SSRG International Journal of Electronics and Communication Engineering
ISSN: 2348-8549/ https://doi.org/10.14445/23488549/IJECE-V11I9P112

Volume 11 Issue 9, 120-134, September 2024

Original Article

Transforming E-commerce with a Novel Multifaceted Data-Decision Framework

Revelle Akshara[1], Ajay Jain[2]

[1,2]*Faculty of Engineering and Technology, Mansarovar Global University, Sehore, Madhya Pradesh.*

[1]*Corresponding Author : akshara.revelle@gmail.com*

Received: 02 July 2024 Revised: 12 August 2024 Accepted: 03 September 2024 Published: 30 September 2024

Abstract *- With the quickly evolving nature of e-commerce, numerous businesses face a range of challenges, such as constrained investment, customer dissatisfaction, delivery delays, product-market misalignment, and a lack of understanding of customer preferences and satisfaction. These issues often contribute to the failure of many enterprises. This paper proposes a Novel Multifaceted Data-Decision Framework designed to navigate these challenges by guiding businesses from data collection to actionable insights using advanced analytics. The framework integrates various essential elements: data collection and processing, descriptive analytics to discern past occurrences, diagnostic analytics to unveil causative factors, and predictive analytics to estimate what could happen in the future. Prescriptive analytics provides detailed advice on how to respond and machine learning classifiers to analyze complex datasets. The framework's effectiveness is illustrated using an e-commerce dataset, showing how businesses of all sizes can leverage analytics for informed decision-making. By adopting this Multifaceted Data-Decision Framework, e-commerce businesses, from small to large-scale, can make informed decisions using data that enhance customer fulfilment, streamline operations, and promote sustainable growth, enabling them to overcome challenges and succeed in a competitive environment.*

Keywords *- Data-decision framework, Descriptive analytics, Diagnostic analytics, E-Commerce, Machine learning classifiers, Predictive analytics, Prescriptive analytics.*

1. Introduction

With the quickly evolving nature of e-commerce, numerous businesses are increasingly challenged to satisfy the growing needs of today's customers while navigating a landscape fraught with obstacles. Central to their success is the ability to ensure customer fulfilment - a complex endeavor shaped by factors ranging from limited resources to shifting market dynamics. In response to these challenges, the imperative for many enterprises to leverage data-driven insights has never been more pressing [1,2]. Data-driven decisions help strengthen many businesses. E-commerce companies can gain insights into customer actions, changes in the market, and how well their operations are running. Informed decisions based on these insights enable businesses to anticipate and mitigate risks, capitalize on opportunities, and adapt to changing market dynamics swiftly [2]. This data-driven approach is fundamental to enhancing the resilience of many e-commerce enterprises, empowering them to navigate uncertainties and drive sustainable growth in a competitive digital environment [3-6].

To address the imperative for data-driven decisions in many e-commerce, this paper introduces a Novel Multifaceted Data-Decision Framework. This framework is meticulously designed to guide many e-commerce businesses through the process of data collection, analysis, and smart decisions. By integrating various analytical techniques as well as ensemble learning, the Data-Decision Framework equips businesses with the tools and methodologies needed to gain worthy insights from data, helping them to make informed decisions. Through the systematic application of this multifaceted framework, many e-commerce businesses can enhance customer satisfaction, optimize operational efficiency, and promote sustainable growth, thereby surmounting challenges and thriving in the competitive digital marketplace.

2. Literature Review

The data science and analytics integrated into e-commerce have been extensively studied, providing valuable insights and frameworks for enhancing business to make informed decisions.

Chen et al. (2012) illustrate how to use data-driven strategies to enhance customer fulfilment and operational outcomes using advanced analytical methods in e-commerce [3].

Waller et al. (2013) discuss the implementation of prescriptive analytics and machine learning classifiers to provide specific action recommendations and manage complex datasets effectively [4]. Gandomi et al. (2015) highlight that big data analytics helps organizations gather valuable insights from vast datasets, streamline operations, and fulfill customers [2].

Akter et al. (2016) highlight the role of descriptive and diagnostic analytics in understanding past occurrences and identifying causative factors, which is crucial for addressing customer discontent and delivery setbacks in e-commerce [5].

Kietzmann et al. (2018) explore how AI transforms advertising by enhancing customer insight through analytics and machine learning techniques [6].

Bertsimas et al. (2020) integrated machine learning with operations research to create a framework for using data to optimize decision-making processes. Their methods, tested in a real-world inventory management scenario, significantly improved decision efficacy by 88%, showcasing the potential of data-driven optimization in operational contexts [7].

Roychowdhury et al. (2020) introduce a machine learning framework for analyzing online shopping behavior, emphasizing predictive models based on user-journey data. It achieves high accuracy in predicting purchase events and categorizes customers into distinct behavioural clusters, offering insights for personalized marketing strategies in e-commerce [8].

Zhuang et al. (2021) offer insights into how analytics are significant in the changes in e-commerce, customer actions, and the integration of advanced technologies [9].

Chen et al. (2021) propose an e-commerce marketing approach that leverages big data analytics to enhance customer fulfilment and optimize operations. Also highlights how big data provides informed decisions, addresses operational challenges, and boosts the competitiveness of e-commerce platforms and enterprises [10].

Alsmadi et al. (2023) analyze existing research to explore the role of analytics in fostering revolution in e-commerce, with a focus on crises like COVID-19. They highlight interdisciplinary perspectives and suggest future research directions to enhance the application of BDA in organizational contexts, emphasizing its transformative potential across various sectors [11].

Ramkumar et al. (2023) explore the transformative value of analytics on e-commerce, highlighting its significance in refining marketing strategies and streamlining operations through real-time insights.

Their qualitative study underscores the potential for businesses to leverage data-driven approaches for personalized customer experiences and strategic decision-making, thereby fostering growth and competitiveness in the digital era [12].

Nijjer et al. (2023) explore the evolving landscape of customer analytics, emphasizing the increasing investment and innovative strategies in leveraging AI and ML for personalized marketing.

Their work outlines the diverse sources of customer data, analytical tools, and the transformative significance of enhancing customer experiences in the digital age [13].

The reviewed studies collectively underscore the pivotal role of strategies based on informed decisions and the need for a robust framework integrating various analytics in revolutionizing e-commerce businesses to succeed in a competitive environment.

3. Proposed Work

The proposed work depicted in Figure 1 entails a comprehensive analysis of an e-commerce dataset aimed at extracting actionable insights to enhance the performance of e-commerce businesses.

3.1. Key Stages

3.1.1. Data Collection and Preprocessing

These are the foundational elements of effective analytics frameworks in e-commerce, enabling the extraction of valuable insights from large datasets.

3.1.2. Descriptive and Diagnostic Analytics

With the pre-processed data, descriptive and diagnostic analytics techniques are applied to gain a detailed analysis of historical changes and behaviors in e-commerce.

3.1.3. Predictive Analytics

Building upon the gained insights, robust machine learning models are developed using predictive analytics techniques to forecast what might have happened in the future.

3.1.4. Prescriptive Analytics

Finally, leveraging the findings from the above analytics, prescriptive analytics techniques are utilized to formulate detailed advice on how to respond to changes in the future to drive growth and competitiveness.

Through this holistic approach, the proposed work aims to empower e-commerce businesses with data-driven insights and actionable recommendations derived from advanced analytics techniques.

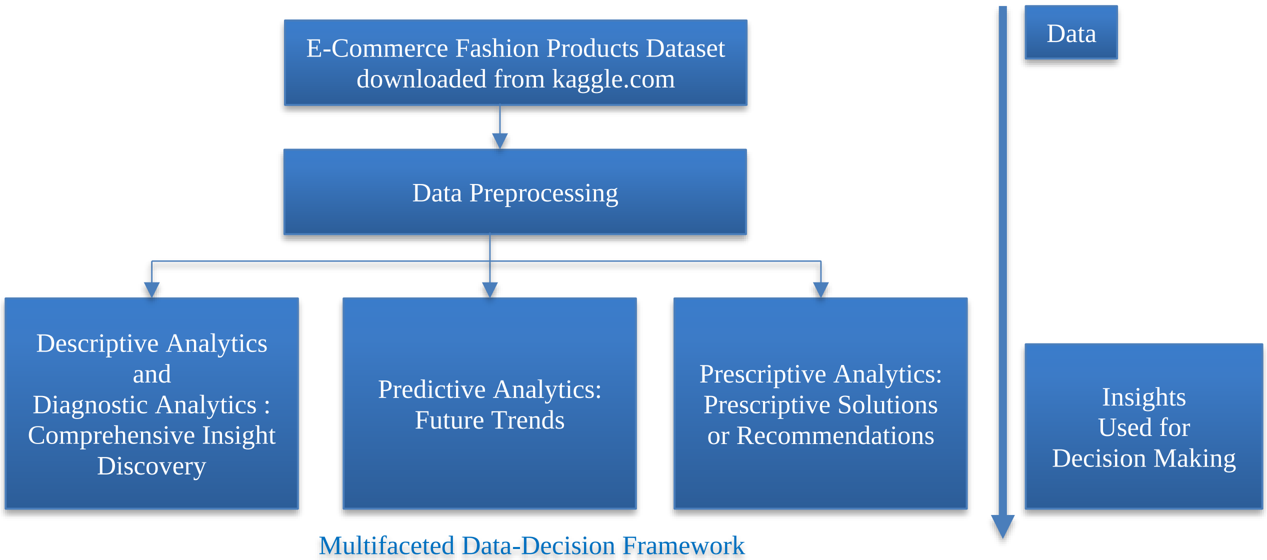

Fig. 1 Architecture of the proposed work

4. Dataset Analysis

The efficacy of the proposed framework is demonstrated using an e-commerce dataset downloaded from Kaggle.com, illustrating how businesses of all sizes can harness analytics for informed decision-making. This dataset comprises some essential attributes related to different products, as depicted in Figure 2.

