint data aids in strategic decision-making and performance tracking. The suggested system operates like a pipeline that collects, pre-processes, extracts features from, classifies, analyzes the sentiment of, and visualizes data. This way of creating a system provides flexibility and modularity, which allow it to scale and make the system resilient to changes. New approaches for classification and sentiment analysis can be easily implemented into the system without breaking anything [2], [12].

Automation will help decrease manual labor and shorten response time. Accuracy will improve thanks to the application of deep learning for classification and sentiment analysis tasks [5], [6]. Making decisions based on data will be facilitated by using dashboards and reporting tools. Scalability makes it possible to integrate more data from different bank channels.

Such an architecture shows how one intelligent system is able to make the process of managing customer complaints more effective, accurate, and strategic in banks, thus enhancing the quality of services and the level of customer satisfaction [3], [12],[14].

The proposed software architecture was theoretically assessed based on findings reported in previous studies. Existing research suggests that LSTM models can achieve competitive performance in customer complaint classification compared with conventional machine learning methods such as SVM and Random Forest [5], [6]. Based on the findings reported in the literature, sentiment analysis can help identify negative customer feedback and support improvements in customer satisfaction [7], [11].

By using visual tools, information about complaints could be obtained, including their sentiment analysis and the workload of certain departments [8], [10]. In general, the combination of NLP techniques, machine learning algorithms, and visualization made the proposed software solution effective and scalable. Thus, the results of the theoretical evaluation prove the viability of the proposed intelligent system [2], [3], [12].

Conclusion

This paper develops an abstract software model for automatic analysis of customer complaints in banks using natural language processing (NLP), machine learning, deep learning, and visualization algorithms. With the help of data processing, classification, and analysis of textual information in an unstructured form, this system will be able to minimize manual labor, optimize processing time, and enhance efficiency [3], [5], [6], [7], [8].

According to previous research and the proposed conceptual framework, NLP models based on LSTM networks may provide effective performance in classifying and prioritizing customer complaints [5], [6], [7]. The use of visualizations facilitates data-based decision-making and strategy development by ensuring a deeper insight into patterns, types, and sentiments found in complaints [5, 6].

Thus, the developed model appears to be effective, intelligent, and applicable for solving issues of customer service and management at financial organizations. Future research may be aimed at investigating the performance of this model on real-life datasets [2], [12], [14]

References

[1] Nikitha, G. N., Chandana, C., Neelashree, N., Nisargapriya, J., & Vishwesh, J. (2020). Bank customer complaints analysis using natural language processing and data mining. *Internat J Prog Res Sci Engin*, *1*(3), 22-25.

[2] Mishra, P., Agrawal, C., & Envey, D. (2023). Advancements and Challenges in Sentiment Analysis: A Comprehensive Review of Frameworks and Future Trends. *International Journal of Engineering Applied Science and Management (IJEASM)*, *4*(12), 1-8.

[3] Pallavi, S. A., Huzaifb, M., Ishaq, A., Reddy, K., & Ahmed, M. (2025). Automated Complaint Ticket Classification in Banks. *Journal of Science, Technology, and Management Innovations*, *1*(01), 57-70.

[4] García-Méndez, S., Mahía, D., & Méndez, J. R. (2024). *Bank transaction text classification using SVM*. arXiv

[5] Pangaribuan, T., Sitorus, A. S., & Simangunsong, A. (2024). LSTM-based automatic classification of bank customer complaints. *JIPI (Jurnal Ilmiah Penelitian dan Pembelajaran Informatika), 9*(1), 150-159.

[6] Baskaran, P. K. (2025). *Enhancing Customer Complaint Classification in Banking: A Deep Learning and Natural Language Processing Approach* (Doctoral dissertation, Dublin, National College of Ireland).

[7] Bhuiyan, R. J., Akter, S., Uddin, A., Shak, M. S., Islam, M. R., Rishad, S. S. I., ... & Hasan-Or-Rashid, M. (2024). Sentiment analysis of customer feedback in the banking sector: A comparative study of machine learning models. *The American Journal of Engineering and Technology*, *6*(10), 54-66.

https://doi.org/10.37547/tajet/volume06issue10-07

[8] Mutesi, D., Ngugi, J., & Djuma, S. (2025). *Enhancing Customer Service in Financial Institutions Through Automated Complaint Classification: A Machine Learning Approach*. Journal of Information and Technology, 5(8), 61–71. https://doi.org/10.70619/vol5iss8pp61-71

[9] Kaur, N. (2021). Sentiment Analysis of E-Banking Customer Reviews Using Nlp. *ShodhKosh J. Vis. Perform. Arts*, *2*(2), 458-465. https://doi.org/10.29121/shodhkosh.v2.i2.2021.5743

[10] Suresh, M., Vincent, G., Vijai, C., Rajendhiran, M., Com, M., Vidhyalakshmi, A. H., & Natarajan, S. (2024). Analyse customer behaviour and sentiment using natural language processing (NLP) techniques to improve customer service and personalize banking experiences. *Educational Administration: Theory And Practice*, *30*(5), 8802-8813.

[11] Mitta, N. R. (2023). Leveraging natural language processing (NLP) for AI-based sentiment analysis in financial markets: real-time insights for trading strategies and risk management. *African Journal of Artificial Intelligence and Sustainable Development*, *3*(2).

[12] Bastani, K., Namvar, A., & Kafashpoor, Z. (2018). *Topic modeling of financial consumer complaints using LDA*. arXiv.

[13] Tang, X., Mou, H., Liu, J., & Du, X. (2021). Research on automatic labeling of imbalanced texts of customer complaints based on text enhancement and layer-by-layer semantic matching. *Scientific Reports*, *11*(1), 11849. https://doi.org/10.1038/s41598-021-91189-0

[14] Gavval, R., Ravi, V., Harshal, K. R., Gangwar, A., & Ravi, K. (2019). CUDA-Self-Organizing feature map based visual sentiment analysis of bank customer complaints for Analytical CRM. *arXiv preprint arXiv:1905.09598*. https://doi.org/10.1049/pbpc037g_ch12