

Prediction and Analysis of Customer Complaints Using Machine Learning Techniques

Ghadah Alarifi, Department of Business Administration, College of Business Administration, Princess Nourah Bint Abdulrahman University, Riyadh, Saudi Arabia

Mst Farjana Rahman, Hajee Mohammad Danesh Science and Technology University, Bangladesh

Md Shamim Hossain, Hajee Mohammad Danesh Science and Technology University, Bangladesh*

 <https://orcid.org/0000-0003-1645-7470>

ABSTRACT

Businesses must prioritize customer complaints because they highlight critical areas where their products or services may be improved. The goal of this study is to use machine learning approaches to anticipate and evaluate customer complaint data. The current study used logistic regression and support vector machine (SVM) to predict customer complaints, and evaluated the datasets using machine learning techniques after collecting five distinct length datasets from the Consumer Financial Protection Bureau (CFPB) website and cleaning the data. Both logistic regression and SVM can accurately predict customer complaints, according to this study, but SVM gives the greatest accuracy. The current study also found that SVM provides the highest accuracy for a one-month dataset and Logistic regression provides for a three-month dataset. In addition, machine learning codes were utilized to display and tabulate consumer complaints across many dimensions.

KEYWORDS

Customers' Complaints, Customer Complaints Analysis, Logistic Regression, Machine Learning, Prediction, Support Vector Machine

INTRODUCTION

According to London (1980), a customer complaint can be seen as a form of feedback that expresses the customer's dissatisfaction with a product or service. The act of making a complaint is often regarded as an indication that the customer cares about the quality of the product or service and wants to see improvements. From the organization's perspective, a complaint is an opportunity to address the customer's concerns and make things right, potentially turning a dissatisfied customer into a loyal one. Therefore, it is important for responsible organizations to take customer complaints seriously and respond to them promptly and appropriately, as it can have a significant impact on their reputation and long-term success. Complaints are an opportunity for organizations to collect

DOI: 10.4018/IJEBR.319716

*Corresponding Author

This article published as an Open Access article distributed under the terms of the Creative Commons Attribution License (<http://creativecommons.org/licenses/by/4.0/>) which permits unrestricted use, distribution, and production in any medium, provided the author of the original work and original publication source are properly credited.

information regarding a consumer's opinions, needs, beliefs, and attitudes. Therefore, the complaint can also be considered in an optimistic sense as providing problem-related documentation of customers regarding a product or service. The emerging developments in international markets have enlarged the responsiveness of an organization to its consumers (Sitko-Lutek et al., 2010; Atalik, 2007). The promptness and effectiveness of customer complaint handling is a vigorous portion of this responsiveness of the organization (Dey et al., 2009). Despite early displeasure, if a consumer complaint is managed and handled appropriately, a business can keep company goodwill and steadily improve long-term relations with customers (Sitko-Lutek et al., 2010; Reichheld and Sasser, 1990). Filip (2013) described that consumer complaints should be measured as a vigorous indicator of business performance valuation, signing difficulties, or failures in core business processes that need swift repossession to shun migration of profitable clients. Businesses must absorb that the costs of losing clients are both profit decline and adverse word of mouth.

Customer complaint managing conquers a fundamental role in the process of customer relationship management (CRM) and positions complaint management as an imperative strategic tool for organizations of all categories (Hossain, 2023; Strauss and Hill, 2001). Knowing the value of complaints about refining the offerings of the organization and building long-term sustainable relationships is a recognized marketing practice. Effectually handling customer complaints rises the probability of superior buyer satisfaction and, consequently, repeat patronage behavior while lessening bad word-of-mouth (Strauss and Hill, 2001; Blodgett, Granbois, & Walters, 1993). Besides, there is evidence that longtime buyers are more profitable since they tend to buy in larger quantities and more repeatedly than new buyers (Reichheld & Sasser, 1990). Complaint management provides firms with a priceless chance to discover areas for development. Allowing such comments of displeasure to be answered with positive responses and changes is critical for any successful company. As a result, when businesses respond to consumer complaints and provide a timely remedy, they are more likely to make more than one consumer happy (Hossain, 2023). According to Huang et al. (2018), when complaints are handled quickly, clients become more happy with the service and loyal to the firm.

The United States' Consumer Financial Protection Bureau (CFPB) was created to help American financial customers report complaints and access support related to their financial matters. As a federal agency, the CFPB is responsible for ensuring that banks, lenders, and other financial institutions treat their customers fairly (Hossain, 2023). By providing a platform for consumers to report complaints and issues, the CFPB plays a vital role in protecting consumers' rights and interests in the financial sector. In summary, the CFPB is a government organization in the US that is dedicated to safeguarding consumer rights and promoting fair practices in the financial industry.

Financial complaint data is readily available for tracking and analyzing how effectively and efficiently financial organizations respond to customer complaints. Each complaint contains various attributes that can be identified and described individually. These features of complaint data have been used for data analysis and predictive purposes, as noted by Fonseka et al. (2016). Managing and comprehending customer complaints is crucial for any business, as emphasized by Strauss and Hill (2001). By utilizing complaint data, financial organizations can identify areas for improvement in their services and address customer concerns more effectively, which can lead to increased customer satisfaction and loyalty.

Specifically, we are looking for answers to the following research questions: (i) How can we use machine learning techniques to predict customer complaints? (ii) Which machine learning models do the best in predicting consumer complaints across various periods?, (iii) What are the most common issues and sub-complaints for different periods? , (iv) Which businesses and products are the most frequently criticized? And (v) how do businesses handle customer complaints?

In this circumstance, the current study aims to predict the complaint behavior of the customers in the financial sector through logistic regression and support vector machine (SVM), and also analyze customer complaints data through machine learning techniques. Both logistic regression and SVM are supervised machine learning methods (Hossain and Rahman, 2022) used in this research because

of their usefulness in classification. Logistic regression is a comparatively straightforward method that is suitable to interpret the outcomes. On the other hand, SVM is a powerful and flexible method for performing classification, regression, and detection of an outlier. Furthermore, the current study contributes to the application of machine learning tools in the financial sector, with the overall goal of improving customer service.

RELATED WORK

The prediction and analysis of customer complaints using various machine learning techniques is a rapidly growing area of research. In today's globally competitive marketplace, organizations must prioritize the needs and wants of their customers in order to sustain and grow (Hossain and Rahman, 2023; Hossain and Rahman, 2022, Hossain and Rahman, 2021). By investing in customer support, organizations can gain a better understanding of their customers' complaints and use this information to enhance their products, services, and overall customer experience. In fact, as noted by Hsiao et al. (2016), customer complaints can be viewed as an opportunity for businesses to improve and differentiate themselves from competitors. However, it's important to recognize that complaints are inevitable. No matter how hard an organization tries or how excellent their offerings are, it's impossible to satisfy every single customer. As a result, complaints can arise in any type of organization, even those providing great services. Nevertheless, by leveraging modern tools and techniques for analyzing and predicting customer complaints, businesses can take proactive measures to address customer concerns and minimize negative feedback. This can ultimately lead to increased customer loyalty and improved business performance.

Day (1980) mentioned that behaviors of customer complaints are triggered by perceived dissatisfaction. Customer dis/satisfaction is usually observed as an individual assessment of the gap between expectations and actual consequences, where undesirably disconfirmed customer expectations lead to customer dissatisfaction (Chang et al., 2011; Day, 1984). Complaints of customers form an acute foundation of information for refining an organization's goods and services. Once organizations' products or services can't satisfy customers' anticipation, customers of that organization are very likely to become disappointed. Customer complaint behavior theory described that most disappointed customers frequently withdraw their support and prompt adverse views on products or services of an organization to other customers. Only minor proportions of unhappy buyers make a complaint and trust in the organizations' capability to resolve their complications impartially (Yang et al., 2018; Hsiao et al., 2016; Jugwanth & Vigar-Ellis, 2013). Managing these customer complaints helps effectively not only recognize the defects in goods or services, but also sustain buyer loyalty.

Customer complaint management (CCM) is a practice of recording and solving complaints of customers. Organizations repeatedly improve customer relationship management (CRM) to manage customer complaints and construct solutions and recommendations (Johnston, 2013). Receiving, handling, and encouraging complaints, and providing a response to customers are vital functions embedded in the system of CRM of organizations. Thus, organizations may advance numerous channels to inspire customers to express their opinion vigorously (Yang et al., 2018).