In the dataset, the 'fulfilled' attribute is a class label that divides products into two groups, either fulfilled (products that have satisfied customers) or not fulfilled (products that have not satisfied customers).

5. Data Preprocessing

Preprocessing steps include handling missing values, data deduplication, and label encoding.

Initially, missing values were identified, and fusion imputation was utilised to handle them as it exhibited exceptional performance, yielding the best preprocessed dataset with a Mean Square Error of 0.18008, outperforming the outcomes of KNN and Regression imputation techniques as depicted in Figure 3.

After obtaining the efficient imputed dataset, duplicates were removed. The dataset has been reduced to 5085 unique entries. After imputation and deduplication, as depicted in Figure 4.

After handling missing values and performing deduplication, label encoding is applied in order to convert the categoricals into numerical ones to make the data ready to efficiently apply predictive analytics. The dataset after label encoding is depicted in Figure 5.

6. Descriptive Analytics and Diagnostic Analytics: Comprehensive Insight Discovery

The architecture of the descriptive analytics and diagnostic analytics in the proposed work is depicted in Figure 6.

Price, Offer%, and Rating are identified as the three essential features by the correlation analysis depicted in Figure 7, and the outcome attribute fulfilled is highly correlated to price.

(15730, 15)

	id	title	Rating	maincateg	price1	actprice1	Offer %	norating1	noreviews1	star_5f	star_4f	star_3f	star_2f	star_1f	fulfilled1
0	16695	Fashionable & Comfortable Bellies For Women (...	3.9	Women	698	999	30.13%	38.0	7.0	17.0	9.0	6.0	3	3	0
1	5120	Combo Pack of 4 Casual Shoes Sneakers For Men ...	3.8	Men	999	1999	50.03%	531.0	69.0	264.0	92.0	73.0	29	73	1
2	18391	Cilia Mode Leo Sneakers For Women (White)	4.4	Women	2749	4999	45.01%	17.0	4.0	11.0	3.0	2.0	1	0	1
3	495	Men Black Sports Sandal	4.2	Men	518	724	15.85%	46413.0	6229.0	1045.0	12416.0	5352.0	701	4595	1
4	16408	Men Green Sports Sandal	3.9	Men	1379	2299	40.02%	77.0	3.0	35.0	21.0	7.0	7	7	1

Fig. 2 Dataset overview: Dimensions and sample entries

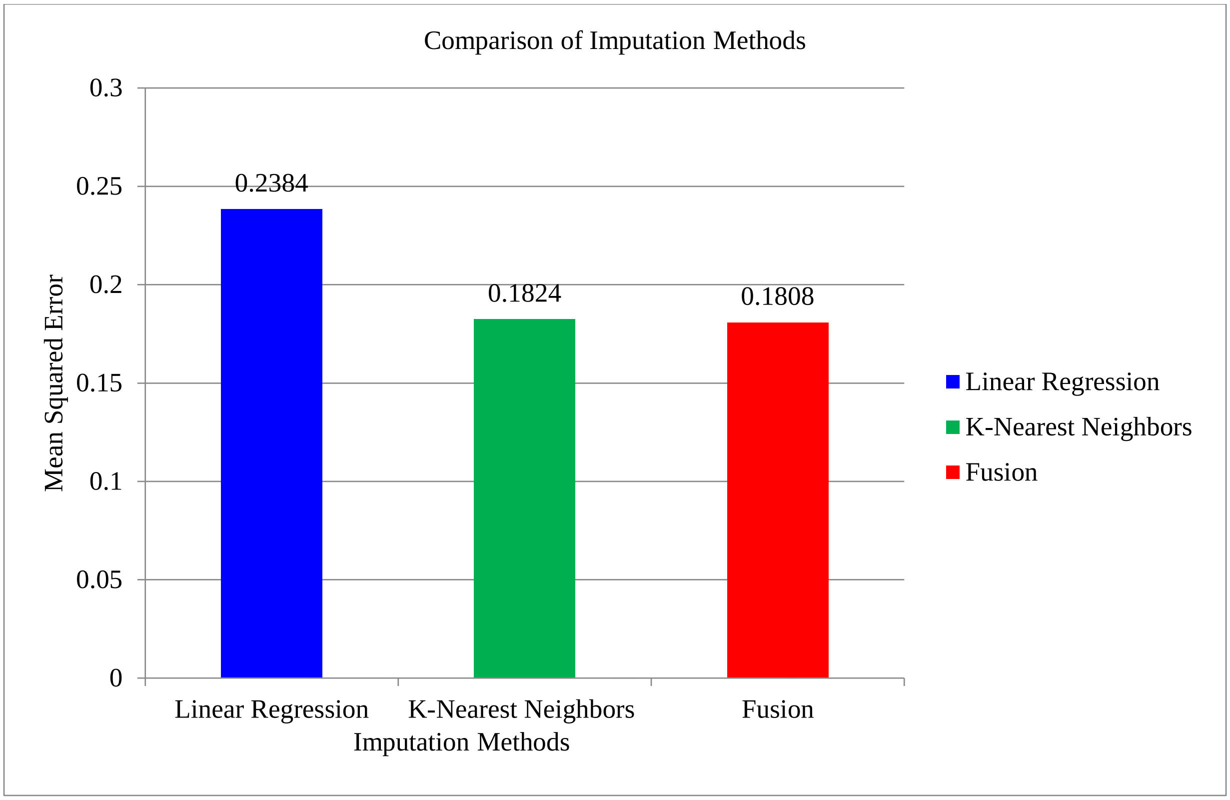

Fig. 3 A Comparative analysis of imputation techniques

(5085, 15)

	id	title	Rating	maincateg	price1	actprice1	Offer %	norating1	noreviews1	star_5f	star_4f	star_3f	star_2f	star_1f	fulfilled1
0	16695	Fashionable & Comfortable Bellies For Women (...	3.9	Women	698	999	30.13	38.0	7.0	17.0	9.0	6.0	3	3	0
1	5120	Combo Pack of 4 Casual Shoes Sneakers For Men ...	3.8	Men	999	1999	50.03	531.0	69.0	264.0	92.0	73.0	29	73	1
2	18391	Cilia Mode Leo Sneakers For Women (White)	4.4	Women	2749	4999	45.01	17.0	4.0	11.0	3.0	2.0	1	0	1
3	495	Men Black Sports Sandal	4.2	Men	518	724	15.85	46413.0	6229.0	1045.0	12416.0	5352.0	701	4595	1
4	16408	Men Green Sports Sandal	3.9	Men	1379	2299	40.02	77.0	3.0	35.0	21.0	7.0	7	7	1

Fig. 4 Dataset after imputation and deduplication

	id	title	Rating	maincateg	price1	actprice1	Offer %	norating1	noreviews1	star_5f	star_4f	star_3f	star_2f	star_1f	fulfilled1
0	16695	1257	3.9	1	698	999	30.13	38.0	7.0	17.0	9.0	6.0	3	3	0
1	5120	802	3.8	0	999	1999	50.03	531.0	69.0	264.0	92.0	73.0	29	73	1
2	18391	686	4.4	1	2749	4999	45.01	17.0	4.0	11.0	3.0	2.0	1	0	1
3	495	2039	4.2	0	518	724	15.85	46413.0	6229.0	1045.0	12416.0	5352.0	701	4595	1
4	16408	2109	3.9	0	1379	2299	40.02	77.0	3.0	35.0	21.0	7.0	7	7	1

