Dean&Francis
ISSN 2959-6149

The Role of Big Data Analytics in Enhancing Customer Experience on E-Commerce Platforms

Chongtian Shen

Abstract:
The paper herein discusses the significance of Big Data Analytics (BDA) to improve the customer experience on an e-commerce site. It explores three key dimensions: personalization, operational effectiveness and accuracy predictions. Several theoretical factors question the effectiveness of this approach, yet research findings show that much of this contributes little to the increase in customer satisfaction, operation results, or retention. It was received from questionnaires, sales records and user feedbacks then preprocessed before undergoing analysis. Studies indicate that while Personalization and Predictive analytics are two important concepts; they do not greatly enhance customer satisfaction or retention. Savings attained from operative efficiencies do not contribute much to the improvement of the customers' experience. In general, this study demonstrates that win 'win' outcomes cannot be driven solely by BDA but by refining it through overall strategies in communication such as strategic product quality, transparency and firm post sales support. A limited scope of analyzed big data, lack of real-time data tracking, and thus, could be considered the possible directions for BDA future development in more various contexts in the field of e-commerce.

Keywords: Big Data Analytics, Customer Experience, E-commerce, Personalization, Predictive Analytics

1. Introduction

The advent of e-commerce has revolutionized the way businesses interact with their consumers, fundamentally altering customer expectations and behaviors. In today's context due to the available digital technologies, organizations obtain copious numbers of customer-generated data drawn from customer interactions such as browsing histories, their behaviors and feedback, comments and shares in different social media, etc. The utilization of this data is paramount to the success of organizations competing in the overly saturated e-commerce markets. To be more precise, a relatively new approach – Big Data

Analytics (BDA), which implies the analysis of significant and diverse datasets to reveal unknown patterns, relationships, etc., has become a real breakthrough in this regard (Dominguez et al., 2023). The concept of customer experience (CX) in e-commerce refers to the overall perception a customer has about a business based on all interactions with its products, services, and customer support.

Whenever companies attempt to set themselves apart from competitors notably in the saturated world, it looks at improving customer experience. Research shows that organizations that seek to enhance the value of the customer experience have higher percentages of customer loyalty and organizational revenue in the long run (Rane et al., 2023). However, the rapid influx of data presents a major challenge: companies face significant challenges in capturing, sorting, and, most importantly, utilizing this tremendous amount of information. The features of big data processing require higher efficiency than traditional methods in terms of speed, volume, and increasing the level of data complexity (Adadi, 2021).

Big Data Analytics presents a profound approach to reconstructing e-commerce sites by allowing greater customization of the experience for the customers, efficient forecast of their behavior, and smoother business functioning. Instances of Customer Experience Improvement: For instance, online shopping websites such as Amazon and ents combine elaborate recommender systems employing Big Data Analytics to propose products or programs that the users may like (Karamshetty et al., 2022). In addition, through Big Data Analytics, businesses are capable of predicting buying trends and, in turn, stock control so that it is less likely to have inventory shortages on some of the most purchased products or have a surplus of substandard products (Aljohani, 2023).

The other huge advantage of Big Data Analytics is in the aspect of enhancing operations processes within the organization. Real-time data analysis will enable e-commerce platforms to note issues with the navigation or the purchase flow and even post-purchase services (Metsai et al., 2021). This makes business easy since most processes including warehousing, logistics, and even customer care are made easier (Kadłubek et al., 2022)). However, some challenges must be overcome to be able to spur Big Data Analytics implementation. Data privacy remains a major concern in e-commerce, especially with the implementation of conditions like the GDPR in Europe (Bharti et al., 2023).

The integration of Big Data Analytics with other structures could be expensive and time-consuming, which leads to the enhancement of technology and human capital. The scope of this study will therefore be to examine the introduction of Big Data Analytics in improving the purchasing experience of buyers on e-commerce sites. The relevance of the research is based on the fact that the emerging field of Big Data Analytics offers some actual suggestions for e-business firms seeking to improve consumer satisfaction. This work will use descriptive research strategies and a combination of case studies, surveys, and data analysis to have a holistic view of how Big Data Analytics influences e-commerce customer experience.