Birim et al. (2016) proposed a model for studying customer complaints in the airline sector and their impact on organizational performance, with a focus on variables such as service quality, fees, and economic conditions. The authors used these variables to predict complaints that could affect future ticket purchases by airlines. Hsiao et al. (2016) integrated a decision tree algorithm into the Six Sigma investigation toolset to enhance the efficiency of handling customer complaints and refining service quality. Chugani et al. (2018) investigated customer complaints in various banks and recommended a data mining model for identifying and resolving problems. Ghazzawi and Alharbi (2019) applied data mining techniques to analyze customer feedback for the Metropolitan Transportation Authority (MTA). Xu et al. (2018) employed professional knowledge and a K-means

algorithm to group customer complaints in the mobile service sector. Meanwhile, Yang et al. (2018) developed a complaint classification model for a telecommunication organization.

In this study, we aim to predict and analyze customer complaints in the financial sector using two machine learning techniques: logistic regression and support vector machines (SVM). By leveraging these tools, we hope to gain insights into the factors that influence customer complaints in this industry and identify opportunities for improving customer satisfaction and loyalty. By studying the work of other researchers in related fields, we hope to build on their findings and contribute to the growing body of knowledge on customer complaints and machine learning.

Logistic Regression

Logistic regression is a statistical technique used in regression analysis to determine the relationship between input and output variables (Hossain et al., 2022). It's a type of regression model that predicts how one or more independent variables affect a dependent variable. The dependent variable can be binary or multinomial, depending on the range of potential values. In binary logistic regression, the dependent variable's values are typically represented as zeros and ones to indicate the two potential outcomes (Al-Mashraie, Chung, and Jeon, 2020). Logistic regression is a well-established independent predicting methodology used in various fields, including marketing (Hossain and Rahman, 2022; De Caigny et al., 2018).

One of the advantages of logistic regression is that it can provide interpretable results. It can estimate the effect of each independent variable on the dependent variable and quantify the magnitude of the effect. This makes logistic regression a useful tool for understanding the factors that influence a particular outcome. Another advantage of logistic regression is that it is relatively easy to implement and interpret. It is a parametric method that assumes a functional form for the relationship between the independent and dependent variables. The logistic function, also known as the sigmoid function, is used to model the probability of the dependent variable being in one of the two categories. The model parameters can be estimated using various optimization algorithms, and the model's goodness-of-fit can be evaluated using various statistical tests. In summary, logistic regression is a powerful statistical tool that can be used to model the relationship between input and output variables (Hossain and Rahman, 2022). It is particularly useful when the dependent variable is binary, and it provides interpretable results that can help us understand the factors that influence a particular outcome.

Support-Vector Machine (SVM)

A support-vector machine, also known as a support-vector network, is a type of supervised machine learning model that is commonly used for classification and regression analysis (Hossain et al., 2022; Boser, Guyon, & Vapnik, 1992). It uses learning algorithms to examine data and identify patterns, and kernel functions are employed to improve the classifier's accuracy. One of the most critical features of SVM is its ability to efficiently map inputs using both linear and nonlinear classification through kernel functions such as linear, polynomial, radial basis function, and sigmoid (Hossain and Rahman, 2022). Although this method usually yields better results, the process of building the prediction model takes longer than other methods because it is an optimization process (Al-Mashraie, Chung, and Jeon, 2020). SVM has been used in various fields, including image analysis, speech recognition, and finance, due to its effectiveness in complex data classification and its ability to handle high-dimensional data (Hossain et al., 2022).

In practice, SVM and Logistic regression perform similarly (Hossain et al., 2022). Ing et al. (2018) mentioned that SVM are a more recent statistical tool that has been shown to outperform classical logistic regression. They also describe that for prediction, logistic regression and support vector machines are helpful classification algorithms. The simpler link between the inputs and outputs is known as logistic regression, although it is susceptible to outliers in the data and does not uncover non-linear correlations. SVM and other machine learning algorithms are getting more popular. SVM with non-linear kernels is a "black box" technique capable of producing more complicated,

multidimensional decision boundaries than logistic regression (Pochet et al., 2016). Hazra et al. (2017) used both support vector machine (SVM) and logistic regression (LR) algorithms to predict lung cancer patients' survival rates, and compared the performance of the two algorithms using accuracy, precision, recall, F1 score, and confusion matrix. They discovered that the logistic regression classifier has the highest accuracy of 77.40 percent, compared to the 76.20 percent accuracy of the support vector machine classifier. The LR and SVM algorithms are currently used in a lot of research. For example, both LR and SVM have been used to assess people's sentiment during COVID-19's lockup period (Majumder et al., 2021), to detect abnormal gait (Chakraborty et al., 2021), and to assess Ovarian cancer major risk factors (Ahmed et al., 2021). Unfortunately, the performance of LR and SVM machine learning algorithms in predicting customer complaints for diverse time period datasets is unclear, and the performance of LR and SVM for complaint datasets has not yet been studied. Thus we applied LR and SVM in our current project.

Our inquiry involved examining the complaints dataset of the Customer Financial Protection Bureau (CFPB). Several studies have also analyzed the CFPB's complaints data from various perspectives. Ayres et al. (2013) conducted a preliminary assessment of the CFPB's consumer complaints on a company-by-company basis and also looked at zip code demographics. On the other hand, Littwin (2015) explored why government institutions should review CFPB consumer complaints and whether the benefits justify the resources needed to do so. Her analysis, based on regression techniques, demonstrated that the CFPB has been successful in settling consumer disputes, educating the public about regulatory operations, and enhancing the agency's reputation.

Bertsch et al. (2020) utilized the CFPB complaint data to construct a high-quality proxy for bank fraud. Bastani et al. (2019) proposed an intelligent technique based on latent Dirichlet allocation (LDA) to analyze CFPB consumer complaints. Their approach involved extracting latent subjects from complaint narratives and examining their tendencies over time. HAYES et al. (2021) analyzed the CFPB's data to find that a higher level of trust in a location is associated with a lower number of complaints filed against financial institutions in that area. They also discovered that banks in low-trust areas were more likely to lower fees charged to consumers than banks in high-trust areas after the CFPB's establishment. Additionally, they observed that the possibility of consumer complaints being sent to a government agency can influence how banks treat their customers, highlighting the impact of stakeholders on corporate policy and the interaction between informal culture and formal institutions.

Although the Consumer Financial Protection Bureau's complaints dataset has been extensively used in research, machine learning techniques have not been employed to predict and analyze consumer complaints data, particularly those from the CFPB. In addition, there has been limited attention given to issues such as consumer complaint prediction and analysis when reporting complaints, despite the growing interest in customer complaint behavior. The aim of this study is to employ machine learning techniques to predict and analyze customer complaints data for various time periods datasets, as well as to demonstrate how these models and techniques can be used to evaluate textual data from CFPB consumer complaint narratives. To the authors' knowledge, no previous studies have used machine learning approaches to assess customer complaints across multiple datasets and time periods. The current study utilizes machine learning techniques to forecast and evaluate complaints.

METHOD

We downloaded our study data from the website of CFPB (<https://www.consumerfinance.gov>). CFPB allows filtering the dataset based on the time period. We filtered and downloaded five datasets for the period of 1 month, 3 months, 6 months, 1 year, and 3 years on 10, January 2021. Table-1 showed the date ranges and number of complaints of each filtered dataset. In each dataset, there are nine products. We downloaded data from one year to three years because three years ago, the type of products was more than nine. Logistic Regression and SVM classification models using python codes were developed and executed in jupyter notebook. Feature transformation and training-testing data

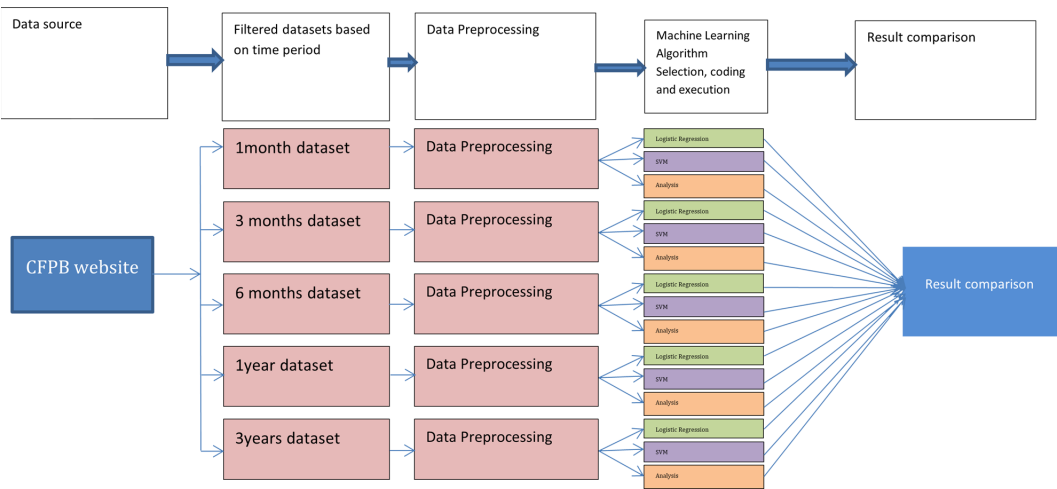
Table 1. Filtered datasets

Dataset	Date range	Number of complaints
1 Month	1 December 2020 to 31 December 2020	45,612
3 Months	1 October 2020 to 31 December 2020	130,956
6 Months	1 July 2020 to 31 December 2020	250,808
1 Year	1 January 2020 to 31 December 2020	440,591
3 Years	1 January 2018 to 31 December 2020	974,757

