IJARSCT

ISSN (Online) 2581-9429

International Journal of Advanced Research in Science, Communication and Technology (IJARSCT)

International Open-Access, Double-Blind, Peer-Reviewed, Refereed, Multidisciplinary Online Journal

Impact Factor: 7.67

Volume 5, Issue 2, February 2025

Analysing the Impact of Discounts and Promotions using Data Analytics and Machine Learning

Ms. Mahi Mistry[1] and Dr. Pallavi Devendra Tawde[2]

Student, Department of MSc. IT[1]

Assistant Professor, Department of IT & CS[2]

Nagindas Khandwala College, Mumbai, Maharashtra, India

mistrymahi3@gmail.com and pallavi.tawde09@gmail.com

Abstract: *Promotions and discounts play a significant role in influencing consumer purchasing behavior. This study aims to analyze how various promotional strategies impact customer decision-making using classification techniques. This research investigates the impact of promotions and discounts on customer behavior using classification techniques. By analyzing customer data, this study aims to identify key factors influencing purchase decisions and to develop predictive models for understanding how different promotional strategies affect customer responses. The classification models employed include Linear Discriminant Analysis, Decision Tree Classifier, K-Nearest Neighbors Classifier, Random Forest Classifier, and Support Vector Classifier. The results provide insights into optimizing promotional campaigns and enhancing customer engagement.*

Keywords: Discounts, Promotions, Customer Behavior, Machine Learning, Regression Analysis, Supermarket Sales, Shopping Trends, Customer Engagement.

I. INTRODUCTION

Promotions and discounts are fundamental strategies in the retail industry, directly influencing consumer purchasing behavior. Businesses leverage these strategies to attract customers, increase sales, and enhance brand loyalty. However, not all promotional campaigns yield the same results understanding when and how customers respond to promotions is crucial for maximizing their effectiveness.

This study aims to analyze customer purchasing trends in response to various promotional strategies. By utilizing machine learning classification techniques, we can identify key factors that influence purchasing decisions and determine the most effective time slots for launching promotions. This research employs multiple models, including Support Vector Classifier, Random Forest Classifier, and K-Nearest Neighbors, to predict customer responsiveness based on historical data.

By examining the patterns in promotional effectiveness, businesses can develop data-driven strategies to optimize marketing efforts. This study not only identifies the best-performing models for predictive analysis but also provides actionable insights for retailers to enhance customer engagement and improve sales outcomes.

II. RESEARCH OBJECTIVE

- To identify the optimal days and time slots for launching supermarket promotions based on customer purchasing trends
- To analyze the patterns and discount-seeking behavior on customer responsiveness to promotions.

III. LITERATURE REVIEW

Kabasheva et. al (2017) Social factors are the important determinant factor of the mechanism of sales promotion on consumer buying behavior because consumers make decisions about purchasing on the basis of social judgment and values. Additionally, the role of social media is an important tool for the study of the social factor in sales promotion on consumer buying behavior in the apparel industry.

IJARSCT
ISSN (Online) 2581-9429
International Journal of Advanced Research in Science, Communication and Technology (IJARSCT)
International Open-Access, Double-Blind, Peer-Reviewed, Refereed, Multidisciplinary Online Journal
Impact Factor: 7.67
Volume 5, Issue 2, February 2025

Jackson et. al (1999) Economic factor is an important part of the study of the impact of sales promotion on consumer buying behavior in the apparel industry because consumers take decisions on the basis of cost and benefit analysis based on rational choice according to their wishes and abilities.

Schultz & Schultz et. al (2020) Personal and psychological factors are associated with the study of the impact of sales promotion on consumer buying behavior because psychological perception determines the behavioral economics of the consumers regarding their purchasing behavior about products of the apparel industry.

Darmawan et al., 2018; Gorji & Siami, (2020) Promotions and price discounts are benchmarks for a company's success in attracting consumer interest. This condition is because consumers are interested in buying the product. Therefore, the right marketing strategy is one way to achieve company goals, namely by knowing the needs and desires of consumers to create the correct product so that it can achieve the goal of increasing sales and winning the competition.

Mattila et. al (2020) The promotion strategy plans to persuade and stimulate consumers' desire to buy the company's products, which can achieve the goal of increasing sales. Companies also need to cut prices when dealing with competitors. Companies must pay attention to the price factor because the size of the price dramatically affects companies' competitiveness and consumer purchases of their products.