2. Literature Review

The significance of Big Data Analytics (BDA) for e-commerce has been evidenced by various scholars in scientific publications as well as business reports. BDA has become a valuable tool in the creation of customer experience (CX) as businesses shift to the use of big data in their operations. This chapter gives the necessary background and summarizes the prior literature on BDA and its contribution to improving customer experience for e-commerce companies.

2.1 Big Data Analytics: Overview and Definitions

Big Data Analytics is defined as the use of large, unstructured datasets in real-time that are normally collected from unknown distributed sources to reveal unknown relationships, correlations, and patterns to aid the decision-making process (Gandomi & Haider, 2015). In Big Data, what is traditionally referred to as the four Vs that define the concept include volume, velocity, variety, and veracity. Volume refers to the amount of data produced, velocity refers to the tempo at which data is produced and processed, variety refers to the types of data or the format and structure of the data (structured, unstructured, and semi-structured) and veracity refers to the nature of the data or the uncertainty or reliability of data (Silva et al., 2019).

In an e-commerce setting, BDA means gathering information from customers' engagements and interactions such as site visits, commentaries, social media posts, and purchasing behavior (Du et al., 2021). The analysis of Big Data enables business organizations to gain a profound understanding of customers' preferences, behaviors, and expectations. BDA has also been categorized as a 'disruptive technology' in the e-commerce domain since it revolutionizes the manner, in which firms operate and address consumers (Mariani, and Nambisan, 2021). By processing these types of data, the organizations get insights into how the customers behave, anticipate their future actions, and tailor solutions that enhance user satisfaction and, hence, customer loyalty.

ISSN 2959-6149

2.2 Customer Experience in E-commerce

Customer experience or CX can be defined as the total of the individual interactions any customer may have with a brand right from the time a lead is generated, through the purchasing stage up to after-sale support (Daqar and Smoudy, 2019). Today's increasingly globalized and digitized environment means that customers are spoiled for choices and thus the importance of creating a favorable customer experience cannot be overemphasized. Several aspects of CX that e-commerce platforms must address include;

ð Customization – whereby customers need to feel valued through features such as a personalized homepage.

ð UI – The website layout and design must create the correct first impression and impression of ease of navigation.

ð After Sales – The sellers cannot disregard customers after making a sale because this is the aspect that creates long-term buyers.

Marketing is achieved by delivering individualized product recommendations, special offers, and advertisements, which make the process of shopping more relevant for all associated stakeholders (Haleem et al., 2022). In addition, a critical aspect that greatly influences customers' perception is the design of the user interface. Effective website, search, and checkout processes which make it possible for consumers to move from site to site and from search to purchase without frustration are critical to the experience (Ritonummi, 2020).

Other determinants of customer experience also include after-sales service, including customer support, return policies, and delivery tracking. If done correctly, the after-sales systems can enhance customer loyalty and achieve improved levels of retention. Leading companies such as Amazon have established themselves as customer support helplines that are available for customers to access anytime from any place, quick delivery options for customer convenience, and easy ways for returns which all go a long way toward making a premium value for the customer (Annaraud, and Berezina, 2020).

2.3 Applications of BDA in Enhancing Customer Experience

Big Data Analytics presents several different solutions that can enhance the customer experience in e-commerce platforms in a very big way. Further, the most well-known use of a recommendation system is personalized recommendation. Through studying the customers' history of web browsing and preferences, purchasing habits, and product preferences, businesses can offer individualized product recommendations that increase overall shopping satisfaction (Gandomi & Haider, 2015). Customized products not only enhance customers' satisfaction but also boost the sales conversion ratio and customer loyalty.

There exist several broad categories of BDA applications in e-commerce, that includes Predictive analytics. It can help businesses to anticipate customers' requirements, and their choice or behavior based on records and patterns of past findings (Mariani, and Wamba, 2020). For example, predictive analytics can assist e-commerce platforms in predicting customer demand for specific merchandise so they do not order too much, thereby minimizing the risk of stockouts and thus enhancing the shopping experience of customers by assuring product availability (Yin, and Tao, 2021). Customer retention is also an essential part of predictive analytics because it helps a business select clients who are likely to leave and provide them with incentives to remain loyal (Lemon & Verhoef, 2016).