preprocessing techniques were implemented. The Block diagram of the machine learning implication process of our study is presented in figure-1. After cleaning the missing value from each dataset separately, we imported the required modules and libraries into the Jupyter notebook. Subsequently, we used the Term Frequency Inverse Document Frequency (TF-IDF) vectorizer algorithm to transform the categorical text into numerical data. TF-IDF vectorizer is a very common procedure to convert text into an expressive demonstration of numbers which is implemented to fit the machine learning algorithm for prediction. Following that, the database was divided into two sets for investigation: training data and text data. In our investigation, Logistic Regression and SVM machine learning codes have been executed and the performance of both models for each dataset was evaluated with the following diverse performance standards:

1. The Cohen’s kappa statistic is a valuable measure for calculating interrater reliability, and if the value of the kappa statistic is 0 or lower, the classifier will be unusable (Hossain et al., 2022).
2. The phi coefficient, also known as the Matthews correlation coefficient (MCC), is widely used in machine learning to measure the quality of binary classifications. The values of MCC -1 and +1 indicate total disagreement and perfect prediction, respectively, between the prediction and observation.
3. The primary implication of Cohen’s kappa is to assess the degree of accurate representation of variables for the training model. Values of Cohen’s kappa closer to 1 indicate a good representation of variables for the training model, while values closer to 0 are ambiguous.

Figure 1. Block diagram of machine learning implication process



4. The mean squared error (MSE) is a measure of the quality of an estimator. It is always non-negative, and values of MSE closer to zero are better.
5. The root-mean-square error (RMSE) is also commonly used to measure the differences between the predicted value of a model and the observed values. Like MSE, the value of RMSE is also non-negative, and closer to zero is better.
6. The F1-score is commonly used to measure the effectiveness of a classifier, while the r^2 _score indicates the coefficient of determination. The best possible r^2 _score is 1.0, and a negative result indicates an arbitrarily worse model.
7. The mean squared logarithmic error (MSLE) indicates the degree of regression loss, which can be defined as a measure of the ratio between the true values and predicted values.
8. Precision is the ratio between true positives and all positives, while recall is the ratio between true positives and relevant elements. In addition, precision indicates the degree of validity of results, and recall refers to the degree of completeness of results. These measures are commonly used in machine learning for classification tasks.

Finally, we visualized and measured a few issues using machine learning codes to analyze the customer complaint completely.

RESULT AND DISCUSSION

Model Discussion

In our study, the study results showed that both the machine learning techniques Logistic Regression and SVM outperformed. All complaints of each dataset were allocated into training and testing sets based on the nine types of complaint-related products. Study results showed that in each dimension both models are very close. Both models provided similar and very decent results. SVM accuracy rates for 1 month, 3 months, 6 months, 1 year, and 3 years were 90.402, 86.368, 86.368, 85.931, 85.701, and 84.317, respectively. Logistic regression accuracy rates were 85.308, 86.102, 85.76, 85.68, and 84.242, respectively. Thus, in our study, SVM provides the best accuracy than logistic regression. We also found that SVM provided the highest accuracy for the one-month dataset and Logistic regression provided for the three-month dataset. The precision values of SVM for 1 month, 3 months, 6 months, 1 year, and 3 years were 0.899, 0.860, 0.856, 0.854, and 0.854, respectively, while the accuracy values of Logistic regression were 0.818, 0.857, 0.852, 0.854, and 0.840, indicating that the degree of validity of SVM model results were slightly higher than logistic regression for all datasets.

The recall values of SVM for month, months, months, 1 year, and 3 years were 0.904, 0.864, 0.859, 0.857, and 0.857, respectively, while the recall values of Logistic regression were 0.853, 0.861, 0.857, 0.857, and 0.842, indicating that the degree of completeness of SVM model was slightly higher than logistic regression for all datasets. We also noticed that the precision and recall values of the 1year dataset for both models were exactly the same. The weighted $f1$ _score of month, months, months, 1 year and 3 year dataset for SVM were 0.898, 0.860, 0.856, 0.854, and 0.841 respectively, and for Logistic regression were 0.830, 0.853, 0.852, 0.854, and 0.840 correspondingly. Therefore, we found that the effectiveness of a classifier for both models was excellent and the weighted $f1$ _score of SVM was also slightly greater than logistic regression. The values of MSE, MSLE, and RMSE likewise indicated that both models were worthy of predicting customer complaints (table 2). Moreover, the score of Matthews correlation coefficient and Cohen's kappa for both models indicate that our predictions were the perfect and accurate demonstration of variables for the training model were enhanced.

Customer Complaints Analysis

We also visualized and tabulated customer complaints using machine learning codes based on the diverse dimensions. Box 1 showed the top ten issues for customer complaints. Figure 1 also exhibited that most of the complaints were associated with incorrect information in the report. Box 2 indicated the highest ten sub issues for getting complaints from customers where none indicates that these complaints are not associated with any mentioned sub-issues. Therefore, organizations can know the priority level of diverse issues and sub-issues for maintaining complaints efficiently. Similarly, Box 3, Box 4, and Box 5 presented the names of the top 10 complaint receiving companies, top complaints receiving products, and the responses of the company to customer complaints, correspondingly. Experian Information Solutions Inc., Transunion Intermediate Holding, and Equifax Inc. received more than 75% of all complaints, according to Box 3. Box 4 also revealed that the three products that received the most complaints were debt collection, credit card or prepaid card use, and credit reporting, credit repair services, or other personal consumer reports. Therefore, Box 3 and 4 indicated, respectively, which companies and products are most commonly criticized. Furthermore, Box 5 demonstrated that businesses resolve more than 75% of customer complaints with an explanation, which is encouraging for businesses. Organizations occasionally provide both non-monetary and monetary redress for complaints.

In table-3,4, and 5, we also statistically displayed the company's response to the consumer, provided consumer consent, and timely response to the consumer to a different product of a different dataset. The company's response to customer concerns was shown in Table 3. Closed with explanation, closed with monetary relief, closed with non-monetary relief, in progress, and untimely answer are the five categories of company responses. Credit reporting, credit restoration services, or other personal consumer reports received the most responses, according to our findings. The majority of the customer concerns were resolved with an explanation. Table 4 depicted the organizations' provision of consumer consent; it is evident from this table that organizations do not provide consent to customers in the vast majority of cases. Table 5 also illustrates that, while corporations respond to the majority of complaints in a timely manner, they do not respond in a timely manner in a few situations. Companies are responding to complaints about credit reporting, credit restoration services, and other personal consumer reports in a timely manner since they get the majority of them. Because all concerns are significant, businesses should respond quickly after carefully examining each one.

All of the boxes and tables presented in our study will assist organizations in understanding customer complaint behavior.

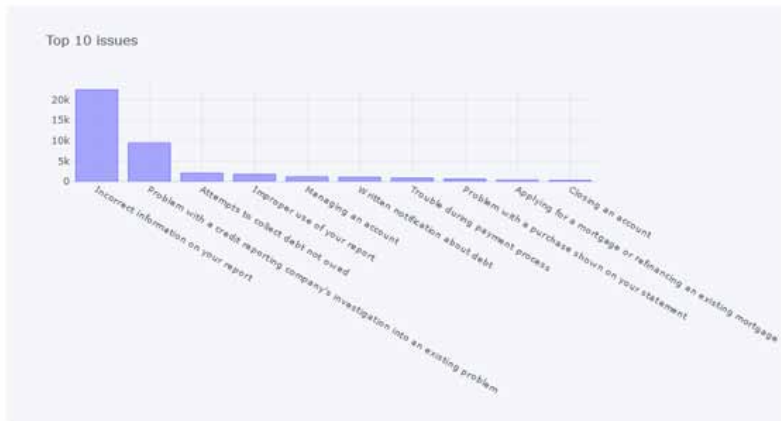
CONCLUSION

At present, financial institutions are facing a variety of challenges and changes, with the need to manage customer complaints being a critical aspect of their operations. Swiftly managing customer complaints in the early stages is crucial for all organizations in preventing various issues that could impact the profitability of the business. Complaint analysis is essential for companies as it involves not only evaluating and understanding complaint data, but also using it effectively for current and future generations. In this regard, the topic of investigating customer complaints has become a substantial study area for organizations to cope with the competition among them. Customer complaints are a regular occurrence for businesses, and it is also a measure of their success, including customer happiness, the number of new customers, and appreciation notes. Thus, managing customer complaints in a thorough manner is crucial to avoid failures and produce product and process improvements. This is precisely where research and practice activities are in high demand, as demonstrated by several studies.