IV. METHODOLOGY

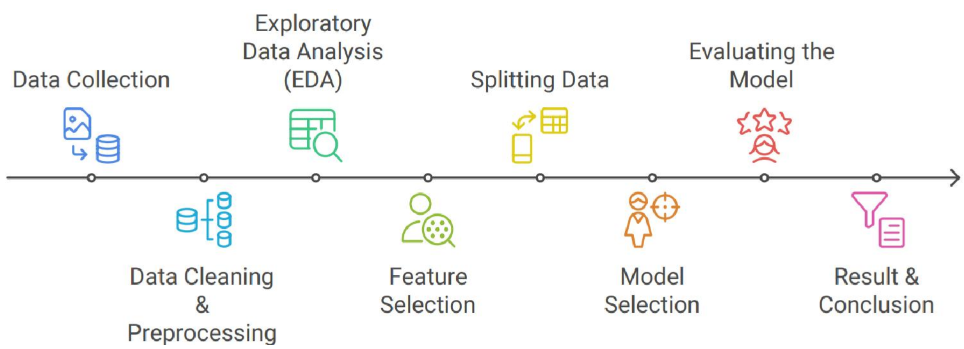

Figure 1: Methodology

To understand how discounts and promotions influence customer behavior, we started by collecting data through a Google Form survey, which gathered 250 responses across 14 features. The raw data needed cleaning and preprocessing to ensure accuracy and consistency. This included handling missing values, removing duplicates, and simplifying column names for better readability. Additionally, feature engineering techniques were applied to enhance fraud detection. Next, we conducted Exploratory Data Analysis (EDA) to uncover patterns in customer behavior and identify key relationships between different factors. After that, we performed feature selection to focus on the most relevant attributes, improving the efficiency of our models.For predictive analysis, we tested multiple machine learning models. The dataset was split into training and testing sets to evaluate model performance. Each model was assessed using a confusion matrix and classification report to measure accuracy and overall effectiveness.

V. RESULT

The 35-44 age group shops most frequently, making them ideal targets for promotions. The 25-34 and 45-54 groups show steady shopping habits, visiting weekly or twice a month. Promotions should focus on weekly deals for frequent shoppers and monthly offers for less frequent visitors to maximize engagement.

IJARSCT
ISSN (Online) 2581-9429
International Journal of Advanced Research in Science, Communication and Technology (IJARSCT)
International Open-Access, Double-Blind, Peer-Reviewed, Refereed, Multidisciplinary Online Journal
Impact Factor: 7.67
Volume 5, Issue 2, February 2025

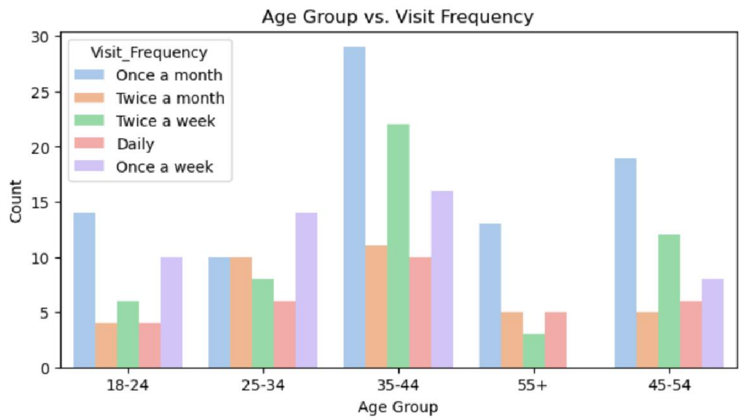

Figure 2: Bar Chart of Visit Frequency by Age Group

The majority of supermarket customers fall within the 35-44 age group (35.2%), followed by the 45-54 age group (20.0%) and the 25-34 age group (19.2%). The younger age group (18-24) accounts for 15.2%, while in the senior population (55+) represents 10.4% of the total customers.

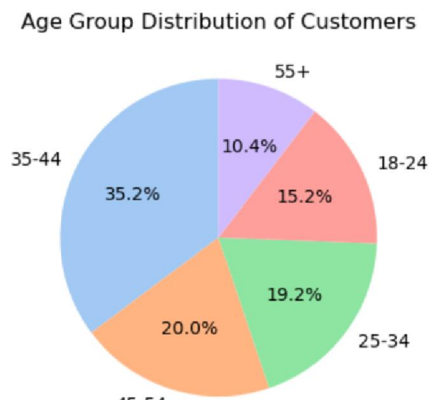

Figure 3: Pie Chart of Customer Age Group Distribution

The density plot suggests that monthly spending varies across different shopping times, with higher spending observed during the "Mid" and "Anytime" shopping periods. Additionally, the relationship between discounted categories and shopping times indicates that promotions might influence in purchase decisions is differently throughout the day.