Fig. 5 Dataset after label encoding

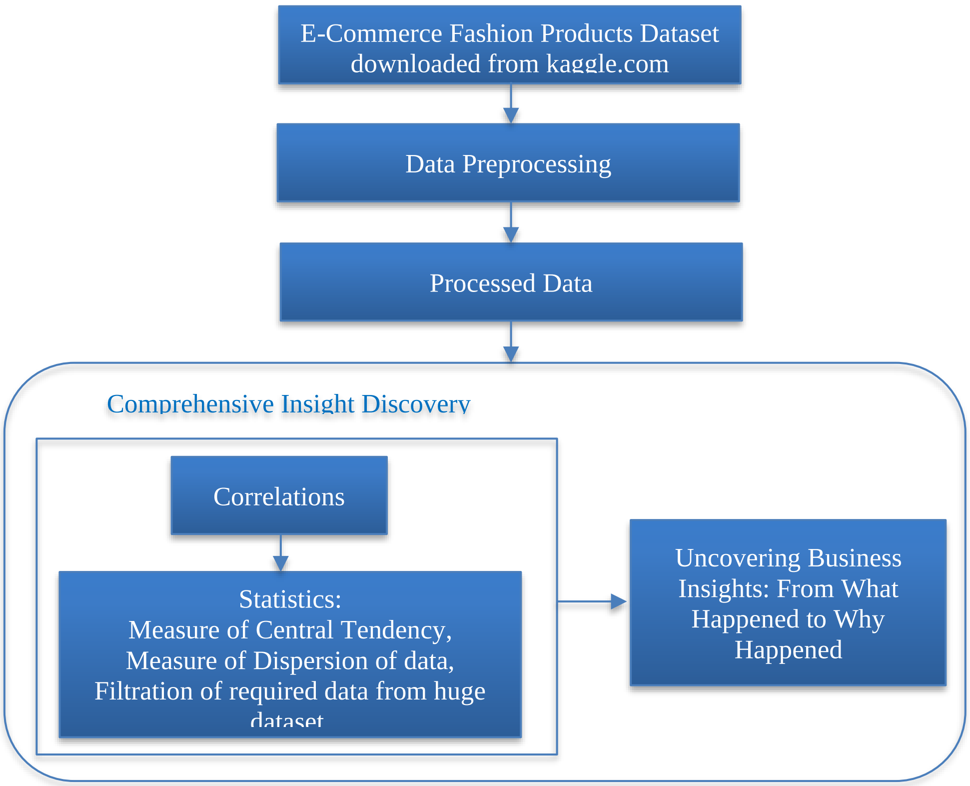

Fig. 6 Architecture of the proposed descriptive analytics and diagnostic analytics

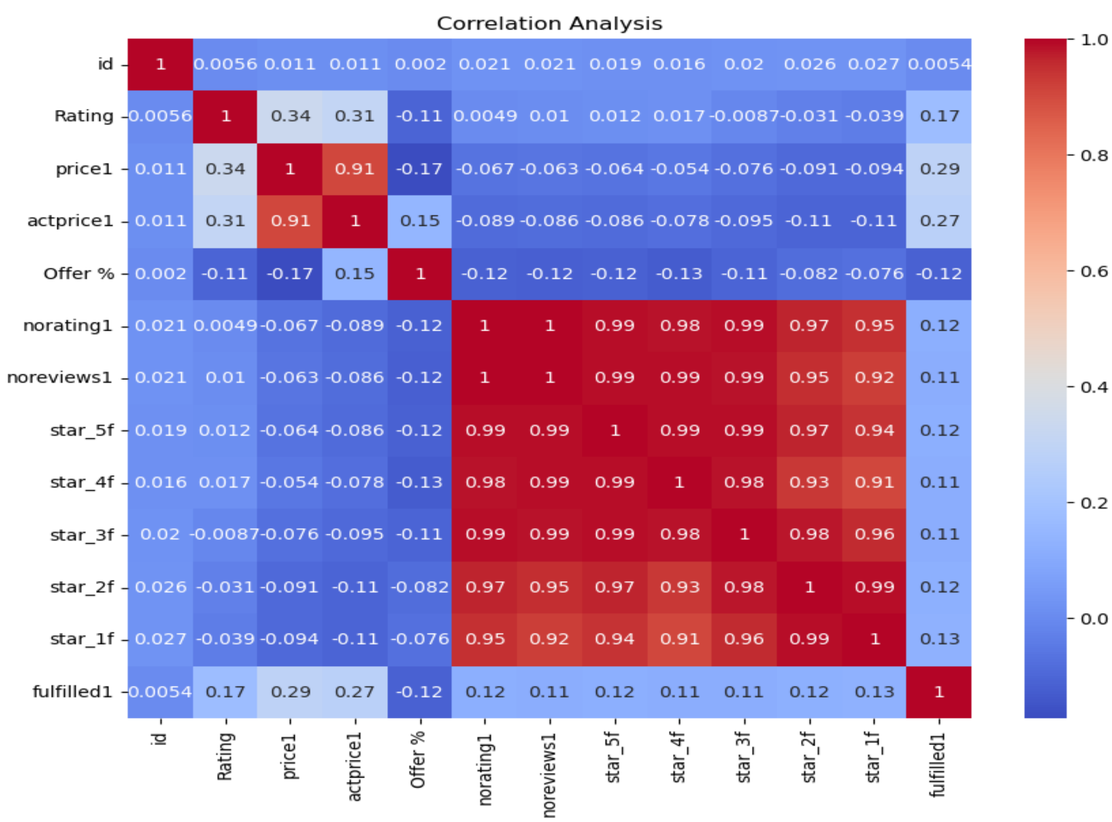

Fig. 7 Correlation analysis

Further analysis aims to identify the price ranges that attract customers, the significance of discount offers on product sales, and customer satisfaction levels with their purchases. To achieve this, price quartiles are implemented separately for both men's and women's products. The graphs in Figures 8 and 9 present data on price quantiles for both men's and women's products, respectively, including the total number of products purchased, the minimum and maximum offer, the count of fulfilled products, the count of fulfilled products with high ratings and their percentage relative to the total fulfilled orders and overall products in each quartile.