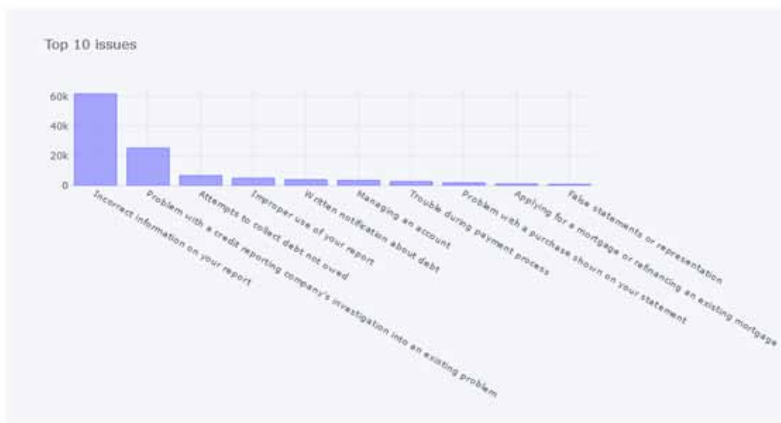
This study has demonstrated the significance of utilizing machine learning algorithms to anticipate and evaluate customer complaint data, with the establishment of logistic regression and SVM models to predict consumer complaints of financial organizations. The results of this study have shown that SVM

Box 1. Top 10 complaint Issue

1 month



3 month



6 month



continued on following page

Box 1. Continued



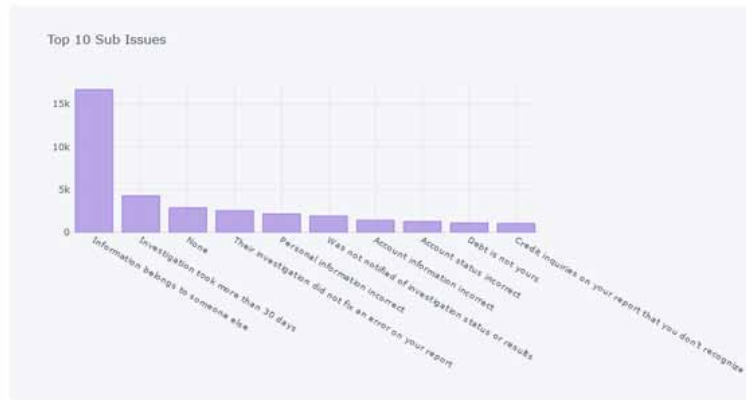
has the highest accuracy in predicting financial users' complaints about products over a one-month dataset, while logistic regression has the most accuracy for a three-month dataset. Additionally, the use of machine learning codes to display and tabulate consumer complaints across many dimensions provides businesses with a better understanding of their customers' needs and preferences, which can aid in improving their products or services. Theoretical and managerial applications of this study have been identified, including the prioritization of customer complaints, improvement of complaint management strategies, and enhancement of customer satisfaction and loyalty. The study's findings demonstrate the potential benefits of utilizing machine learning algorithms to anticipate and evaluate customer complaint data. Overall, businesses can benefit from the insights provided by this study, enabling them to identify patterns and trends in customer complaints, address issues proactively, and ultimately improve their products or services. This, in turn, can result in higher customer satisfaction and loyalty, which is crucial for business success.

Theoretical Applications

The current research has some theoretical implications. Firstly, the study demonstrates the importance of machine learning algorithms in predicting customer complaints over time. By utilizing these

Box 2. Top 10 complain Sub-Issue

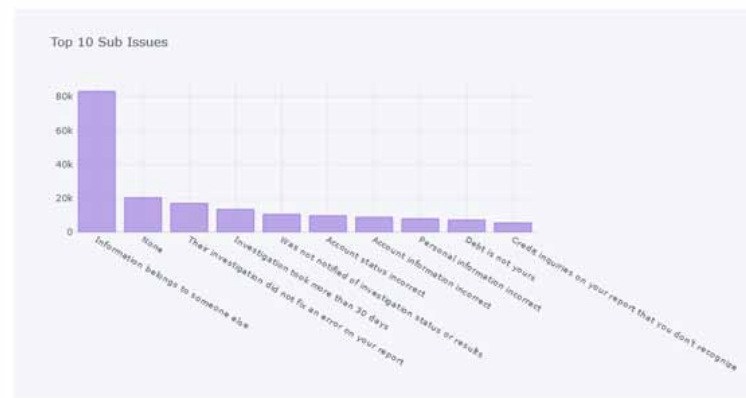
1 month



3 months



6 months



continued on following page

Box 2. Continued

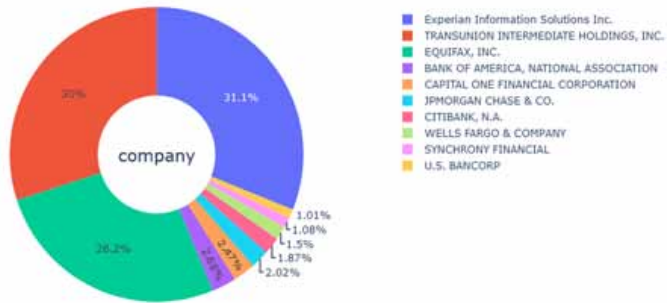


algorithms, businesses can gain a more accurate understanding of their customers' needs and preferences. Secondly, the study highlights the significance of customer complaints as a source of valuable insights for businesses. By prioritizing customer complaints and utilizing machine learning techniques, businesses can gain valuable insights into the critical areas where their products or services may need improvement. Thirdly, the study provides evidence of the usefulness of machine learning approaches such as logistic regression and support vector machine (SVM) in anticipating and evaluating customer complaint data. This is particularly important for businesses looking to proactively address issues and improve their products or services. Fourthly, the study emphasizes the importance of analyzing customer complaints across various dimensions such as product type, location, and time period. By doing so, businesses can identify specific areas that need improvement and tailor their complaint management strategies accordingly. Fifthly, the study demonstrates how machine learning algorithms can be used to visualize customer complaints statistics for huge datasets. This provides businesses with a more efficient and accurate way to manage and prioritize customer complaints. Sixthly, the study highlights the importance of utilizing machine learning techniques to manage customer complaints for financial organizations. This can result in higher customer satisfaction and loyalty, which is crucial for business success. Seventhly, the study provides evidence that machine learning models can accurately forecast consumer complaints over time. This has