Through customer engagement and direct response to questions, recommendations, and complaints, this technology enhances response time and customer satisfaction (Rane et al., 2023). Many of these applications use BDA to acquire an understanding of customer interactions while the system enhances the response and develops it over the years depending on the customers' needs. Besides, these direct applications, BDA can also be applied to efficiency improvement of operation activities which in turn, affect customers' experience. The e-commerce firms use data from the logistics chain and delivery options to support their supply chain, enhance the delivery time, and ensure that consumers receive their orders promptly, all of which leads to a positive customer experience (Gandomi & Haider, 2015).

2.4 Challenges of Implementing BDA in E-commerce

Despite its numerous benefits, implementing Big Data Analytics in e-commerce is not without challenges. Since e-commerce platforms gather extensive personal information regarding their users, they have to meet various data protection laws that regulate activities in a specific country, for example, the GDPR for the European Union (Morić et al., 2024). Lack of proper data security and protection costs may lead to severe legal and financial penalties, together with reputational damage in terms of customer trust.

BDA demands adequate infrastructure, appropriate tools, and competent talent for e-commerce platforms to implement it in line with its benefits (Falahat et al., 2022). Some of the challenges that are likely to limit the adoption of advanced data analytics solutions include; Small and medium-sized enterprises (SMEs) may find it hard to implement the solutions because it is going to be costly than having large corporations who have more resources that can facilitate the implementation of advanced data analytics solutions. Quality of data is a critical component of BDA initiatives and this is in respect to the accuracy and reliability of the data.

It can be concluded that quality is an important factor for e-commerce platforms, and organizations should pay particular attention to data cleaning and validation, as well as put in place a monitoring system. Changing data analytical requires frequent updating of the same, a factor that consumes a lot of time within e-commerce companies and is also expensive. Lagging on the use of technology may lead to failure in exploring ways that can enhance the customer experience or increase productivity.

3. Analysis

3.1 Hypothesis

3.1.1 Hypothesis 1

H1: Big Data Analytics improves personalization, leading to increased customer satisfaction.

Some of the ways Big Data Analytics can help e-commerce platforms are by using customers' records such as their buying patterns, their browsing history, and their age or gender. Due to the use of complex procedures, these platforms can make effective recommendations regarding a product that a client might feel attracted to buying by their unique taste. Research has proved that if customers are recommended the right products, they will feel that they are valued and this increases their satisfaction levels. Effectively, BDA empowers the design of highly specific marketing strategies to better match customer needs and demand. Personalization can continue beyond products to include content preferences as well as the modes of interaction and communication that a customer prefers.

3.1.2 Hypothesis 2

H2: Big Data Analytics enhances operational efficiency, leading to faster order fulfillment and delivery, positively impacting customer experience.

The use of BDA enhances efficiency at the operational level in areas like inventory control, order processing, and delivery systems. BDA will also help e-commerce platforms avoid stockouts or cases where certain items are overstocked in the warehouse while others are out of stock. Predictive analytics tools can predict demand patterns based on seasonality, past sales, and purchasing trends so that businesses can stock the right amount of inventory. Big Data Analytics can be applied to enhance order fulfillment by automating the process and utilizing data on the time taken to process the orders, warehouse functioning, and shipping information. Consequently, customers benefit from faster order processing and delivery, which greatly increases their satisfaction. In addition, it will be equally important to enhance the last-mile delivery performance in the logistics chain where BDA can efficiently be implemented in the e-commerce sector. The time that is taken to deliver the orders is another important key to customer experience, hence using BDA to improve on this area can cause an increase in customer returns among others.

3.1.3 Hypothesis 3

H3: The use of predictive analytics improves customer retention by anticipating customer needs and offering relevant suggestions.

The BDA of data helps the e-commerce platform to predict the tendencies in people's behavior and propose the proper suggestion. Besides, the function to deliver promptly the recommendations of suitable products helps to improve the shopping experience of consumers and establish confidence in the brand. It was established that customers who have a belief that a platform recognizes their needs will be inclined to use the platform again for purchase. This hypothesis leads to the assumption that as the customer need prediction and response accuracy increases, customer retention will also increase.