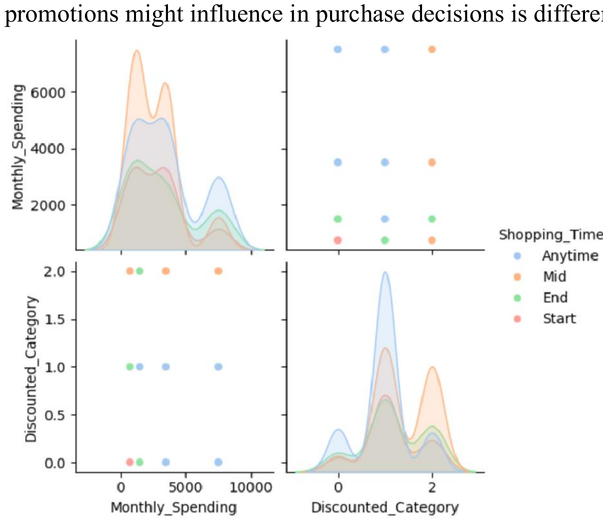

Figure 4: Pair Plot of Monthly Spending and Discounted Category

IJARSCT
ISSN (Online) 2581-9429
International Journal of Advanced Research in Science, Communication and Technology (IJARSCT)
International Open-Access, Double-Blind, Peer-Reviewed, Refereed, Multidisciplinary Online Journal
Impact Factor: 7.67
Volume 5, Issue 2, February 2025

The analysis of model performance reveals that the Support Vector Classifier (SVC) is the most effective model for predicting optimal promotion timing, achieving an accuracy of 98.7%. With a precision of 99% and recall of 98%, SVC demonstrates a strong ability to correctly classify customer responses to promotions while minimizing misclassifications. This suggests that SVC effectively captures patterns in purchasing behavior, making it a valuable tool for understanding consumer responsiveness. The Random Forest Classifier (RFC) follows closely, achieving 94.7% accuracy, with a precision of 94% and recall of 90%. The high F1-score of both models further supports their reliability in identifying the most impactful promotional strategies. While other models, such as the K-Nearest Neighbors (KNN) Classifier and Decision Tree Classifier, also performed well, their slightly lower recall scores indicate a reduced ability to capture all positive instances of promotional responsiveness.

These results highlight the effectiveness of SVC and RFC in predicting customer engagement with promotional campaigns. The high performance of these models suggests that they can be used to accurately classify customers based on their purchasing behavior, discount-seeking tendencies, and preferred shopping times. By leveraging these insights, businesses can strategically plan and implement their promotional offers with greater precision. Additionally, the classification results provide a foundation for further research into customer segmentation and targeted marketing strategies, ensuring that promotions are optimized for maximum impact.