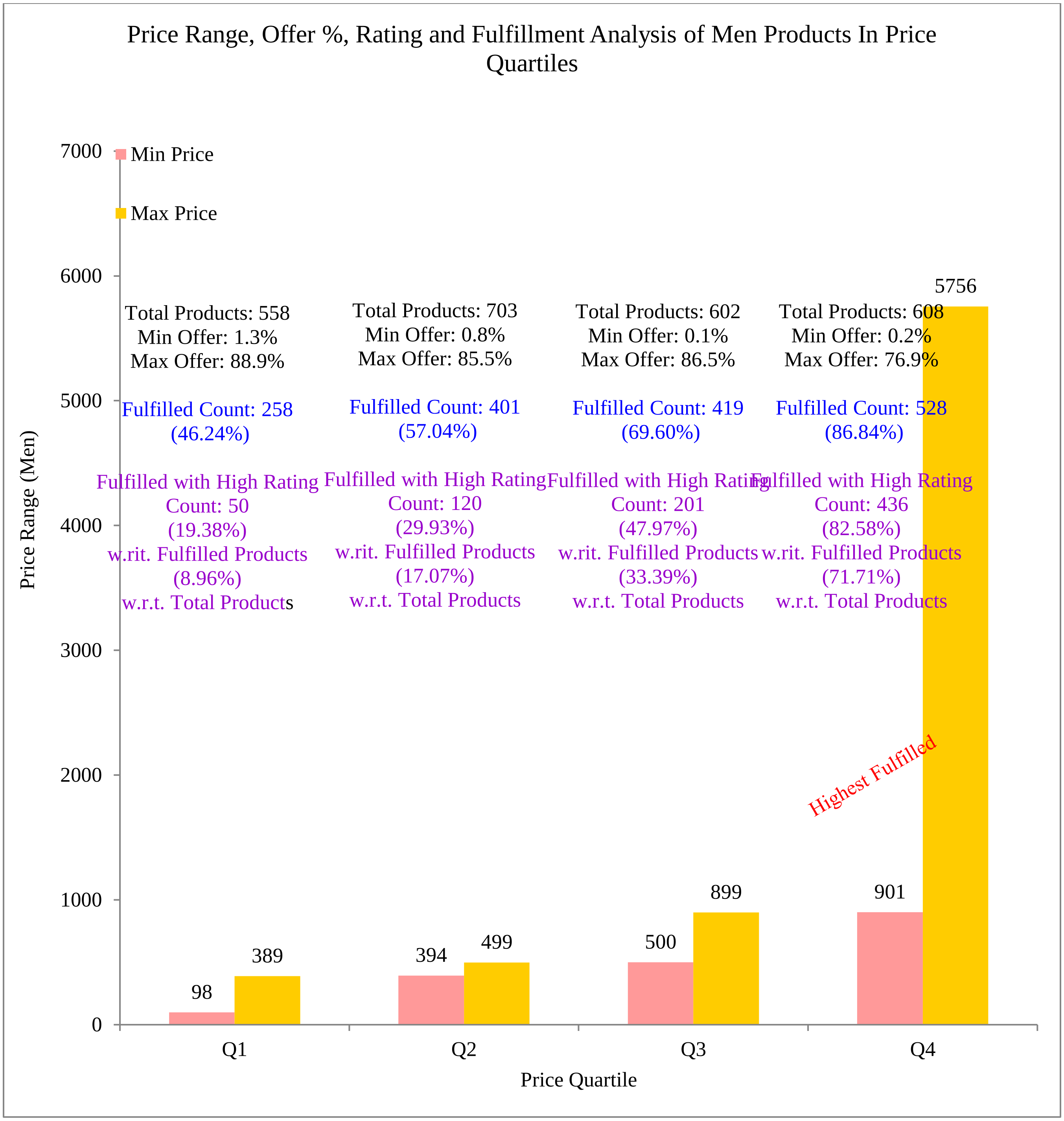

Fig. 8 Analysis of key attributes of men's products in price quartiles

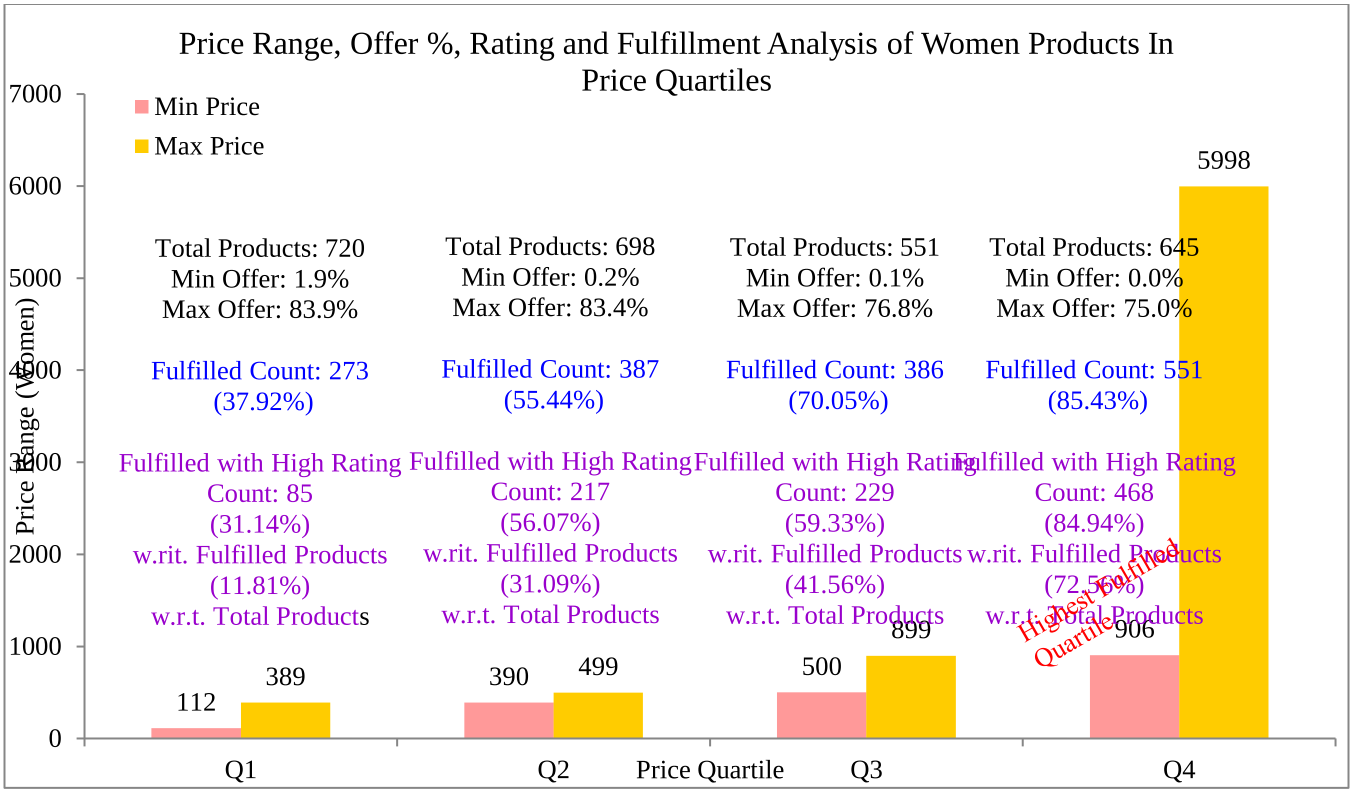

Fig. 9 Analysis of key attributes of women's products in price quartiles

Insights from the analysis across price quartiles depicted in Figures 8 and 9 are presented in Table 1.

Table 1. Insights from the analysis of key attributes of men's and women's products in price quartiles

Price Range	Total no. of products purchased by customers	Customer Satisfaction in percentage
Men's Products		
390 to 500	703	29%
900 to 6000	608	82%
Women's Products		
100 to 400	720	31%
400 to 500	698	56%
900 to 6000	645	85%

6.1. Key observations from the Analysis of key Attributes of Men's and Women's Products across Price Quartiles

- A higher number of men's products were found in the price ranges of 390 to 500 and 900 to 6000, while a higher number of women's products were found in the price ranges of 100 to 500 and 900 to 6000.
- Fulfillment for men's products is low in the 390 to 500 price range but high in the 900 to 6000 range. For women's products, fulfillment is very low in the 100 to 400 and 400 to 500 price ranges but high in the 900 to 6000 range.

6.2. Recommendations from the Analysis of key Attributes of Both Men and Women Products across Price Quartiles

- Focusing on quality improvement of men's products in the 390 to 500 price range and women's products in the 100 to 500 price range could enhance customer satisfaction and attract more buyers.
- For products in the 900 to 6000 price range, it is crucial to maintain the quality that currently satisfies customers and consider adjusting offers to maximize business profits.

7. Predictive Analytics: Future Trends

Predictive analytics techniques involve building machine learning models to predict customer behavior. The architecture of the proposed predictive analytics is depicted in Figure 10.

Random Forest, Gradient Boosting, AdaBoost, XGBoost, LightGBM, Voting, Bagging, and Stacking are popular ensemble learning algorithms that are applied to predict customer responses based on historical data.

7.1. Evaluation Metrics

The model's predictive performance is assessed using accuracy, precision, recall, and F1 score, which collectively offer an efficient evaluation [14].

7.1.1. Accuracy

Proportion of total correct predictions.

$$Accuracy = \frac{TP+TN}{TP+TN+FP+FN} \quad (1)$$

7.1.2. Precision

Proportion of true predicted positives among all predicted positives.

$$Precision = \frac{TP}{TP+FP} \quad (2)$$

7.1.3. Recall (Sensitivity)

Proportion of true predicted positive among all actual positives.

$$Recall = \frac{TP}{TP+FN} \quad (3)$$

7.1.4. F1 Score

Harmonic mean of precision and recall.

$$F1 = 2 * \frac{Precision*Recall}{Precision+Recall} \quad (4)$$

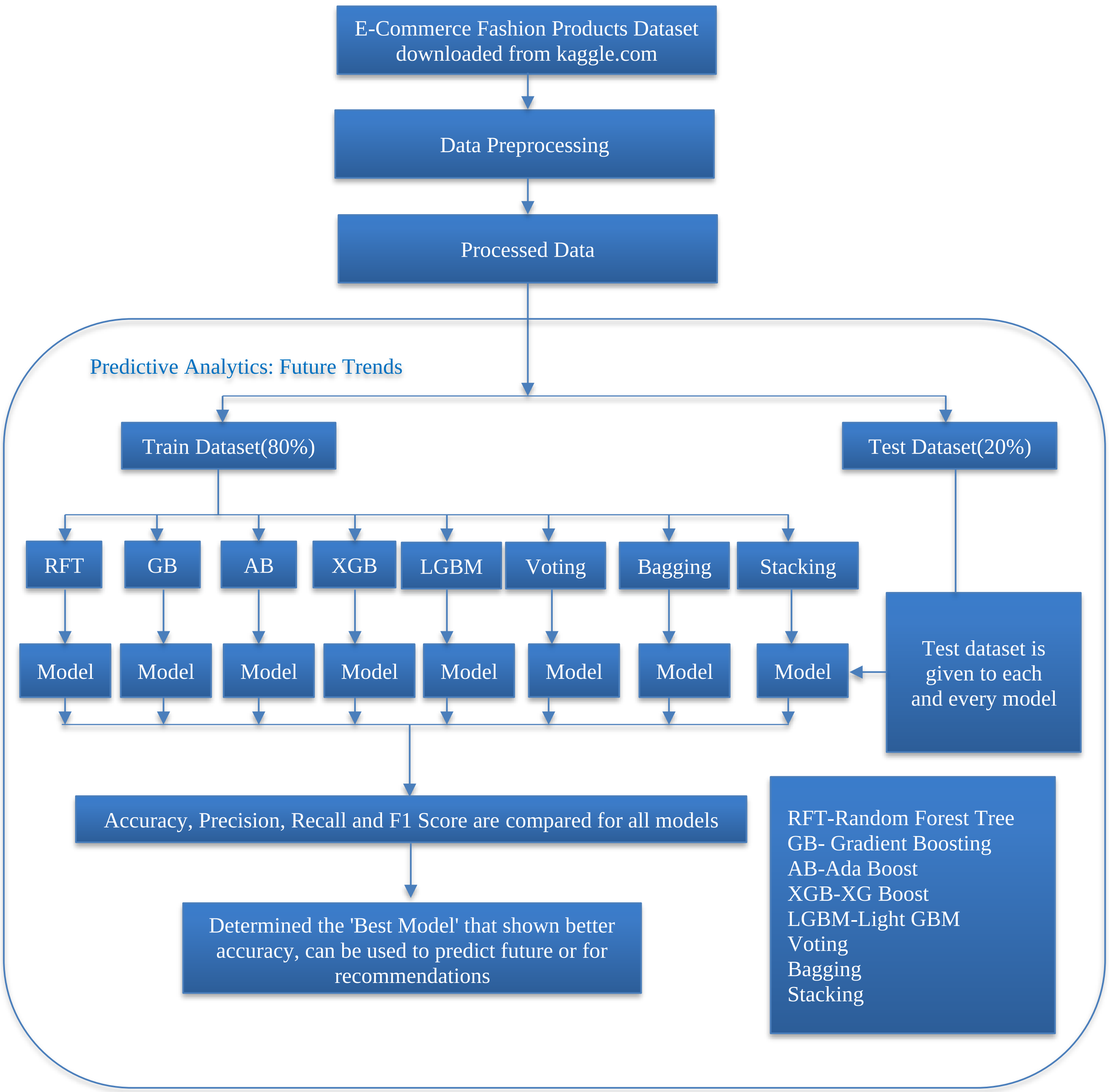

Fig. 10 Architecture of the proposed predictive analytics

7.2. Random Forest Classifier

It uses numerous decision trees on random sets of data to improve accuracy and control overfitting, and the final decision is based on what most trees "vote" for [15].