Box 3. Top 10 Companies receiving complaints

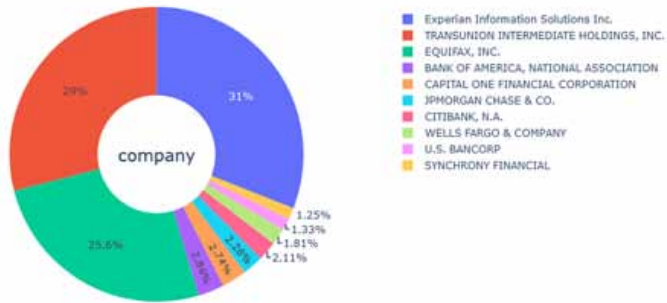
1 month

Top 10 companies receiving most complaints



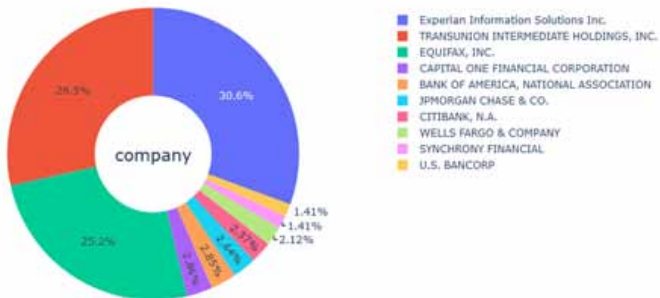
3 months

Top 10 companies receiving most complaints



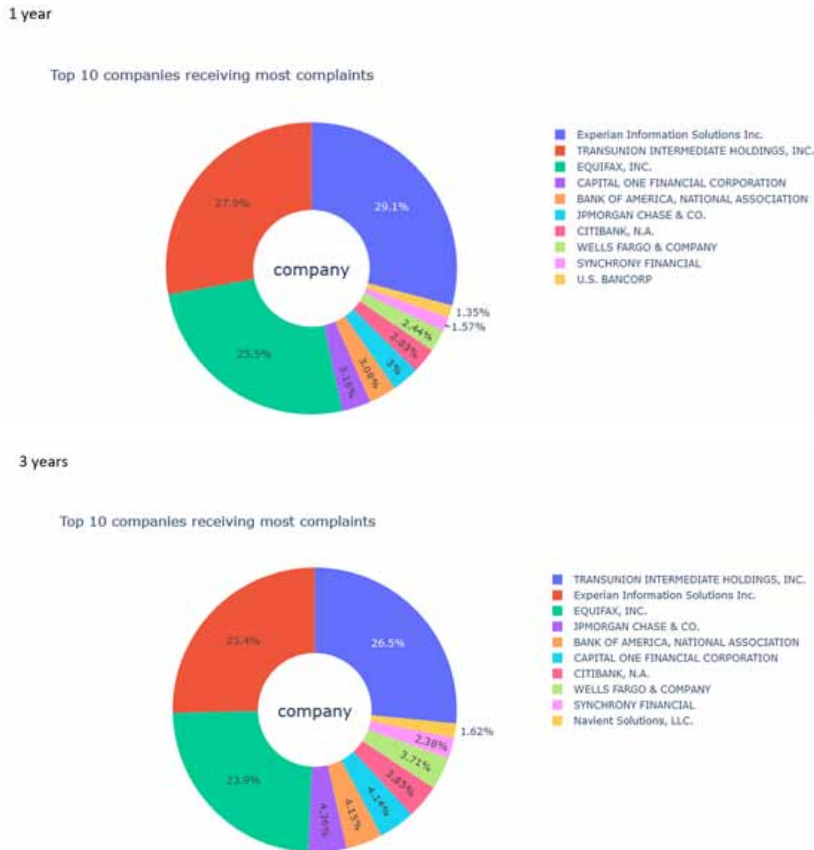
6 months

Top 10 companies receiving most complaints



continued on following page

Box 3. Continued



important implications for businesses looking to improve their products and services. Eighthly, the study demonstrates the usefulness of machine learning algorithms in classifying customer complaints automatically. This can save time and resources for businesses and help them to manage complaints more effectively. Ninthly, the study highlights the importance of prioritizing customer complaints for businesses. By doing so, businesses can gain valuable insights into the critical areas where their products or services may need improvement. Moreover, the study provides evidence that businesses can improve their complaint management strategies by utilizing machine learning techniques to identify patterns and trends in customer complaints.

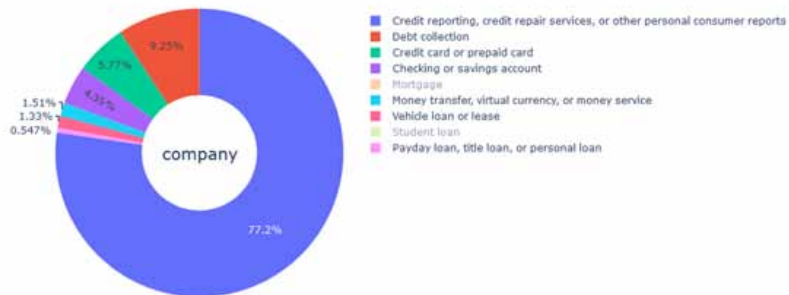
Managerial Applications

The current research has some managerial implications. Firstly, businesses can use the findings of the study to determine which machine learning techniques are most appropriate for various time period datasets. This can help them to improve their complaint management strategies and prioritize customer complaints more effectively. Secondly, the study's findings on the most accurate machine learning algorithms for different time periods can be used by businesses to enhance their complaint management strategies. Thirdly, businesses can use machine learning algorithms to identify patterns and trends in customer complaints, allowing them to proactively address issues and improve their products or services. Fourthly, businesses can analyze customer complaints across various dimensions such as

Box 4. Products receiving top complaints

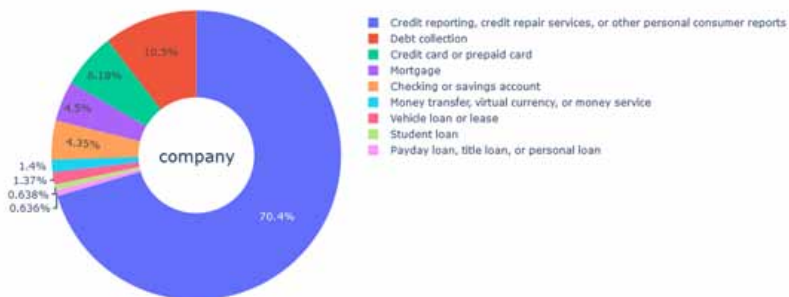
1 month

Top 10 products receiving most complaints



3 months

Top 10 products receiving most complaints



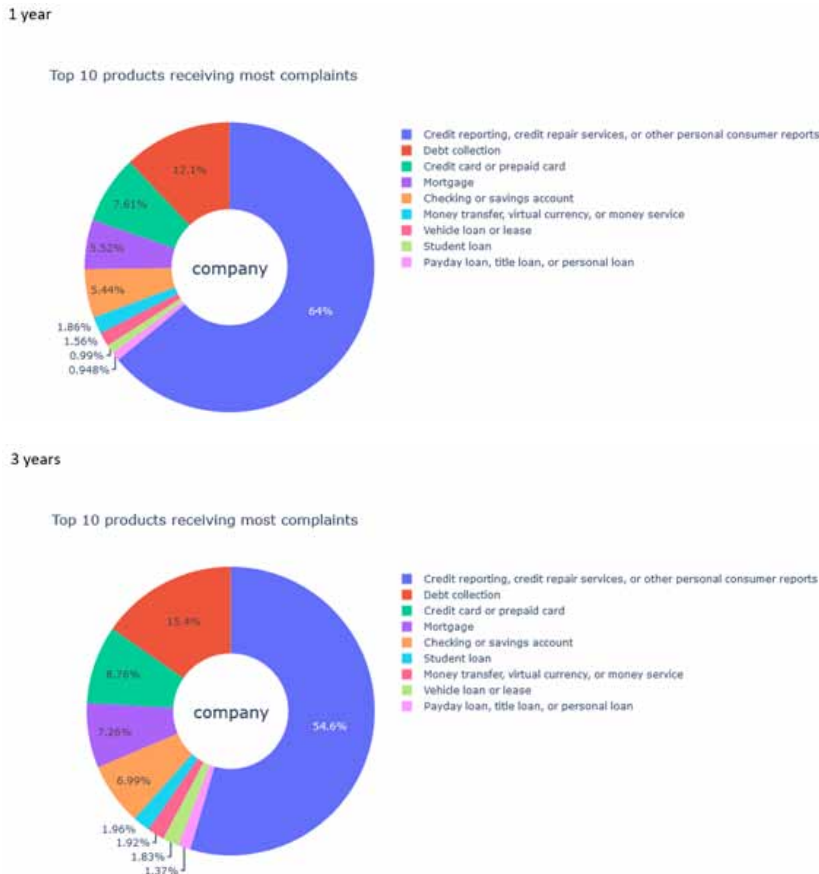
6 months

Top 10 products receiving most complaints



continued on following page

Box 4. Continued



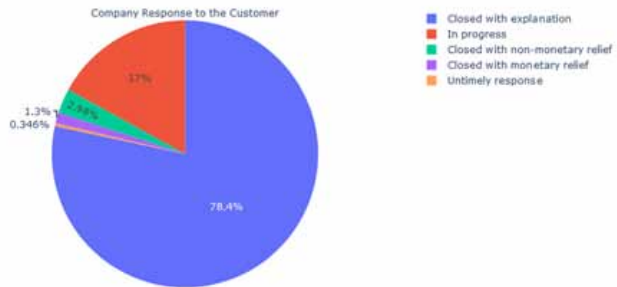
product type, location, and time period, allowing them to identify specific areas that need improvement. Fifthly, businesses can use machine learning algorithms to visualize customer complaints statistics for huge datasets, which can help them to manage and prioritize customer complaints more effectively. Sixthly, businesses can prioritize customer complaints to gain valuable insights into the critical areas where their products or services may need improvement. This can result in higher customer satisfaction and loyalty, which is crucial for business success. Seventhly, businesses can improve their complaint management strategies by utilizing machine learning techniques to identify patterns and trends in customer complaints. Eighthly, the study's findings can be used to develop more effective complaint management systems for financial organizations, which can result in higher customer satisfaction and loyalty. Ninthly, businesses can use machine learning algorithms to classify customer complaints automatically, saving time and resources and helping them to manage complaints more effectively. Moreover, businesses can use the study's findings to develop a complaints decision support tool that utilizes machine learning models to accurately forecast consumer complaints over time.