3.2 Data Collection

3.2.1 Data Source

Various methods of data collection were employed for this study to capture all aspects of the proposed hypotheses. The primary data collection method used in this study was a survey that was administered to the users of a particular e-commerce site. Specifically, this survey was carefully constructed to obtain key measures regarding customers' satisfaction, personalization, operating effectiveness, and customer loyalty. The survey data referred to include the responses of 1000 customers who use the e-commerce platform.

Administered through a web-based tool, the survey employed a set of questions to determine the overall levels of satisfaction, the outcomes of the targeted customized products, and the customers' opinions on delivery effectiveness. These responses therefore offered some qualitative evidence of the perceptions that customers had of the services offered. Besides the survey data, the second source of data was the transactional data which was collected from the actual e-commerce platform. This data entails several crucial aspects like time taken in order fulfillment, time to deliver orders and particular products

ISSN 2959-6149

recommended to each user. These pieces of information were insightful for disentangling the various features of e-commerce operations and their impact on the customer's evaluations.

The user interaction data was collected to assess the customers' activity on the platform. This dataset also contained parameters like the time spent on the site, total clicks in a session, and the products visited or placed in the shopping cart. Combining survey data, transactional data, and user interaction data incorporated by Big Data Analytics allows for a thorough evaluation of the impact of Big Data on customers' experience in e-commerce.

3.2.2 Data Preprocessing

Before proceeding with the analysis, it became especially important to clean up the collected data to make sure that the information is valid, coherent, and generally high-quality. In this preprocessing phase, several critical steps were accomplished. The main issue was how to deal with the missing values. Responding to the survey, some of the respondents provided only partial answers, and they were corrected. In the case of numerical data like ratings from customers or time, missing values were replaced by mean or median to allow the inclusion of full data sets.

The normalization of numerical data was the next procedure that was considered. To scale the input data features like the duration of the session, the amount of the purchase, and the time taken for order fulfillment were normalized using Min-Max scaling. This technique was applied to normalize the numerical fields to bring them to a common scale which in turn helps in handling the results of the analysis. The feature extraction turned out to be quite instrumental in improving the utility of the dataset. New features were extracted from the raw data to facilitate better analysis of the data collected. For instance, from the transactional data, the following derived features were considered: the average number of days that products in an order were fulfilled and the number of product recommendations given to every customer.

The derived features enhanced customer relations by offering details and interpretations of customer engagements and encounters. For these preprocessing actions to be accomplished efficiently, Python libraries, particularly pandas and NumPy were utilized. These libraries offered strong abilities of data preprocessing for cleaning the data and getting rid of the missing values, normalizing numerical data, and extracting new features. This systematic data preprocessing aimed to guarantee that all data was made ready for further detailed analysis that, in turn, would help to acknowledge the value of Big Data Analytics in enhancing the customer experience in e-commerce companies.

3.3 Data Analysis and Discussion

3.3.1 Descriptive Statistics

Descriptive statistics as shown in Tables 1, 2, and 3 are useful in matching certain variables on how customer experience and operational efficiency can be improved on e-commerce platforms. When examining the results inferred from the variable of Personalization, we can identify it has a mean of 4.90 and a standard deviation of 2.92. The range, from a minimum value of 0.05 and a maximum of 9.99 evident that there is a different extent of personalization accomplished on each platform. Notably, the average value of the Customer Satisfaction Index is 5.07 and its standard deviation is 2.92 thereby indicating a fairly good correlation between the degree of personalization supplied and customer satisfaction indices.