The formula for aggregating results in a Random Forest is:

$$\hat{y} = mode\ \{ h_1\ (x), h_2\ (x), \dots, h_n(x)\} \quad (5)$$

7.3. AdaBoost Classifier

It constructs a chain of weak learners, with each learner correcting the mistakes of the earlier ones by increasing the weight of the misclassified data points, helping the model improve its accuracy over time [16].

The formula for aggregating results in Adaptive Boosting is:

$$\hat{y} = sign(\textstyle\sum_{i=1}^{N} \propto_i\ h_i(x)) \quad (6)$$

Where N is the number of weak learners (trees) and $\propto_i$ is the weight assigned to the ith weak learner.

7.4. Gradient Boosting Classifier:

It is similar to AdaBoost, which optimizes a loss function directly instead of focusing on reducing classification errors [17].

The formula for aggregating results in Gradient Boosting is:

$$\hat{y} = \textstyle\sum_{i=1}^{N} \gamma_i h_i(x) \quad (7)$$

Where N is the number of weak learners (trees) and γ_i is the weight assigned to the ith weak learner.

7.5. XGBoost Classifier

It is a fast and efficient version of gradient boosting, designed with several optimizations [18].

The formula for aggregating results in Extreme Gradient Boosting is:

$$\hat{y} = \textstyle\sum_{k=1}^{K} f_k(x) \quad (8)$$

Where $f_k(x)$ are the outputs of individual trees.

7.6. LightGBM Classifier

It is another gradient boosting framework which speeds up training by using Gradient-based One-Side Sampling (GOSS) to select a subset of data points [19].

The formula for aggregating results is similar to XGBoost as:

$$\hat{y} = \textstyle\sum_{k=1}^{K} f_k(x) \quad (9)$$

Where $f_k(x)$ are the outputs of individual trees.

7.7. Voting Classifier

It combines multiple machine learning models and predicts the class label by hard voting (majority vote) [20].

The formula for aggregating results in Hard Voting is:

$$\hat{y} = mode\ \{ h_1\ (x), h_2\ (x), \dots, h_n(x)\} \quad (10)$$

7.8. Bagging Classifier

It trains several base models separately on different random subsets of the training data, with replacement and predicts the class label by taking the majority vote [20].

The formula for aggregating results in Bagging is:

$$\hat{y} = \frac{1}{N} \textstyle\sum_{i=1}^{N} h_i(x) \quad (11)$$

7.9. Stacking Classifier

It improves predictions by using a meta-learner to combine the results of several base models. Instead of just voting on the predictions from each model, stacking trains a meta-learner to learn how to best combine these predictions [21].

The formula for aggregating results in Stacking is:

$$\hat{y} = \left(g(h(x))\right) \quad (12)$$

Where g is the meta-classifier and $h(x)=[h_1(x), h_2(x), \dots, h_N(x)]$ are predictions from base models.

In all the above formulae, $\hat{y}$ it corresponds to the final prediction, and $h_i(x)$ is the prediction of the ith tree or learner or model for input x

For evaluating the performance of each technique, confusion matrices were constructed, as depicted in Fig. 11 to Figure 18. Next, we perform a comparative analysis of evaluation metrics across all techniques to identify the most effective ensemble model, as depicted in Figure 19.