Limitations and Future Research Directions

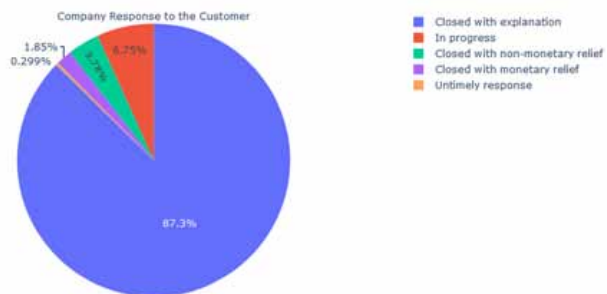
The current study's shortcomings also provide possibilities for future investigation. The current research on using machine learning algorithms to predict and manage financial consumer complaints has some limitations that provide directions for future research. One limitation

Box 5. Company response to complaints

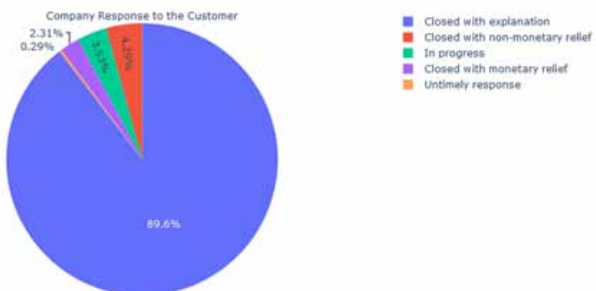
1 month



3 months



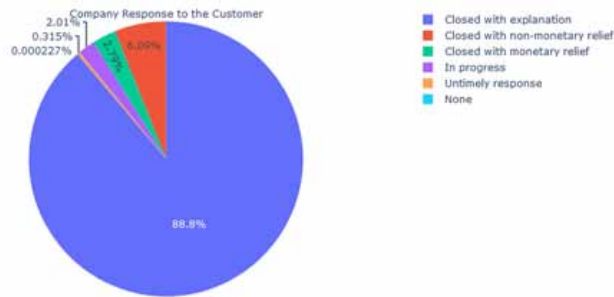
6 months



continued on following page

Box 5. Continued

1 year



3 years

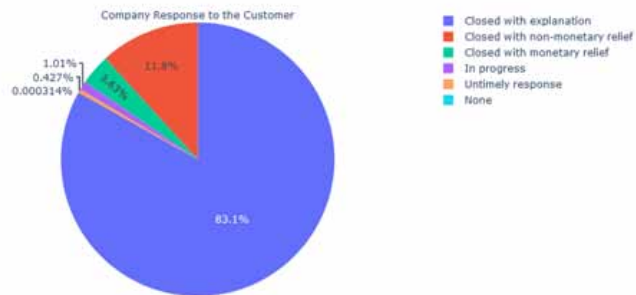


Table 2. performance of models

Dataset	Accuracy		Precision		Recall		f1_score (weighted)		mean_squared_error		RMSE		r2_score		cohen_kappa_score		matthews_corrcoef		mean_squared_log_error	
	SVM	Logistic regression	SVM	Logistic Regression	SVM	Logistic Regression	SVM	Logistic Regression	SVM	Logistic regression	SVM	Logistic regression	SVM	Logistic regression	SVM	Logistic regression	SVM	Logistic regression	SVM	Logistic regression
1 month	90.402	85.308	0.899	0.818	0.904	0.853	0.898	0.830	0.789	1.026	0.888	1.013	0.393	0.299	0.771	0.633	0.773	0.648	0.061	0.079
3 months	86.368	86.102	0.860	0.857	0.864	0.861	0.860	0.853	1.062	1.078	1.031	1.038	0.480	0.477	0.770	0.758	0.771	0.760	0.082	0.077
6 months	85.931	85.76	0.856	0.852	0.859	0.857	0.856	0.852	1.133	1.161	1.064	1.081	0.489	0.465	0.774	0.766	0.775	0.767	0.089	0.091
1 year	85.701	85.68	0.854	0.854	0.857	0.857	0.854	0.854	1.153	1.130	1.073	1.063	0.502	0.509	0.776	0.774	0.777	0.775	0.088	0.087
3 years	84.317	84.242	0.854	0.840	0.857	0.842	0.841	0.840	1.208	1.236	1.099	1.111	0.550	0.545	0.774	0.773	0.774	0.773	0.085	0.088

Table 3. Company response to consumer

Dataset	company_response_to_consumer	Product								
		Checking or savings account	Credit card or prepaid card	Credit reporting, credit repair services, or other personal consumer reports	Debt collection	Money transfer, virtual currency, or money service	Mortgage	Payday loan, title loan, or personal loan	Student loan	Vehicle loan or lease
1 Month	Closed with explanation	998	1309	27841	3210	436	1199	162	177	406
	Closed with monetary relief	175	239	16	18	64	57	11	3	11
	Closed with non-monetary relief	58	239	626	315	35	36	8	14	28
	In progress	658	711	5099	425	114	519	49	59	129
	Untimely response	5	15	41	57	8	15	8	3	6
3 Months	Closed with explanation	3925	5301	84484	11691	1374	4776	683	665	1453
	Closed with monetary relief	694	991	131	64	188	252	31	15	51
	Closed with non-monetary relief	190	798	2253	1193	114	173	43	66	121
	In progress	882	970	5287	607	146	656	63	78	155
	Untimely response	7	27	102	167	12	38	13	11	15
6 Months	Closed with explanation	9463	11747	158393	24430	3614	10778	1635	1418	3190
	Closed with monetary relief	1674	2473	285	155	415	575	83	23	121
	Closed with non-monetary relief	420	1736	4869	2546	270	390	93	138	289
	In progress	883	970	5290	610	146	663	63	78	155
	Untimely response	11	29	187	339	17	56	37	22	29
1 Year	Closed with explanation	18398	23762	260304	46894	6794	21815	3640	3708	5926
	Closed with monetary relief	3759	5390	568	318	743	1003	186	70	234
	Closed with non-monetary relief	912	3343	15383	4869	471	736	202	394	519
	In progress	883	970	5294	610	147	663	63	78	155
	None	0	0	1	0	0	0	0	0	0
	Untimely response	16	46	327	622	34	93	85	110	53
3 Years	Closed with explanation	16913	20118	144247	42864	5110	21304	3870	5500	5068
	Closed with monetary relief	4171	4688	489	327	568	797	172	105	242
	Closed with non-monetary relief	848	2715	26901	4932	351	738	212	476	415
	In progress	318	365	1976	204	44	221	20	27	56
	None	0	0	1	0	0	0	0	0	0
	Untimely response	19	40	263	642	41	81	92	129	55

Table 4. provided consumer consent

Dataset	consumer_consent_provided?	Product								
		Checking or savings account	Credit card or prepaid card	Credit reporting, credit repair services, or other personal consumer reports	Debt collection	Money transfer, virtual currency, or money service	Mortgage	Payday loan, title loan, or personal loan	Student loan	Vehicle loan or lease
1 Month	Consent not provided	95	190	13773	564	12	146	9	5	29
	Consent provided	227	427	4952	596	65	355	23	19	88
	Consent withdrawn	4	4	43	11	3	7	1	0	1
	None	754	1199	13050	2425	458	807	146	185	310
	Other	24	48	1129	87	12	51	7	4	10
3 Months	Consent not provided	592	1146	46526	3448	188	830	110	108	230
	Consent provided	1484	2631	18169	4076	562	2082	267	231	633
	Consent withdrawn	13	22	105	32	9	23	3	1	3
	None	1123	2004	19049	4050	695	1273	243	322	509
	Other	61	99	5144	307	34	165	17	14	30
6 Months	Consent not provided	1783	3121	88285	9156	849	2146	324	375	666
	Consent provided	4133	7076	44693	10959	2082	5466	871	681	1715
	Consent withdrawn	32	61	196	75	28	44	4	5	8
	None	1124	2006	19056	4058	696	1280	245	322	511
	Other	141	176	9940	632	67	358	39	31	67
1 Year	Consent not provided	3478	6699	143292	18889	1762	4308	770	1138	1366
	Consent provided	8604	15510	89355	23348	4248	11358	2162	2228	3407
	Consent withdrawn	77	130	414	146	45	89	12	19	19
	None	1124	2009	19062	4061	697	1284	246	322	514
	Other	264	326	17122	994	139	735	74	60	120
3 Years	Consent not provided	3049	5823	79255	16578	1277	4306	830	1707	1146
	Consent provided	7552	13204	66010	24342	3380	10254	2397	3609	3018
	Consent withdrawn	39	91	235	101	25	53	10	18	13
	None	414	749	6696	1399	225	423	86	104	185
	Other	243	338	10422	977	150	893	91	47	105

of the current study is that it only examines customer complaints data for a single financial institution, which may not be representative of the entire financial industry. Future research could replicate this study with data from multiple financial institutions to obtain a more comprehensive understanding of the effectiveness of machine learning algorithms in predicting and managing