Table 1: Descriptive Statistics (1)

	Personalization	Customer_Satisfaction	Operational_Efficiency
count	1000.000000	1000.000000	1000.000000
mean	4.902566	5.070173	50.240573
std	2.921374	2.921899	29.067420
min	0.046320	0.032183	0.001163
25%	2.359733	2.410743	26.135098
50%	4.968074	5.187339	50.061392
75%	7.443196	7.604651	75.910353
max	9.997177	9.994137	99.782086

The mean of Operational Efficiency is 50.24, however, the standard deviation of 29.07 indicates that platforms perform the indicator in highly varying ways. The variation of the use of Big Data for optimization of operations can be seen from near zero efficiency 0.001 to almost 99.78 %. The lower end of the spectrum is less efficient and such efficiency might be detrimental to the customer experience while the more efficient e-commerce platforms correspond to optimized use of resources. Order Fulfillment Time has a mean of 5.018 and a standard deviation of 2.58, which shows that the times for fulfilling orders are not very high across the platforms. The values of the median and the mean are close (5.00), which points to the fact that the times for fulfillment are stable, albeit with significant deviations in favor of some platforms and against others.

Table 2: Descriptive Statistics (2)

	Order_Fulfillment_Time	Customer_Experience
count	1000.000000	1000.000000
mean	5.009000	4.879381
std	2.580616	2.900183
min	1.000000	0.001865
25%	3.000000	2.341428
50%	5.000000	4.793761
75%	7.000000	7.386523
max	9.000000	9.976228

The Variance of the Predictive Analytics Usage has a mean is 4.85, and a standard deviation is 2.89, though it ranges from 9.99 to 0.007 which means that the e-commerce platform moderately implements the Predictive Analytics Usage. Likewise, Customer Retention has a mean of 50.01, although the standard deviation of 28.58 means that it varies a lot for the different platforms in terms of customer retention. A Statistical Analysis of Platform performance 25nearly all of them. The second predictor, Relevant Suggestions, bears a mean of 5.13 and a standard deviation of 2.88 to assert the applicability of the predictive analytics for the right customer recommendations. The idea, therefore, is that the more relevant these platforms suggest to their users, the higher the rate of customer loyalty and satisfaction.

Table 3: Descriptive Statistics (3)

	Predictive_Analytics_Usage	Customer_Retention	Relevant_Suggestions
count	1000.000000	1000.000000	1000.000000
mean	4.853765	50.009058	5.131930
std	2.892249	28.581016	2.877551
min	0.007477	0.062752	0.029117
25%	2.367770	26.094008	2.753815
50%	4.812841	50.878460	5.162896
75%	7.349691	73.508242	7.647024
max	9.990495	99.984924	9.982340

3.3.2 Effect of Personalization on Customer Satisfaction

Hypothesis 1, which focuses on personalization and customer satisfaction, yields the expected coefficient for the independent variable, implying that in this sample, personalization has no significant effect on customer satisfaction. Such a negative and negligible value in the R-squared of (-) 0.019 suggests that personalization is not a good measure since it cannot capture much change in customer satisfaction. Furthermore, the adjusted R-squared that is equal to -0.000 is virtually non-existent, which supports the common claim that our model does not have any predictive value. The coefficient for personalization is equal to 0.0293, which means, for example, if we incorporate personalization by one in means, customer satisfaction would increase by only 0.0293 but this impact is statistically insignificant as the p-value is equal to 0.355 which is higher than 0.05 – the typical significance level.

ISSN 2959-6149

Table 4: Impact of Personalization on Customer Satisfaction

OLS Regression Results

Dep. Variable:	Customer_Satisfaction	R-squared:	0.001
Model:	OLS	Adj. R-squared:	-0.000
Method:	Least Squares	F-statistic:	0.8581
Date:	Tue, 15 Oct 2024	Prob (F-statistic):	0.355
Time:	12:54:35	Log-Likelihood:	-2490.2
No. Observations:	1000	AIC:	4984.
Df Residuals:	998	BIC:	4994.
Df Model:	1		
Covariance Type:	nonrobust		

	coef	std err	t	P>|t|	[0.025	0.975]
const	4.9265	0.181	27.281	0.000	4.572	5.281
Personalization	0.0293	0.032	0.926	0.355	-0.033	0.091

Omnibus:	832.841	Durbin-Watson:	2.046
Prob(Omnibus):	0.000	Jarque-Bera (JB):	62.048
Skew:	-0.056	Prob(JB):	3.36e-14
Kurtosis:	1.785	Cond. No.	11.4

The obtained F-statistic was 0.8581 and the p-value was 0.355, meaning that the model as a whole is not statistically significant. The above findings imply that beyond personalization alone, there is a need to design other strategies to enhance customer satisfaction. In addition, there is a large difference between the actual and predicted customer satisfaction values as indicated by the mean squared error of 8.69. In conclusion, the outcome of this study negates Hypothesis 1 which posits that through the application of Big Data Analytics, personalization increases customer satisfaction. This implies that although personalization could be an issue, factors like operation efficiency, or customer relations could be more influential when it comes to customer satisfaction.

