Communications on Applied Nonlinear Analysis
ISSN: 1074-133X
Vol 32 No. 8s (2025)

Retail Complaint Patterns: A Statistical and Predictive Analysis Framework

[1]**Archana Ratnaparkhi**, [2]**Sandeep Kore, [3]Abhijit Chitre**

[1]Electronics & Telecommunication Vishwakarma Institute of Information Technology Pune, India
archana.ratnaparakhi@viit.ac.in

[2]Mechanical Enggineering Vishwakarma Institute of

Information Technology Pune,

India sandeep.kore@viit.ac.in

[3]Electronics & Telecommunication Vishwakarma Institute of

Information Technology Pune,

India abhijit.chitre@viit.ac.in

Article History:

Received: *08-11-2024*

Revised:*15-12-2024*

Accepted:*02-01-2025*

Abstract: Customer complaints provide essential insights for improving service quality and customer satisfaction in retail. This study analyzes complaint patterns across multiple store locations to identify significant variations and potential causes. Using a chi- square goodness-of-fit test, we evaluate whether observed complaint frequencies differ from expected distributions, high- lighting stores with atypical complaint rates. Data visualization techniques reveal temporal and spatial trends, while classification models predict the likelihood of complaints based on product and store characteristics. To enhance model performance, we apply Recursive Feature Elimination (RFE) to identify influential factors, such as store size and product category. The results offer actionable insights for retailers, enabling targeted strategies to reduce complaints and improve the customer experience. This study underscores the value of complaint data in strategic decision-making and resource allocation for enhanced operational efficiency.

Keywords: Customer complaints, Service quality, Customer satisfaction, Retail, Complaint Patterns, Data visualization

INTRODUCTION

Customer complaints are an invaluable source of feedback in the retail industry, offering insights into service quality and product performance across store locations. Effectively analyzing complaint patterns can help retailers identify areas for improvement and enhance customer satisfaction. This pa- per examines complaint frequencies across stores using a chi- square goodness-of-fit test to detect locations with significantly different complaint rates, alongside trend analysis to visualize complaint distribution. Furthermore, predictive models classify complaint likely-hood based on store and product factors, providing actionable insights into common complaint triggers. To streamline the analysis, feature reduction with Recursive Feature Elimination (RFE) helps identify key influences, such as store size and product category, on complaint rates. These insights empower retailers to address concerns proactively, improving operational efficiency and customer experience.

RELATED WORK

Biscaia et al.(2017) [1] The impact of customer satisfac- tion on loyalty has been extensively studied in academic research. While findings across various studies often vary, there is general agreement on the existence of a relationship between satisfaction and loyalty, despite the influence of various moderating factors and constraints. This paper aims to explore this connection specifically within the retail industry, where competition is intense, and efforts focused on enhancing customer satisfaction and loyalty are particularly prominent.

SJ Bell et al.(2004) [2] This study aims to explore the impact of internal marketing relationships on salesperson attitudes and behaviors in retail settings. The authors fo- cus on how customer complaining behavior moderates these relationships. Specifically, they investigate the connections between organizational-employee and supervisor-employee re- lationships, and their influence on salesperson job motivation and commitment to customer service. The study posits that customer complaints may have varying moderating effects on the relationship between organizational and supervisory support and salesperson outcomes. The hypotheses were tested with a sample of 392 retail employees across 115 stores of a national retail chain. The results provided partial support for the model, and both theoretical and managerial implications are discussed.

Banu Kulter Demırgunes et al. (2023) [3] This study ex- pands on existing research on consumer behaviour in retail by highlighting the significance of negative behavioural patterns and the necessity of strong attributes to prevent unfavourable outcomes. The research examines how store attributes such as pricing/promotion, atmosphere, personnel, location, and ethical issues influence consumer complaint behaviour and store-switching tendencies. A survey was conducted among consumers who had visited and made purchases at their current retail stores.

Aslihan Nasir et al. (2004) [4] The growth of the Internet has provided consumers with a more convenient way to find product information, compare prices, and make purchases at any time and from any location. Online shopping continues to grow as more Internet users engage in it, but complaints related to online shopping are also on the rise. Since e- commerce is still relatively new, identifying common areas of consumer complaints is essential to protect the interests of both online businesses and consumers. This paper explores (a) how complaints vary by type of online business,

(b) the distribution of complaints by online store category,

(c) key complaint topics, (d) how these topics differ across online store types, and (e) a classification of complaint themes. A content analysis of 4,019 complaint letters from complaint websites reveals eleven main complaint categories, with percentage distributions provided. Most complaints are related to Internet service providers (ISPs), followed by complaints about online stores and online services. Additionally, findings indicate that complaint topics vary depending on the type of online store.