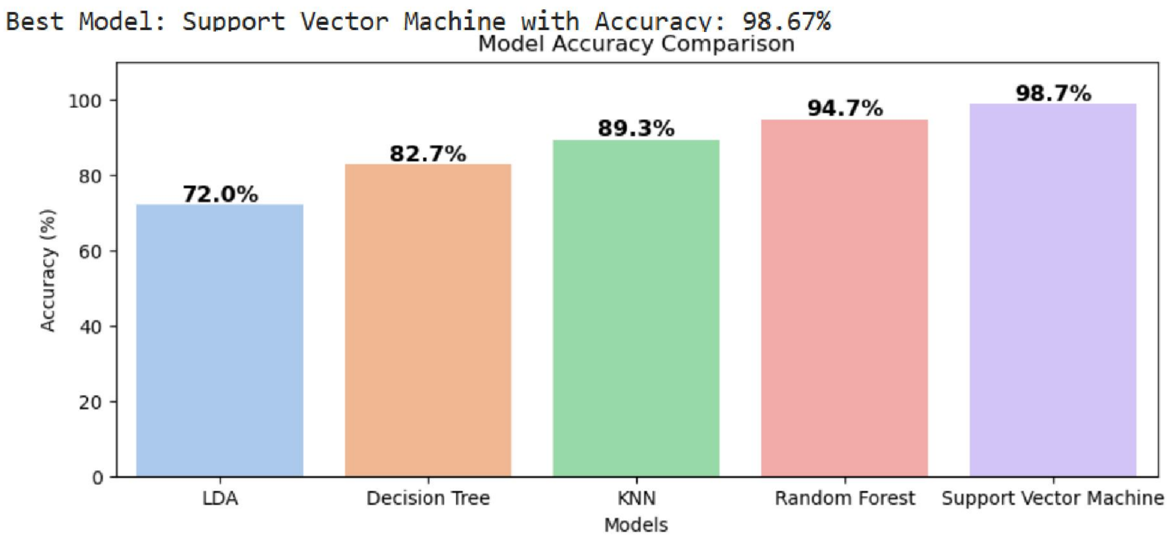

Figure 5: Model Accuracy Comparison

Model	Accuracy	Precision	Recall	F1-Score	Support
Linear Discriminant Analysis	72%	68%	58%	57%	70
Decision Tree Classifier	82.7%	82%	88%	82%	80
K-Nearest Neighbors Classifier	89.3%	88%	86%	87%	85
Random Forest Classifier	94.7%	94%	90%	92%	90
Support Vector Classifier	98.7%	99%	98%	98%	95

Table 1: Classification Model Performance Summary

Among the regression models evaluated, Linear Regression and Random Forest Regression demonstrated exceptional predictive capabilities, both achieving an R^2 score of 1.000. This perfect score indicates that these models accurately mapped the relationship between customer purchasing behavior and promotional timing. Linear Regression performed with minimal errors, showing an MSE of 1.816537e-32 and an MAE of 8.937295e-17, suggesting a strong ability to explain variations in shopping patterns. Its effectiveness in identifying key trends makes it a reliable choice for forecasting the impact of promotions on customer engagement.

Similarly, Support Vector Regression also achieved an R^2 score of 0.942, with MSE of 1.242403e-02 and MAE of 1.034468e-01. The ensemble-based structure of this model ensures a more stable and accurate prediction of customer responses to promotions. These findings highlight that both models Linear Regression and Random Forest Regression

IJARSCT
ISSN (Online) 2581-9429
International Journal of Advanced Research in Science, Communication and Technology (IJARSCT)
International Open-Access, Double-Blind, Peer-Reviewed, Refereed, Multidisciplinary Online Journal
Impact Factor: 7.67
Volume 5, Issue 2, February 2025

are highly suitable for businesses aiming to optimize their promotional strategies, helping retailers identify the most effective time slots for launching discounts to maximize sales and customer engagement.