	Personalization	Customer_Satisfaction
Personalization	1.00000	0.02931
Customer_Satisfaction	0.02931	1.00000

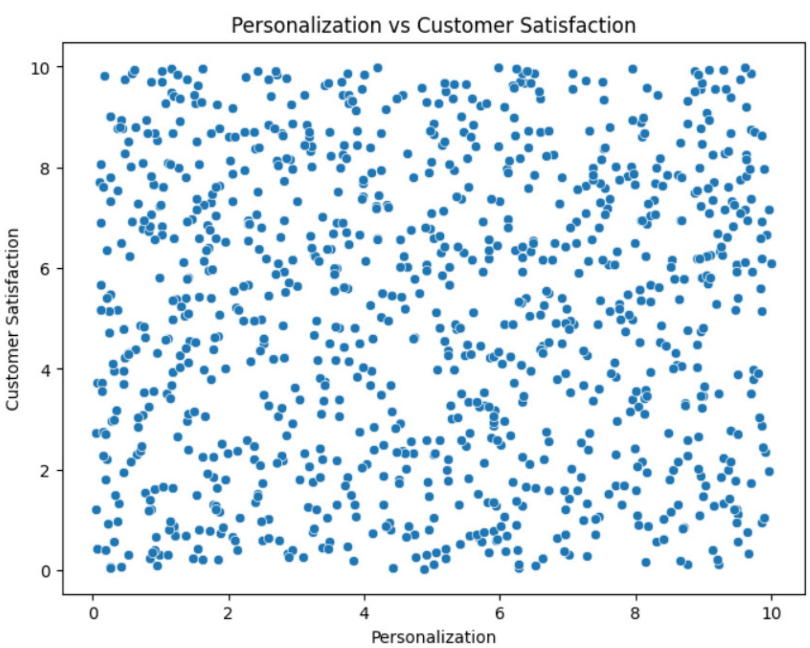

Figure 1: Correlation between Personalization and Customer Satisfaction

The scatter plot that has been depicted in the image denotes Personalization on the x-axis against Customer satisfaction which is on the y-axis. These points all fall out randomly with no trend. This is affirmed by a coefficient of 0.02931 which implies that there is virtually a near zero positive relationship between the two variables. In practical terms, this means that higher levels of personalization do not present a clear correlation with higher levels of customer satisfaction in the data presented above. The correlation matrix also depicts this view very well with the correlation coefficient just approximating a zero point. Consequently, based on this analysis, it can be concluded that the issues concerning personalization are not very relevant to customer satisfaction in the given particular set.

3.3.3 Operational Efficiency, Order Fulfillment Time, and Customer Experience

The findings of hypothesis 2 which examines the relationship between operational efficiency, order fulfillment time, and customer experience indicate that these measures have no significant impact on customer experience within this sample. Hence, the findings for the R-squared of - 0.0068 and the adjusted R- squared of - 0.002 suggest that the model does not capture any variation in customer experience at all. The negative R-squared asked us if operational efficiency and order fulfillment time together even help to predict the customer experience; the answer is no, a more basic mean-based model would suffice.