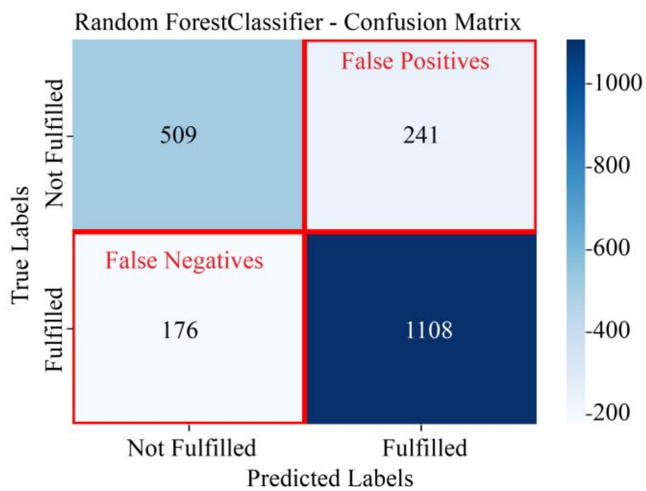

Fig. 11 Confusion matrix of Random Forest classifier

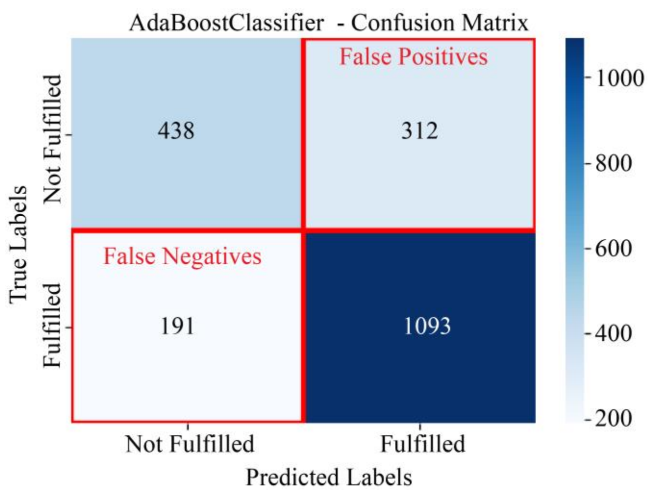

Fig. 12 Confusion matrix of AdaBoost classifier

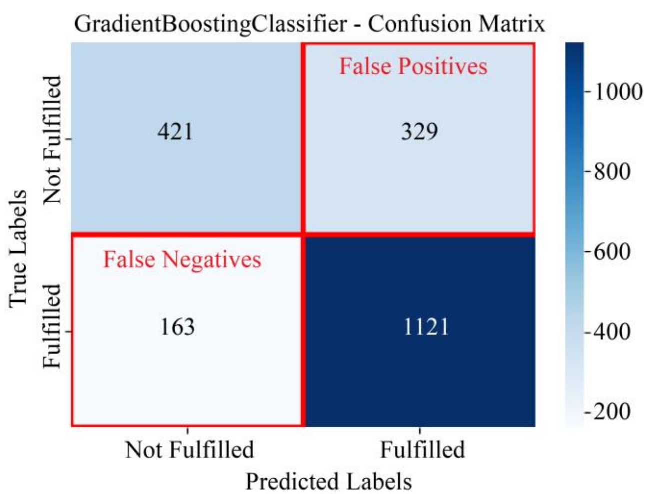

Fig. 13 Confusion matrix of Gradient Boosting classifier

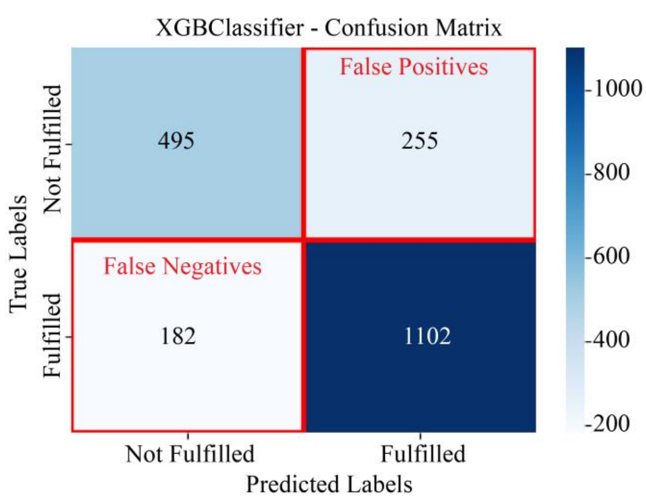

Fig. 14 Confusion matrix of XGBoost classifier

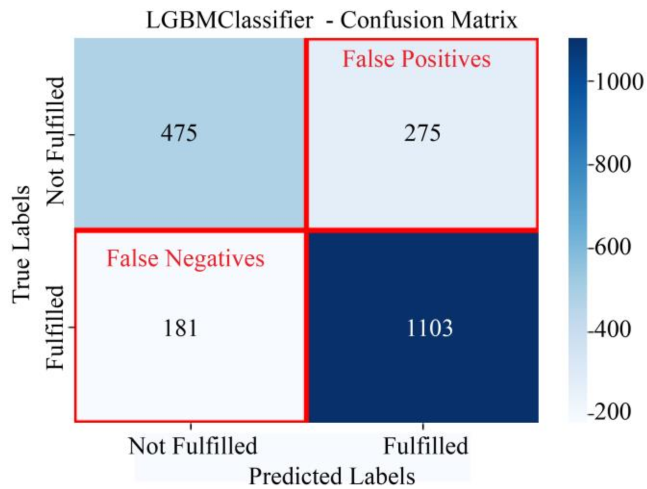

Fig. 15 Confusion matrix of LightGBM classifier

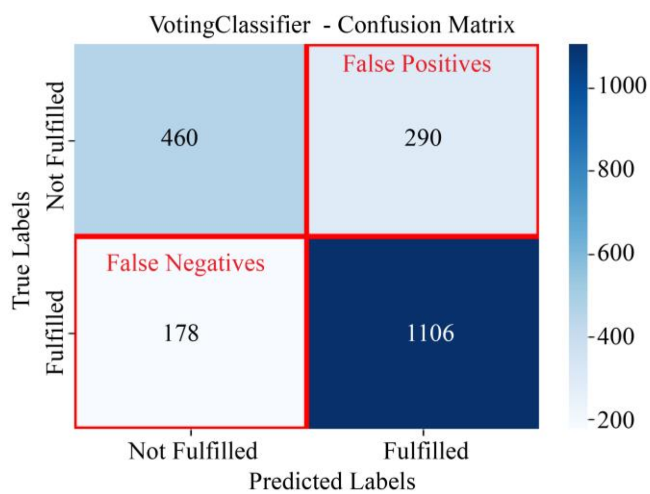

Fig. 16 Confusion matrix of Voting classifier

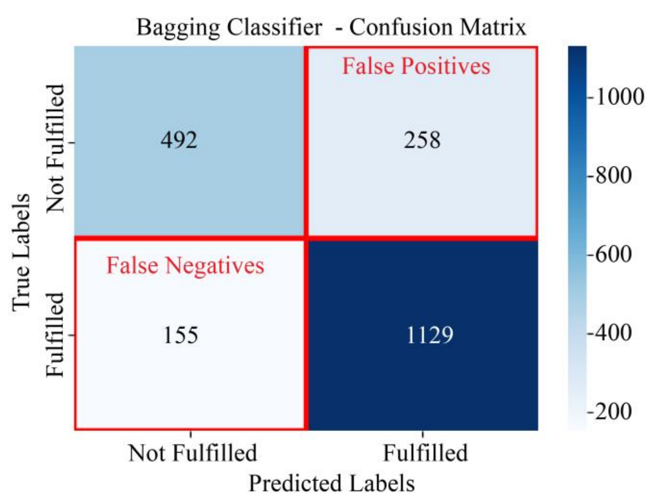

Fig. 17 Confusion matrix of Bagging classifier

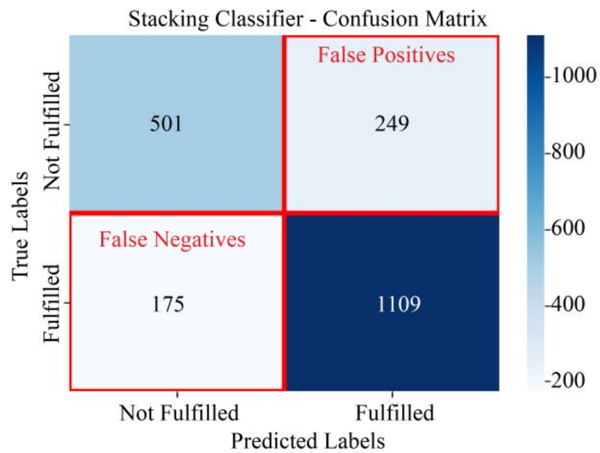

Fig. 18 Confusion matrix of Stacking classifier

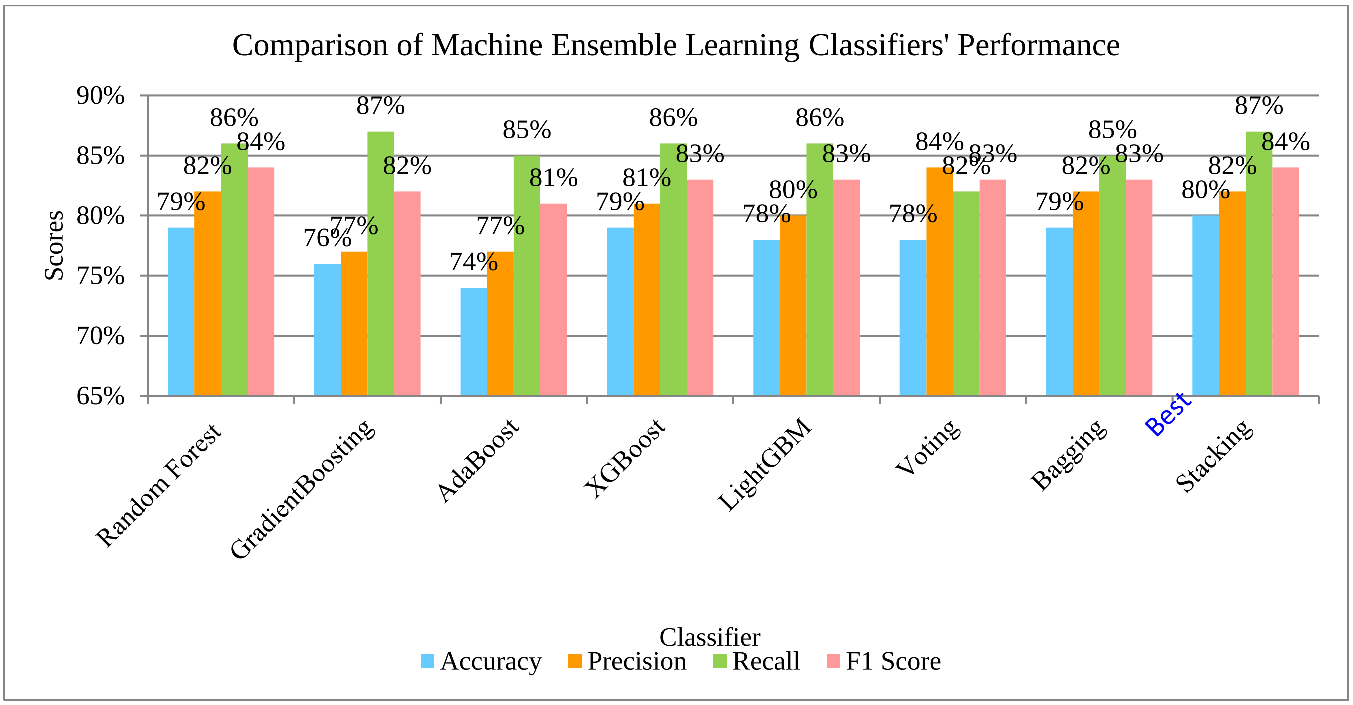

Fig. 19 Performance comparison of Ensemble Learning algorithms

From the above analysis, the Stacking classifier has emerged as the top-performing model with the highest accuracy of 80%, precision of 82%, recall of 87%, and F1-score of 84%.

8. Prescriptive Analytics: Prescriptive Solutions or Recommendations

Prescriptive analytics involves formulating actionable strategies based on insights derived from predictive models and data analysis. By leveraging predictive analytics results and considering business objectives and constraints, prescriptive analytics aims to optimize decision-making processes and maximize desired outcomes [22]. The architecture of the proposed prescriptive analytics is depicted in Figure 20.

In prescriptive analytics, the Stacking classifier, the best-performed model, was utilized to conduct further analysis. First, the feature importance of the Stacking model was assessed to understand the factors influencing fulfilment rates. Subsequently, an experimental analysis was performed by modifying the prices of 35 randomly selected not-fulfilled entries and observing the significance of fulfilment rates.

8.1. Feature Importance

Feature importance analysis is crucial in identifying the factors that have more impact on the Stacking model's predictions, which can enhance business insights and decision-making processes [23].

In this case, the analysis revealed that the price feature had the highest importance, as depicted in Figure 21, indicating that price plays a critical role in determining whether a product gets fulfilled.

8.2. Experimental Analysis: Price Adjustment

To further explore the relationship between price and product fulfilment, an experiment was conducted. Prices of 35 randomly selected not-fulfilled entries were increased and decreased by 5%, 10%, 15%, and 20%, respectively. The number of products that were fulfilled for each price adjustment was observed and visualized in Fig. 22 and Fig. 23.