Table 5. Timely response to consumer

Dataset	timely_response?	Product								
		Checking or savings account	Credit card or prepaid card	Credit reporting, credit repair services, or other personal consumer reports	Debt collection	Money transfer, virtual currency, or money service	Mortgage	Payday loan, title loan, or personal loan	Student loan	Vehicle loan or lease
1 Month	No	12	22	80	124	11	31	13	3	10
	Yes	1882	2491	33543	3901	646	1795	225	253	570
3 Months	No	35	69	227	428	33	99	37	12	28
	Yes	5663	8018	92030	13294	1801	5796	796	823	1767
6 Months	No	68	160	468	872	65	194	122	30	56
	Yes	12383	16795	168556	27208	4397	12268	1789	1649	3728
1 Year	No	236	597	870	1718	131	391	289	143	113
	Yes	23732	32914	281007	51595	8058	23919	3887	4217	6774
3 Years	No	354	263	827	1811	137	317	326	189	132
	Yes	21915	27663	173050	47158	5977	22824	4040	6048	5704

consumer complaints. Additionally, In addition, in a future study, a hybrid prediction model could be proposed to embed the appearances of two or more procedures to augment the performance of the model. Future research can be designed to measure the performance of the model for the full complaints dataset. Also, the future study could emphasize including diverse models such as K Neighbors Classifier, Naive Bayes, kNN, and random forest, etc. Customer sentiment analysis is critical for any company to understand the success of their product, and future research may examine consumer sentiment of complaints consumers have about their products using both supervised and unsupervised machine learning algorithms. Another limitation of this study is that it only uses one type of complaint data (i.e., textual complaints) and does not consider other sources of complaint data such as social media or phone calls. Future research could expand on this study by examining the effectiveness of machine learning algorithms in predicting and managing complaints from various sources. Additionally, this study only focuses on predicting and managing customer complaints in the financial industry. Future research could examine the effectiveness of machine learning algorithms in other industries such as healthcare or retail. Furthermore, the study only uses a limited number of machine learning algorithms. Future research could explore the effectiveness of other machine learning algorithms in predicting and managing customer complaints. Lastly, this study only considers a one- and three-month dataset. Future research could examine the effectiveness of machine learning algorithms in predicting and managing customer complaints for longer periods, such as six months or a year. In summary, the limitations of this study provide valuable directions for future research to expand and

improve upon the current findings. By addressing these limitations, researchers can enhance the understanding of the effectiveness of machine learning algorithms in predicting and managing customer complaints across various industries and over longer periods.

ACKNOWLEDGMENT

This work was supported by the Princess Nourah Bint Abdulrahman University, Riyadh, Saudi Arabia, under grant number PNURSP2023R 277.

REFERENCES

- Abdul Rahim, N. F., Ahmed, E. R., Sarkawi, M. N., Jaaffar, A. R., & Shamsuddin, J. (2019). Operational risk management and customer complaints: The role of product complexity as a moderator. *Benchmarking*, 26(8), 2486–2513. doi:10.1108/BIJ-04-2018-0089
- Ahmed, M. R., Rehana, H., & Asaduzzaman, S. (2021). Ovarian cancer substantial risk factor analysis by machine learning: A low incoming country perspective. *Biointerface Research in Applied Chemistry*, 11(1), 8457–8466. doi:10.33263/BRIAC111.84578466
- Al-Mashraie, M., Chung, S. H., & Jeon, H. W. (2020). Customer switching behavior analysis in the telecommunication industry via push-pull-mooring framework: A machine learning approach. *Computers & Industrial Engineering*, 106476, 106476. doi:10.1016/j.cie.2020.106476
- Atalik, O. (2007). Customer complaints about airline service: A preliminary study of Turkish frequent flyers. *Management Research News*, 30(6), 409–419. doi:10.1108/01409170710751908
- Ayres, I., Lingwall, J., & Steinway, S. (2013). Skeletons in the Database: An Early Analysis of the CFPB's Consumer Complaints. SSRN *Electronic Journal*. 10.2139/ssrn.2295157
- Bastani, K., Namavari, H., & Shaffer, J. (2019). Latent Dirichlet allocation (LDA) for topic modeling of the CFPB consumer complaints. *Expert Systems with Applications*, 127, 256–271. doi:10.1016/j.eswa.2019.03.001
- Bertsch, C., Hull, I., Qi, Y., & Zhang, X. (2020). Bank misconduct and online lending. *Journal of Banking & Finance*, 116, 105822. doi:10.1016/j.jbankfin.2020.105822
- Birim S, Anitsal MM, Anitsal İ. (2016). A Model Of Business Performance In The Us Airline Industry: How Customer Complaints Predict The Performance?. *Business Studies Journal*, 8(2).
- Blodgett, J.G., Granbois, D.H., & Walters, R.G. (1993). The Effects of Perceived Justice on Negative Word-of-Mouth and Repatronage Intentions. *Journal of Retailing*, 69, 399 – 428.
- Boser, B. E., Vapnik, V. N., & Guyon, I. M. (1992). Training Algorithm Margin for Optimal Classifiers. *Perception*, 144–152.
- Cambra-Fierro, J., Melero, I., & Sese, F. J. (2015). Managing Complaints to Improve Customer Profitability. *Journal of Retailing*, 91(1), 109–124. doi:10.1016/j.jretai.2014.09.004
- Chakraborty, S., Jain, S., Nandy, A., & Venture, G. (2021). Pathological Gait Detection Based on Multiple Regression Models Using Unobtrusive Sensing Technology. *Journal of Signal Processing Systems for Signal, Image, and Video Technology*, 93(1), 1–10. doi:10.1007/s11265-020-01534-1
- Chang, J., Khan, M. A., & Tsai, C.-T. S. (2011). Dining Occasions, Service Failures and Customer Complaint Behaviours: An Empirical Assessment. *International Journal of Tourism Research*, 14(6), 601–615. doi:10.1002/jtr.874
- Chugani, S., Govinda, K., & Ramasubbareddy, S. (2018). Data Analysis of Consumer Complaints in Banking Industry using Hybrid Clustering. In *2018 Second International Conference on Computing Methodologies and Communication (ICCMC)*, (pp. 74–78). IEEE. doi:10.1109/ICCMC.2018.8487638
- Day, R. L. (1980). Research perspectives on consumercomplaining behavior. In C. Lamb & P. Dunne (eds). *Theoretical Developments in Marketing*, American Marketing Association Press: Chicago.
- Day, R. L. (1984). In T. C. Kinnear (Ed.), *Modeling choices among alternative responses to dissatisfaction*. In *Advances in Consumer Research* (pp. 469–499). Association for Consumer Research.
- De Caigny, A., Coussement, K., & De Bock, K. W. (2018). A new hybrid classification algorithm for customer churn prediction based on logistic regression and decision trees. *European Journal of Operational Research*, 269(2), 760–772. doi:10.1016/j.ejor.2018.02.009
- Dey, P., Hariharan, S., & Ho, W. (2009). Innovation in healthcare services: A customer-focused approach. *International Journal of Innovation and Learning*, 6(4), 387–405. doi:10.1504/IJIL.2009.024135
- Filip, A. (2013). Complaint Management: A Customer Satisfaction Learning Process. *Procedia: Social and Behavioral Sciences*, 93, 271–275. doi:10.1016/j.sbspro.2013.09.188