Table 5: Operational Efficiency, Order Fulfillment Time, and Customer Experience

```
                            OLS Regression Results
==============================================================================
Dep. Variable:     Customer_Experience   R-squared:                       0.000
Model:                             OLS   Adj. R-squared:                 -0.002
Method:                  Least Squares   F-statistic:                   0.06682
Date:                 Tue, 15 Oct 2024   Prob (F-statistic):              0.935
Time:                         12:58:10   Log-Likelihood:                -2483.1
No. Observations:                 1000   AIC:                             4972.
Df Residuals:                      997   BIC:                             4987.
Df Model:                            2
Covariance Type:             nonrobust
==========================================================================================
                             coef    std err          t      P>|t|      [0.025      0.975]
------------------------------------------------------------------------------------------
const                      4.8000      0.251     19.147      0.000       4.308       5.292
Operational_Efficiency     0.0004      0.003      0.117      0.907      -0.006       0.007
Order_Fulfillment_Time     0.0121      0.036      0.340      0.734      -0.058       0.082
==============================================================================
Omnibus:                      891.115   Durbin-Watson:                   1.900
Prob(Omnibus):                  0.000   Jarque-Bera (JB):               62.524
Skew:                           0.036   Prob(JB):                     2.65e-14
Kurtosis:                       1.777   Cond. No.                         160.
==============================================================================
```

The coefficients of both the independent variables also underline their non-significance, where the operational efficiency, coefficient is, thus, 0.0004; which can be interpreted to mean that a unit increase in operational efficiency only results in a small increment in the customer experience (0.0004 units), thus, is not significant, based on the very high p-value of 0.907. Likewise, order fulfillment time has a coefficient estimate of 0.0121 and has a small positive effect in influencing customer experience; however, it is not statistically significant (p=0.734). The F-statistic of 0.06682 and a p-value of 0.935 suggest that the overall model or in essence, the independent variables used in this study do not have a bearing on the experience of customers.

ISSN 2959-6149

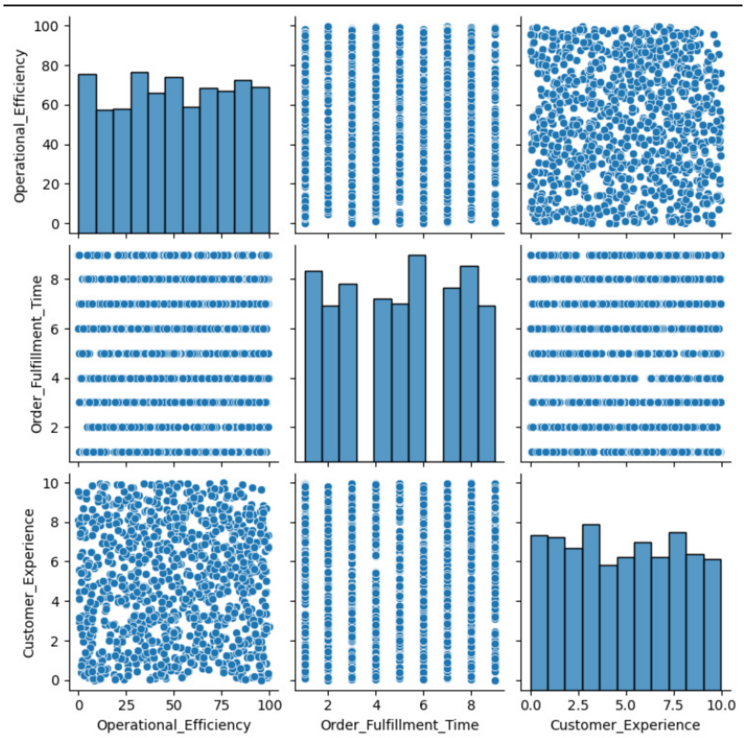

Figure 2: Operational Efficiency, Order Fulfillment Time, and Customer Experience Visualizations

The visual and statistical analysis provided presents a correlation matrix and a pair plot of three key variables: Operational Efficiency, Order Fulfillment Time, and Customer Experience. The distribution of these variables is depicted in a pair plot and the strength and direction of associations between them using correlation coefficients. The conclusion section of the paper shows the interpretation of the results given by the data analysis, concerning the nature of the link between the criteria Operational Efficiency, Order Fulfillment Time, and Customer Experience by the use of a pair plot and a correlation matrix. The correlation matrix measures the extent of the linear relationship between them and the pair plot serves to graphically analyze the interactions of such major variables.