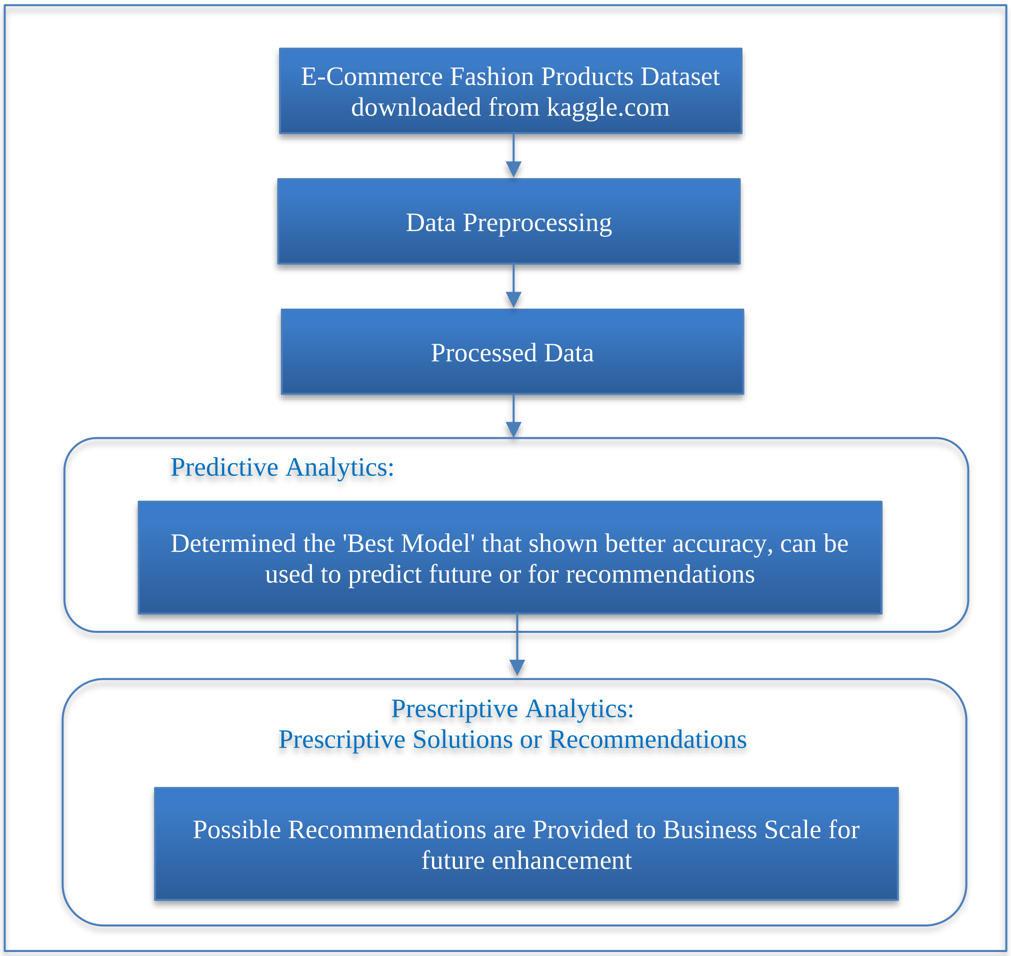

Fig. 20 Architecture of the proposed prescriptive analytics

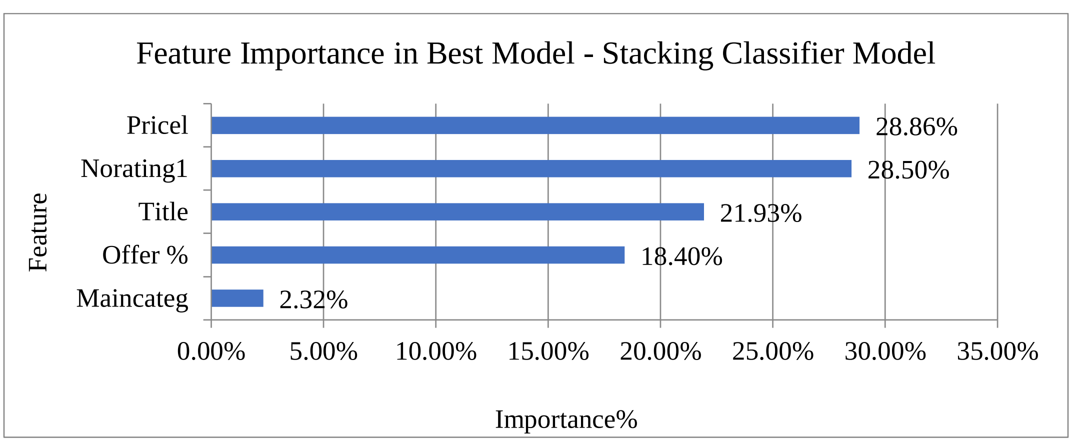

Fig. 21 Feature importance analysis in Stacking classifier

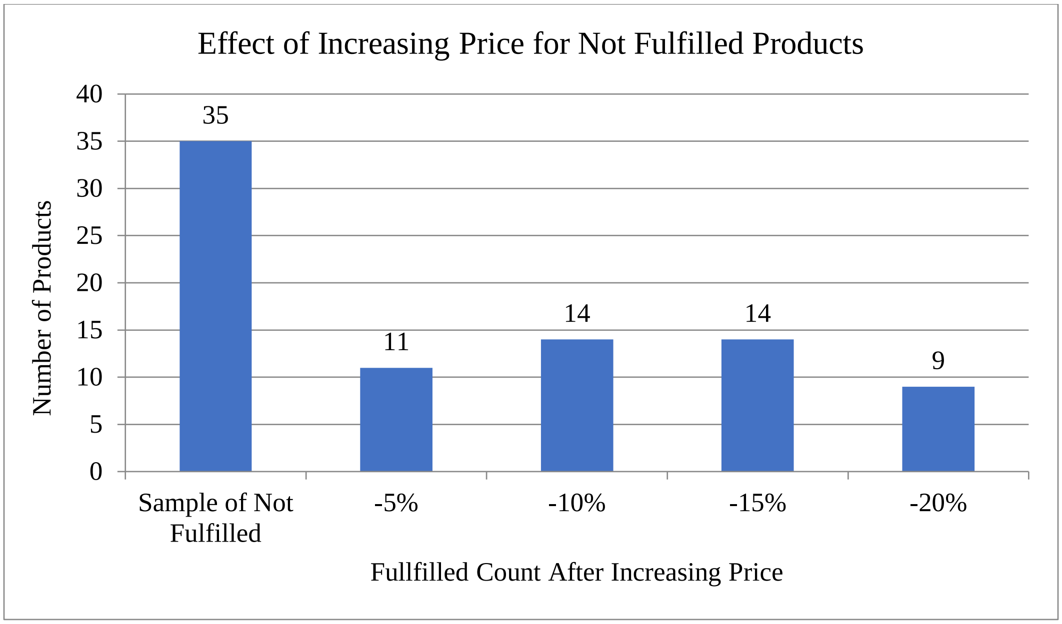

Fig. 22 Effect of increasing price on product fulfilment

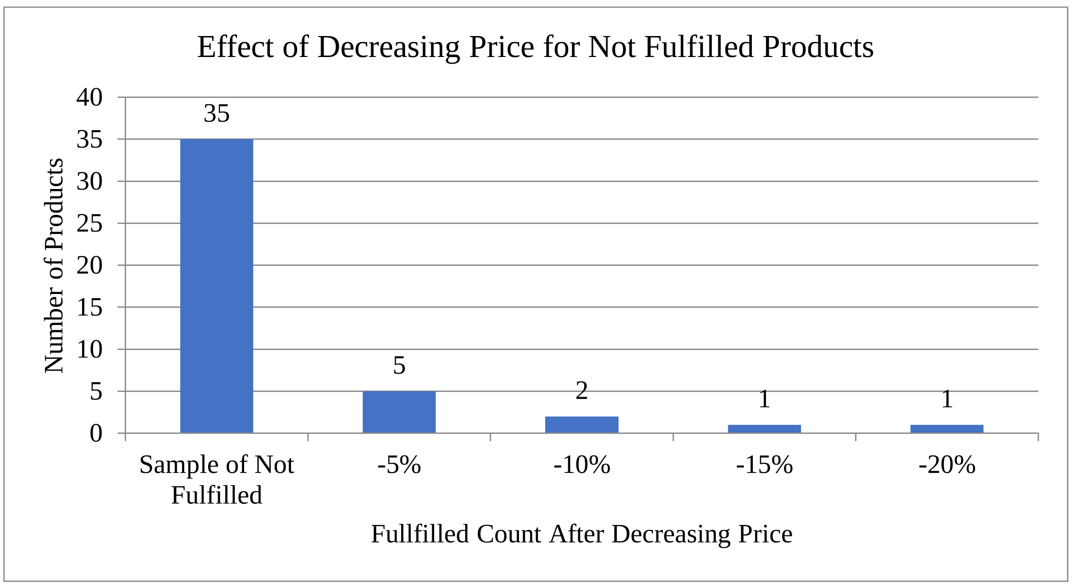

Fig. 23 Effect of decreasing price on product fulfilment

8.3. Key Observations from the Effect of Increasing and Decreasing Price on Product Fulfilment:

- *Optimal Price Increase:* A moderate price increase of 10-15% yields the highest fulfilment rates, suggesting customers may perceive these products as higher quality or better value.
- *Price Sensitivity:* When the price increase reaches 20%, the fulfilment rate declines, indicating a threshold where customers perceive the price as too high.
- *Limited Effectiveness of Price Decrease:* Lowering prices has a minimal significance on increasing fulfilment rates, indicating that for these products, price reduction alone does not significantly influence purchasing behavior.
- *Customer Perception:* Customers may associate higher prices with better product quality or value, leading to increased fulfilment rates for moderately higher prices.