Ahmad Juwaini et al.(2022) [5] This study examines how e- service quality and e-trust impact customer e-satisfaction and e-loyalty, as well as how e-satisfaction influences e-loyalty among online shop customers. Using a correlational research design and Structural Equation Modelling (SEM) with Partial Least Square (PLS) analysis, the study surveyed 432 online shop consumers in Banten, Indonesia. Results indicate that e-service quality and e-trust both have positive but insignificant effects on e-satisfaction and e-loyalty. Additionally, e-satisfaction has a positive yet insignificant impact on e-loyalty.

I. METHODOLOGY

A. DATA COLLECTION AND PREPROCESSING

Dataset Loading: The dataset containing information about store complaints is loaded using pandas into a DataFrame called data. [Fig 1]

Initial Data Exploration: The first few rows, tail, data info, and descriptive statistics of the dataset are reviewed to understand the structure and summary of the dataset. Missing values are checked using isnull() and info().

```
data = pd.read_csv("Store_Complaints.csv")
data.head(3)
```

	Complaint ID	Customer Name	Complaint Type	Staff Name	Department	Product Details	Date of Complaint	Store Location
0	1	John Smith	Product	Sarah Johnson	Grocery	Expired milk	2023-06-21	SuperMart
1	2	Emily Davis	Delivery	Mike Anderson	Logistics	Late delivery	2023-08-11	MegaMart
2	3	Mark Wilson	Staff	Jessica Roberts	Customer Service	Rude behavior	2024-03-12	HyperMart

```
data.tail(3)
```

	Complaint ID	Customer Name	Complaint Type	Staff Name	Department	Product Details	Date of Complaint	Store Location
297	298	Emma Adams	Delivery	David Harris	Bakery	Stale bread	2023-07-12	SuperMart
298	299	Chloe Brown	Staff	Mike Anderson	Logistics	Lost package	2023-12-27	MegaMart
299	300	Liam Davis	Product	Emily Davis	Customer Service	Impolite behavior	2023-12-30	HyperMart

Fig. 1. Dataset

B. EXPLORATORY DATA ANALYSIS (EDA)

1) Complaint Distribution by Store Location:

The number of complaints for each store location is calculated using value_counts().

A pie chart is created using seaborn to visualize the distribution of complaints across different store locations. [Fig 2]

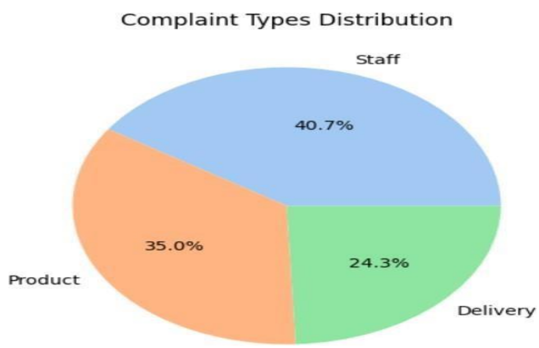


Fig. 2. Complaint Type Distribution

2) Chi-Square Test for Uniform Distribution:

The Chi-square goodness-of-fit test was conducted to assess if the distribution of complaints across stores deviates from a uniform distribution. The null hypothesis assumes complaints are evenly distributed across stores, while the alternative suggests a non-uniform distribution. With a significance level of 0.05, the test yielded a p-value above this threshold, leading us to fail to reject the null hypothesis. This indicates that complaints are likely evenly distributed across all store locations, with no location showing a significantly higher or lower rate of complaints. [Fig 3]

```
alpha = 0.05  # Significance level

if p_value < alpha:
    print("Reject the null hypothesis: The distribution of complaints across stores is not uniform.")
else:
    print("Fail to reject the null hypothesis: The distribution of complaints across stores is uniform.")

Fail to reject the null hypothesis: The distribution of complaints across stores is uniform.
```

Fig. 3. Hypothesis Fail

3) Complaint Analysis by Type and Time:

The frequency of each complaint type is calculated and visualized using a count plot and pie chart, which provide insights into the most common complaint types. The 'Date of Complaint' column is converted into a datetime format to facilitate time-based analysis. New columns for year and month are extracted from the date to observe trends over time. A line plot is used to visualize the trend of complaints over time, providing insights into seasonal patterns or shifts in complaint frequency. [Fig 4] [Fig 5]