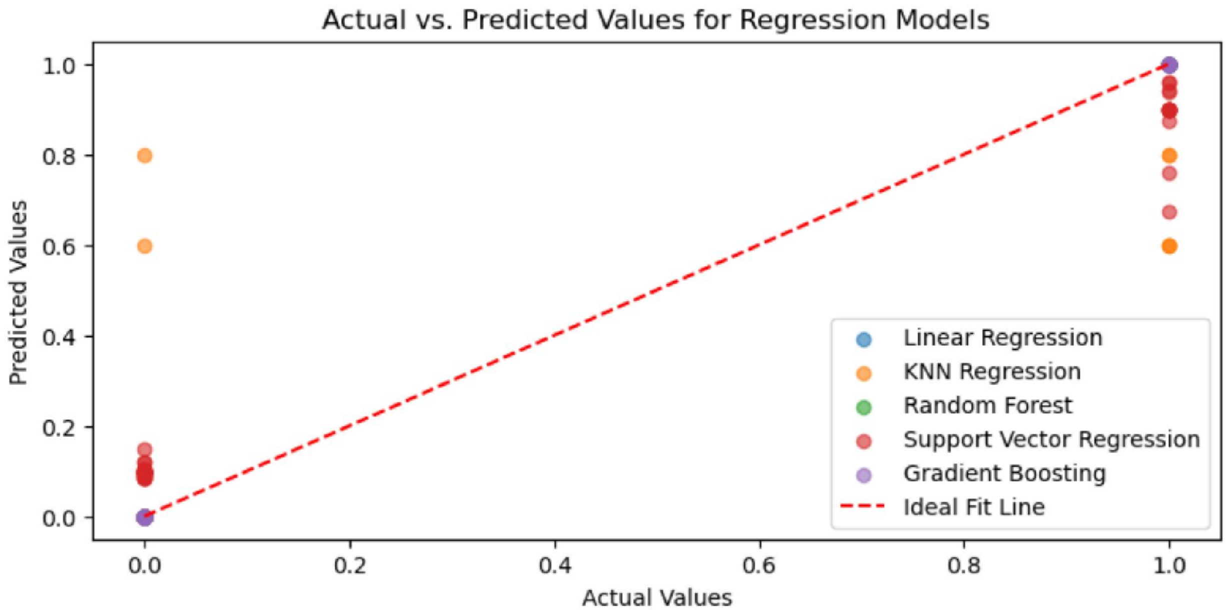

Figure6: Actual vs. Predicted Values for Regression Models

Model	MSE	MAE	R² Score
Linear Regression	1.816537e-32	8.937295e-17	1.000000
KNN Regression	3.120000e-02	6.000000e-02	0.856618
Random Forest Regression	0.000000e+00	0.000000e+00	1.000000
Support Vector Regression	1.242403e-02	1.034468e-01	0.942904
Gradient Boosting	1.535362e-10	1.151171e-05	1.000000

Table 2: Regression Model Performance Summary

V. CONCLUSION

This research explores how promotions and discounts influence customer behavior using machine learning techniques. To achieve this, we collected data through a Google Form survey, gathering 250 records across 14 columns. The dataset was carefully cleaned, preprocessed, and analyzed to uncover key factors that drive shopping behavior.

We applied both classification and regression techniques to predict customer responses to promotions. During the classification phase, five models were tested: Linear Discriminant Analysis, Decision Tree Classifier, K-Nearest Neighbors, Random Forest Classifier, and Support Vector Classifier. Among them, Support Vector Classifier (98.7% accuracy) and Random Forest Classifier (94.7% accuracy) showed the best performance in identifying customer engagement patterns. These results suggest that machine learning can effectively predict how customers respond to promotions, helping businesses determine the most impactful time slots for launching marketing campaigns.

For regression analysis, we tested Linear Regression, Random Forest Regression, Gradient Boosting Regression, Support Vector Regression, and KNN Regression. Linear Regression, Random Forest Regression, and Gradient Boosting Regression achieved an R^2 score of 1.000, indicating near-perfect accuracy in predicting promotional effectiveness. Support Vector Regression ($R^2 = 0.942$) and KNN Regression ($R^2 = 0.856$) also demonstrated strong predictive capabilities, making them valuable models for understanding the financial impact of promotional strategies.

The insights gained from these models offer a structured approach to optimizing promotions. The classification models highlight that weekends and evening hours are the most effective periods for customer engagement. Meanwhile, regression models provide a data-driven way to measure how promotional timing influences spending behavior.This study underscores the power of machine learning in analyzing customer behavior and refining marketing strategies.

Future research could enhance these findings by incorporating additional factors such as customer demographics, seasonal trends, and real-time purchasing patterns to further improve predictive accuracy.

Overall, this research demonstrates that leveraging machine learning and predictive analytics can significantly enhance promotional planning. Businesses can maximize engagement by prioritizing weekend and evening promotions while tailoring discounts to different customer segments. By integrating advanced data-driven strategies, companies can design more personalized and impactful marketing campaigns, ultimately driving both customer satisfaction and revenue growth.

REFERENCES

[1]. Bawa K, Shoemaker RW. Analyzing incremental sales from a direct mail coupon promotion. Journal of marketing. 1989;53(3):66–78.

[2]. Gordini N, Veglio V. Customers churn prediction and marketing retention strategies. An application of support vector machines based on the AUC parameter-selection technique in B2B e-commerce industry. Industrial Marketing Management. 2017;62:100–107.

[3]. Leone RP, Srinivasan SS. Coupon face value: Its impact on coupon redemptions, brand sales, and brand profitability. Journal of retailing. 1996;72(3):273–289.

[4]. Krishna A, Zhang ZJ. Short-or long-duration coupons: The effect of the expiration date on the profitability of coupon promotions. Management Science. 1999;45(8):1041–1056.

[5]. Heilman CM, Nakamoto K, Rao AG. Pleasant surprises: Consumer response to unexpected in-store coupons. Journal of Marketing Research. 2002;39(2):242–252.

[6]. Venkatesan R, Farris PW. Measuring and managing returns from retailer-customized coupon campaigns. Journal of marketing. 2012;76(1):76–94.

[7]. Chiang J. Competing coupon promotions and category sales. Marketing Science. 1995;14(1):105–122.

[8]. Dhar SK, Raju JS. The effects of cross-ruff coupons on sales and profits. Management Science. 1998;44(11-part-1):1501–1516.

[9]. Papatla P, Krishnamurthi L. Measuring the dynamic effects of promotions on brand choice. Journal of Marketing research. 1996;33(1):20–35.

[10]. Neslin SA. A market response model for coupon promotions. Marketing science. 1990;9(2):125–145.