8.4. Recommendations

- *Strategic Price Increases:*
  - ✓ *Targeted Strategy:* Increase prices by 10-15% for products that are not being fulfilled. This approach is likely to enhance fulfilment rates without adversely affecting customer perception.
  - ✓ *Impact Monitoring:* Closely monitor the effects of these changes to ensure that the desired outcomes in fulfilment rates are achieved.

- *Avoid Aggressive Price Hikes:* Avoid price increases of 20% or more, as they can lead to decreased fulfilment rates. Customers may perceive these products as overpriced, impacting their purchasing decisions.
- *Re-evaluate Price Decrease Strategies:*
  Given the limited significance of price decreases on fulfilment rates, focus on enhancing product quality, marketing, and customer service to improve product attractiveness.
  - ✓ *Targeted Promotions:* Use targeted promotions or discounts instead of widespread price cuts to incentivize purchases without significantly affecting perceived product value.
- Ensure that any price increase is accompanied by clear communication of the value and quality improvements. Highlight product benefits, customer testimonials, and unique selling points to justify the higher price.
- *Data-Driven Pricing:* Continuously analyze sales and fulfilment data to identify optimal pricing strategies. Employ advanced learning techniques to predict the significance of different price changes on fulfilment rates, allowing for dynamic and responsive pricing strategies.

By adopting the above recommendations, e-commerce businesses can optimize their pricing strategies to improve fulfilment rates and enhance overall customer satisfaction.

9. Conclusion and Future Scope

As e-commerce continues to evolve quickly, many grapple with a myriad of obstacles, which often lead to their downfall. To tackle these challenges, this paper proposes a novel Data-Decision Framework, which represents a pivotal advancement in addressing the multifaceted challenges faced by e-commerce businesses. By integrating comprehensive data collection, processing, and advanced analytics methodologies, including descriptive, diagnostic, predictive, and prescriptive analytics, the framework offers a structured approach to transforming raw data into actionable insights. Through the application of machine learning classifiers, businesses can gain a deeper understanding and predictive capabilities crucial for strategic decision-making. The effectiveness of this framework, demonstrated through an e-commerce dataset, underscores its potential to enhance operational efficiency, align products with market demands, and improve customer satisfaction across businesses of varying scales. By leveraging these analytical tools, e-commerce enterprises can navigate complexities, optimize resources, and sustain growth in competitive markets. Future research should explore further refinements and applications of this framework to continually enhance its efficacy and adaptability in dynamic e-commerce environments.

References

[1] Thomas H. Davenport, and Jill Dyché, “Big Data in Big Companies,” *International Institute for Analytics*, vol. 3, pp. 1-31, 2013. [Google Scholar] [Publisher Link]

[2] Amir Gandomi, and Murtaza Haider, “Beyond the Hype: Big Data Concepts, Methods, and Analytics,” *International Journal of Information Management*, vol. 35, no. 2, pp. 137-144, 2015. [CrossRef] [Google Scholar] [Publisher Link]

[3] Hsinchun Chen, Roger H.L. Chiang, and Veda C. Storey, “Business Intelligence and Analytics: From Big Data to Big Impact,” *MIS quarterly*, vol. 36, no. 4, pp. 1165-1188, 2012. [CrossRef] [Google Scholar] [Publisher Link]

[4] Matthew A. Waller, and Stanley E. Fawcett, “Data Science, Predictive Analytics, and Big Data: A Revolution that Will Transform Supply Chain Design and Management,” *Journal of Business Logistics*, vol. 34, no. 2, pp. 77-84, 2013. [CrossRef] [Google Scholar] [Publisher Link]

[5] Shahriar Akter et al., “How to Improve Firm Performance Using Big Data Analytics Capability and Business Strategy Alignment?,” *International Journal of Production Economics*, vol. 182, pp. 113-131, 2016. [CrossRef] [Google Scholar] [Publisher Link]

[6] Jan Kietzmann, Jeannette Paschen, and Emily Treen, “Artificial Intelligence in Advertising: How Marketers Can Leverage Artificial Intelligence Along the Customer Journey,” *Journal of Advertising Research*, vol. 58, no. 3, pp. 263-267, 2018. [CrossRef] [Google Scholar] [Publisher Link]

[7] Dimitris Bertsimas, and Nathan Kallus, “From Predictive to Prescriptive Analytics,” *Management Science*, vol. 66, no. 3, pp. 1025-1044, 2019. [CrossRef] [Google Scholar] [Publisher Link]

[8] Sohini Roychowdhury et al., “Categorizing Online Shopping Behavior from Cosmetics to Electronics: An Analytical Framework,” *arXiv*, 2020. [CrossRef] [Google Scholar] [Publisher Link]

[9] Weiqing Zhuang et al., “Big Data Analytics in E-commerce for the US and China through Literature Reviewing,” *Journal of Systems Science and Information*, vol. 9, no. 1, pp. 16-44, 2021. [CrossRef] [Google Scholar] [Publisher Link]

[10] Mingxi Chen, and Qunying Du, “E-commerce Marketing Strategy Based on Big Data Statistical Analysis,” *2021 13th International Conference on Measuring Technology and Mechatronics Automation (ICMTMA)*, Beihai, China, pp. 686-689, 2021. [CrossRef] [Google Scholar] [Publisher Link]

[11] Ayman Abdalmajeed Alsmadi et al., “Big Data Analytics and Innovation in E-Commerce: Current Insights and Future Directions,” *Journal of Financial Services Marketing*, pp. 1-18, 2023. [CrossRef] [Google Scholar] [Publisher Link]

[12] Anagha Ramkumar et al., “Big Data Analytics and its Application in E-commerce,” *AIP Conference Proceedings*, 2023. [CrossRef] [Google Scholar] [Publisher Link]

[13] Shivinder Nijjer et al., “Customer Analytics: Deep Dive into Customer Data,” *Encyclopedia of Data Science and Machine Learning*, pp. 1092-1107, 2023. [CrossRef] [Google Scholar] [Publisher Link]
[14] Foster Provost, and Tom Fawcett, *Data Science for Business: What you need to know about Data Mining and Data-Analytic Thinking*, O'Reilly Media, 2013. [Google Scholar] [Publisher Link]
[15] Leo Breiman, “Random Forests,” *Machine Learning*, vol. 45, pp. 5-32, 2001. [CrossRef] [Google Scholar] [Publisher Link]
[16] Jerome H. Friedman, “Greedy Function Approximation: A Gradient Boosting Machine,” *Annals of Statistics*, vol. 29, no. 5, pp. 1189-1232, 2001. [CrossRef] [Google Scholar] [Publisher Link]
[17] Yoav Freund, and Robert E. Schapire, “A Decision-Theoretic Generalization of on-Line Learning and an Application to Boosting,” *Journal of Computer and System Sciences*, vol. 55, no. 1, pp. 119-139, 1997. [CrossRef] [Google Scholar] [Publisher Link]
[18] Tianqi Chen, and Carlos Guestrin, “XgBoost: A Scalable Tree Boosting System,” *Proceedings of the 22nd ACM SIGKDD International Conference on Knowledge Discovery and Data Mining*, pp. 785-794, 2016. [CrossRef] [Google Scholar] [Publisher Link]
[19] Guolin Ke et al., “LightGBM: A Highly Efficient Gradient Boosting Decision Tree,” *Advances in Neural Information Processing Systems*, 2017. [Google Scholar] [Publisher Link]
[20] Leo Breiman, “Bagging Predictors,” *Machine Learning*, vol. 24, pp. 123-140, 1996. [CrossRef] [Google Scholar] [Publisher Link]
[21] David H. Wolpert, “Stacked Generalization,” *Neural Networks*, vol. 5, no. 2, pp. 241-259, 1992. [CrossRef] [Google Scholar] [Publisher Link]
[22] R. Divakar et al., “Prescriptive and Predictive Analysis of Intelligible Big Data,” *2022 6th International Conference on Computing Methodologies and Communication (ICCMC)*, Erode, India, pp. 845-851, 2022. [CrossRef] [Google Scholar] [Publisher Link]
[23] Sankalp Jain, and Naresh Kumar Nagwani, “A Study on Significance of Feature Selection on Product Valuation,” *2021 6th International Conference for Convergence in Technology (I2CT)*, Maharashtra, India, pp. 1-5, 2021. [CrossRef] [Google Scholar] [Publisher Link]