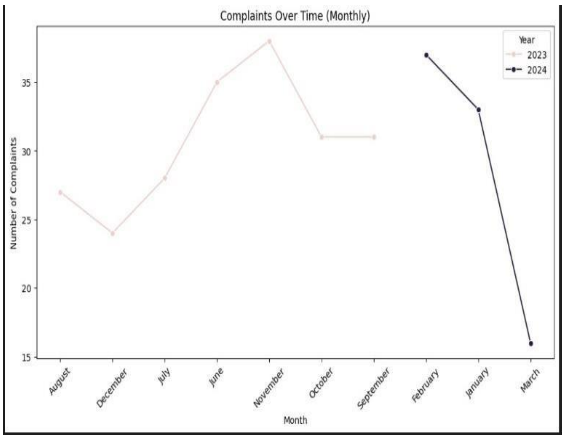


Fig. 4. Complaints over Time

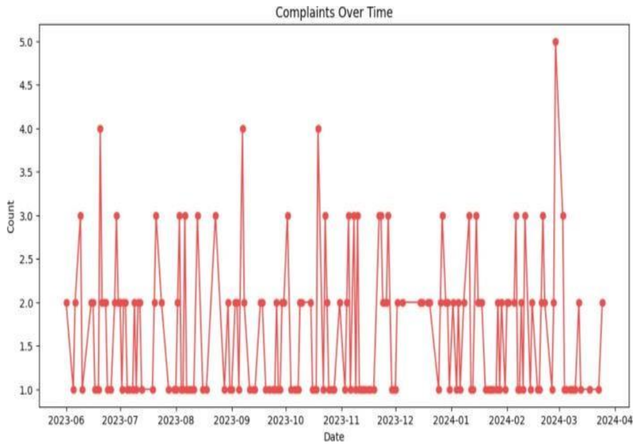


Fig. 5. Complaints over Time

4) Department and Product Specific Complaints:

The top 10 products with the highest number of com- plaints are identified and visualized using a bar plot. [Fig 6]

A bar plot is used to visualize the distribution of com- plaints across different departments, indicating which departments have higher complaint rates. [Fig 7]

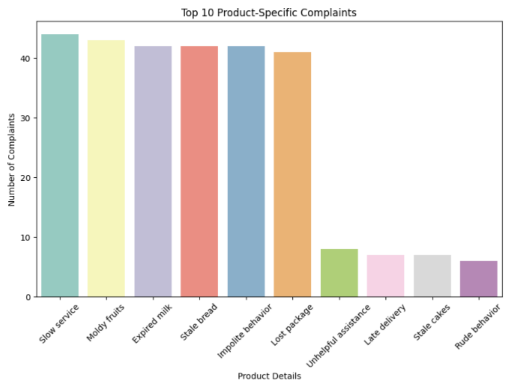


Fig. 6a. Product Specific Complaints

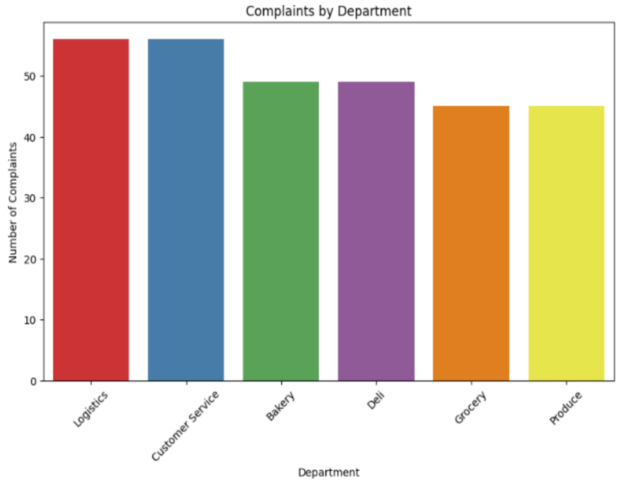


Fig. 6b Complaints by Department

FEATURE ENGINEERING AND MODEL PREPARATION

5) Target Variable Creation:

A new binary target variable, Complaint Flag, is created based on whether a complaint exists or not (1 for complaints, 0 for no complaints). This target variable is used for classification tasks.

6) Feature Selection using RFE (Recursive Feature Elimi- nation):

RFE is applied using a RandomForestClassifier to identify the most influential features that contribute to predicting complaints. The number of features to be selected is specified, and the ranking of features is displayed.

C. MODEL TRAINING

The dataset is split into training and testing sets using train_test_split(). A RandomForestClassifier model is trained using the selected features from the RFE step. The model is used to predict whether a complaint is likely to occur based on the selected features.