- Fonseka, W. R. A., Nadeesha, D. G. M., Thakshila, P. M. C., Jeewandara, N. A., Wijesinghe, D. M., Sahabandu, R. V. D. S., & Asanka, P. P. G. D. (2016). Use of data warehousing to analyze customer complaint data of Consumer Financial Protection Bureau of United States of America. *2016 IEEE International Conference on Information and Automation for Sustainability (ICIAfS)*. IEEE. doi:10.1109/ICIAfS.2016.7946520
- Ghazzawi, A., & Alharbi, B. (2019). Analysis of Customer Complaints Data using Data Mining Techniques. In *Procedia Computer Science* (Vol. 163, pp. 62–69). Elsevier B.V. doi:10.1016/j.procs.2019.12.087
- Hayes, R. M., Jiang, F., & Pan, Y. HAYES. (2021). Voice of the Customers: Local Trust Culture and Consumer Complaints to the CFPB. *Journal of Accounting Research*, 59(3), 1077–1121. doi:10.1111/1475-679X.12364
- Hazra, A., Bera, N., & Mandal, A. (2017). Predicting Lung Cancer Survivability using SVM and Logistic Regression Algorithms. *International Journal of Computers and Applications*, 174(2), 19–24. doi:10.5120/ijca2017915325
- Hossain, M.S. & Rahman, M. F. (2022). Customer Sentiment Analysis and Prediction of Insurance Products' Reviews Using Machine Learning Approaches, *FIIB. Business Review*.
- Hossain, M. S. (2023). Behavioral Analytics of Consumer Complaints. In S. Nagaraj & K. Kumar (Eds.), *AI-Driven Intelligent Models for Business Excellence* (pp. 42–67). IGI Global. doi:10.4018/978-1-6684-4246-3.ch003
- Hossain, M. S., & Rahman, M. F. (2021). Website Quality, Perceived Flow, Trust, and Commitment: Developing a Customer Relationship Management Model. In R. C. Ho, A. Hou Hong Ng, & M. Nourallah (Eds.), *Impact of Globalization and Advanced Technologies on Online Business Models* (pp. 202–226). IGI Global., doi:10.4018/978-1-7998-7603-8.ch013
- Hossain, M. S., & Rahman, M. F. (2023). Detection of readers' emotional aspects and thumbs-up empathy reactions towards reviews of online travel agency apps, *Journal of Hospitality and Tourism Insights*. 10.1108/JHTI-10-2022-0487
- Hossain, M. S., Rahman, M. F., Uddin, M. K., & Hossain, M. K. (2022, June 01). Rahman, M. F., Uddin, M. K. & Hossain, M. K. (2022). Customer sentiment analysis and prediction of halal restaurants using machine learning approaches. *Journal of Islamic Marketing*. Advance online publication. doi:10.1108/JIMA-04-2021-0125
- Hsiao, Y.-H., Chen, L.-F., Choy, Y. L., & Su, C.-T. (2016). A novel framework for customer complaint management. *Service Industries Journal*, 36(13-14), 675–698. doi:10.1080/02642069.2016.1272592
- Huang, W., Mitchell, J., Dibner, C., Ruttenberg, A., & Tripp, A. (2018). How Customer Service Can Turn Angry Customers into Loyal Ones. *Harvard Business Review*, (January), 1–8. <https://hbr.org/2018/01/how-customer-service-can-turn-angry-customers-into-loyal-ones>
- Ing, E., Su, W., Schonlau, M., & Torun, N. (2018). SVMs and logistic regression to predict temporal artery biopsy outcomes. *Canadian Journal of Ophthalmology*. doi:10.1016/j.jcjo.2018.05.006 PMID:30851764
- Johnston, R. (2001). Linking complaint management to profit. *International Journal of Service Industry Management*, 12(1), 60–69. doi:10.1108/09564230110382772
- Jugwanth, B., & Vigar-Ellis, D. (2013). Customer complaint behaviour and companies' recovery initiatives: The case of the hello peter website [J]. *Ellis*, 20, 143–165.
- Landon, E. L. (1980). The Direction of Consumer Complaint Research. *Advances in Consumer Research. Association for Consumer Research (U. S.)*, 7, 335–338.
- Majumder, S., Aich, A., & Das, S. (2021). Sentiment Analysis of People During Lockdown Period of COVID-19 Using SVM and Logistic Regression Analysis. *SSRN Electronic Journal*. 10.2139/ssrn.3801039
- Pochet, N. L. M. M., & Suykens, J. A. K. (2006, June). Support vector machines versus logistic regression: Improving prospective performance in clinical decision-making. *Ultrasound in Obstetrics & Gynecology*, 27(6), 607–608. Advance online publication. doi:10.1002/uog.2791 PMID:16715467
- Rahman, M.F. & Hossain, M. S. (2022). The impact of website quality on online compulsive buying behavior: Evidence from online shopping organizations, *South Asian Journal of Marketing*. 10.1108/SAJM-03-2021-0038
- Reichheld, F., & Sasser, W. (1990). Zero defects: Quality comes to services. *Harvard Business Review*, 68(3), 105–111. PMID:10107082

Reichheld, F. F., & Sasser, E. W. Jr. (1990). Zero Defects: Quality Comes to Services. *Harvard Business Review*, 68, 105–111. PMID:10107082

Shooshtari, N. H., Stan, S., & Clouse, S. F. (2018). Receiving, Recording, and Responding to Customer Complaints: The Effects of Formalizing Customer Complaint Handling Policies in Retail Firms. *Services Marketing Quarterly*, 39(3), 225–239. doi:10.1080/15332969.2018.1472982

Sitko-Lutek, A., Chuancharoen, S., Sukpitikul, A., & Phusavat, K. (2010). Applying social network analysis on customer complaint handling. *Industrial Management & Data Systems*, 110(9), 1402–1419. doi:10.1108/02635571011087455

Strauss, J., & Hill, D. J. (2001). Consumer complaints by e-mail: An exploratory investigation of corporate responses and customer reactions. *Journal of Interactive Marketing*, 15(1), 63–73. doi:10.1002/1520-6653(200124)15:1<63::AID-DIR1004>3.0.CO;2-C

Xu, J., Zhao, J., Zhao, N., Xue, C., Fan, L., Qi, Z., & Wei, Q. (2018). The Research and Construction of Complaint Orders Classification Corpus in Mobile Customer Service. In Lecture Notes in Computer Science (including subseries Lecture Notes in Artificial Intelligence and Lecture Notes in Bioinformatics) (Vol. 11109 LNAI, pp. 351–361). Springer Verlag. doi:10.1007/978-3-319-99501-4_31

Yang, Y., Xu, D.-L., Yang, J.-B., & Chen, Y.-W. (2018). An evidential reasoning-based decision support system for handling customer complaints in mobile telecommunications. *Knowledge-Based Systems*. t 10.1016/j.knosys.2018.09.029

Ghadah Alarifi is an assistant professor in the Department of Business Administration at Princess Nourah Bint Abdulrahman University in Riyadh, Saudi Arabia. Dr. Alarifi's research focuses on entrepreneurship, innovation, social impact, and sustainability, and she has written many articles and book chapters about these topics. Dr. Alarifi is regarded as an accomplished academic and researcher who is dedicated to excellence in her teaching and research. Her work has made a significant contribution to the fields of business and entrepreneurship, particularly in Saudi Arabia.

Mst Farjana Rahman completed her MBA from the Department of Marketing, Hajee Mohammad Danesh Science and Technology University (HSTU), Bangladesh in 2018. Her research interests include online business, e-marketing, online shopping, CRM, and machine learning in marketing.

Md. Shamim Hossain is an accomplished academic and researcher in the field of marketing, with a research focus on the applications of machine learning in marketing, operations management, online business, e-marketing, self-service technologies (SSTs), e-commerce, m-banking, online customer behavior, and other areas of business. He received his Ph.D. in Business Management from the University of International Business and Economics (UIBE) in Beijing, China, and is currently an Associate Professor in the Department of Marketing at Hajee Mohammad Danesh Science and Technology University (HSTU) in Bangladesh. Dr. Hossain has published numerous research papers in prestigious journals, including the *Journal of Retailing and Consumer Services*, the *Journal of Food Quality*, the *Discrete Dynamics in Nature and Society*, the *Mathematical Problems in Engineering*, the *Journal of Healthcare Engineering*, the *Journal of Sport Psychology*, the *Journal of Hospitality and Tourism Insights*, the *Journal of Islamic Marketing*, the *FIIB Business Review*, the *Economic Research-Ekonomska Istraivanja*, and the *International Journal of Engineering Business Management*. He has also contributed book chapters to several high-quality textbooks published by IGI-Global in the United States. In addition to his research work, Dr. Hossain is a dedicated teacher and has taught a variety of courses in marketing at HSTU. He is highly regarded for his engaging teaching style and his ability to make complex marketing concepts accessible and interesting. Dr. Hossain is a pioneer in the field of machine learning-based marketing and is actively engaged in research related to this area. He can be reached at shamimuibe@yahoo.com for any further inquiries or collaborations related to his research or teaching.