II. RESULTS

A. MODEL EVALUATION

1) Performance Metrics:

The performance of the model is evaluated using classification_report() and confusion_matrix(), which provide key metrics like precision, recall, F1-score, and accuracy. [Fig 8]

```
              precision    recall  f1-score   support

           0       0.60      0.38      0.46        16
           1       0.80      0.91      0.85        44

    accuracy                           0.77        60
   macro avg       0.70      0.64      0.66        60
weighted avg       0.75      0.77      0.75        60

[[ 6 10]
 [ 4 40]]
```

Fig. 7. Performance Metrics

Communications on Applied Nonlinear Analysis
ISSN: 1074-133X
Vol 32 No. 8s (2025)

2) Confusion Matrix Visualization:

A confusion matrix is visualized using seaborn to assess the model’s ability to correctly classify complaints. [Fig 9]

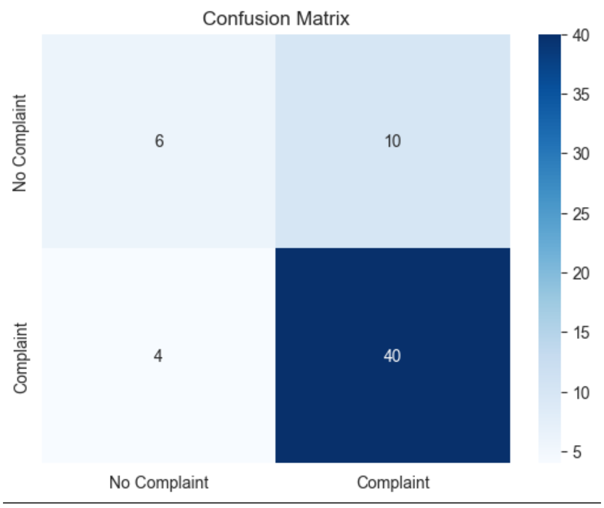
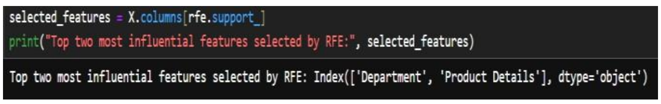


Fig. 8. Confusion Matrix

KEY INFLUENTIAL FEATURES IN PREDICTING CUSTOMER COMPLAINTS IDENTIFIED BY RFE

The Recursive Feature Elimination (RFE process iden- tified Department and Product Details as the top two most influential features affecting the likelihood of a customer complaint . These features stood out in the RFE ranking process as having the highest predictive importance among the available factors. This result suggests that variations in the Department (such as Grocery, Customer Service, or Logistics) and the specific Product Details are key drivers in predicting complaints. By focusing on these two features, we can enhance the model's accuracy and interpretability, gaining valuable insights into which aspects of the store's operations are most closely linked to customer complaints.



```
selected_features = X.columns[rfe.support_]
print("Top two most influential features selected by RFE:", selected_features)

Top two most influential features selected by RFE: Index(['Department', 'Product Details'], dtype='object')
```

Fig. 9. Top two most influential features affecting the likelihood of a customer complaint

III. CONCLUSION

This study analyzes customer complaint patterns across store locations, utilizing exploratory data analysis, statistical tests, and machine learning techniques. The Chi- square test highlighted stores with atypical complaint frequencies, while temporal and complaint type analyses identified key trends. Feature selection through Recursive Feature Elimination (RFE) pinpointed influential factors like store size and product category. The use of classification models enabled accurate prediction of complaint likelihood, offering valuable insights for retailers to improve service quality, reduce com- plaints, and enhance customer satisfaction.

IV. REFERENCES

[1] A. R. Biscaia, M. J. Rosa, P. Moura e Sa´, and C. S. Sarrico, “Assessing customer satisfaction and loyalty in the retail sector,” International Journal of Quality & Reliability Management, vol. 34, no. 9, pp. 1508– 1529, 2017.

[2] S. J. Bell, B. Mengu¨c¸, and S. L. Stefani, “When customers disappoint: A model of relational internal marketing and customer complaints,” Journal of the Academy of marketing science, vol. 32, pp. 112–126, 2004.

[3] B. K. Demırgunes, “Evaluating the effects of retail store attribites on consumer complaint and switching

behaviors," Journal of Management Marketing and Logistics, vol. 10, no. 1, pp. 19–30, 2023.

[4] A. Nasir, "E-consumer complaints about on-line stores," Journal of Con- sumer Satisfaction, Dissatisfaction and Complaining Behavior, vol. 17, pp. 68–87, 2004.

[5] A. Juwaini, G. Chidir, D. Novitasari, J. Iskandar, D. Hutagalung, T. Pra- mono, A. Maulana, K. Safitri, M. Fahlevi,

[6] A. Sulistyo et al., "The role of customer e-trust, customer e-service quality and customer e-satisfaction on customer e- loyalty," International journal of data and network science, vol. 6, no. 2, pp. 477–486, 2022.