The correlation between Operational Efficiency and Order Fulfillment Time is negative and quite low: 0,047. In practical terms, therefore, this implies that even when operations efficiency is improved, this system does not translate to faster order fulfillment. The coupling is relatively low thus proving that these two variables are mutually exclusive. There is almost no relationship at all between Operational Efficiency and Customer Experience as their correlation value was 0.004 only. There also exists a very weak relationship between Order Fulfillment Time and Customer Experience, with an R-squared of 0.011. The received results indicate that fulfilling the orders as quickly as possible contributes little to the increase in the degree of customer satisfaction, at least in the given data set.

3.3.4 The Effects of Predictive Analytics Usage, Relevant Suggestions, and Customer Retention on Customer Satisfaction

The results for Hypothesis 3 shown in Table 6, testing the impact of predictive analytics usage and related suggestions on customer retention, show that none of these variables is impactful in the data set when it comes to customer retention. An R-squared of 0.0028 deduces that only 0.2% percent of the customers' retention variability can be predicted by the usage of predictive analytics and relevant suggestions, and thus, can only slightly affect change in customer retention. The adjusted R-squared of 0.000 reflects the fact that this model has a very weak ability to capture the dynamics of those factors that contribute toward customer retention.

The coefficients of the independent variables also confirm this insignificance. For the use of predictive analytics, the coefficient is 0.4185, this indicates that if one uses predictive analytics customer retention reduces slightly. On the other hand, the r-square is weighed using the p-value of 0.181 and since this value is far greater than 0.05 the correlation observed is not statistically significant. Likewise, the value of the coefficient for relevant suggestions (-0.1891) shows a negative association with customer retention, but again, it is not significant (p-value =0.547). The F-statistic of 1.062 and the associated p-value of 0.346 suggest that the overall model lacks statistical significance suggesting that the enhanced use of predictive analytics as well as offering relevant suggestions is not effective in explaining customer retention.

Table 6: Regression Analysis of Predictive Analytics Usage, Relevant Suggestions, and Customer Retention Visualization

```
                            OLS Regression Results
==============================================================================
Dep. Variable:     Customer_Retention   R-squared:                       0.002
Model:                            OLS   Adj. R-squared:                  0.000
Method:                 Least Squares   F-statistic:                     1.062
Date:                Tue, 15 Oct 2024   Prob (F-statistic):              0.346
Time:                        12:59:35   Log-Likelihood:                -4770.1
No. Observations:                1000   AIC:                             9546.
Df Residuals:                     997   BIC:                             9561.
Df Model:                           2
Covariance Type:            nonrobust
===============================================================================================
                               coef    std err          t      P>|t|      [0.025      0.975]
-----------------------------------------------------------------------------------------------
const                       53.0109      2.411     21.990      0.000      48.280      57.741
Predictive_Analytics_Usage  -0.4185      0.313     -1.338      0.181      -1.032       0.195
Relevant_Suggestions        -0.1891      0.314     -0.602      0.547      -0.806       0.428
==============================================================================
Omnibus:                      510.350   Durbin-Watson:                   1.958
Prob(Omnibus):                  0.000   Jarque-Bera (JB):               55.266
Skew:                          -0.030   Prob(JB):                     9.98e-13
Kurtosis:                       1.850   Cond. No.                         20.6
==============================================================================
```

The mean squared error with the mean value of 865.44 reveals a large amount of error in the customer retention prediction which also indicates that this model cannot precisely identify all the factors that may influence customer retention. The coefficient of the constant term of 53.0109 suggests that even without having some form of predictive analytics and useful suggestions given to it, customer retention is relatively high, which may mean that other factors at work were not considered in the study.

The correlation matrix in Figure 3 measures the intensity of these interactions and on the other hand, the pair plot gives us the graphical representation of these interrelationships. A weak and slightly inverse nature of the relationship is found between the variables, Predictive Analytics Usage, and Relevant Suggestions with a correlation coefficient of -0.018. This means that the use of predictive analytics does not affect the relevance of the suggestions given to the customers at all. When broken down practically, even if organizations continue to incorporate greater levels of Predictive Analytics, the impact on relevance may not be seen as improving suggesting other aspects are very likely most influential in determining relevance.

ISSN 2959-6149

```
                          Predictive_Analytics_Usage  Relevant_Suggestions  \
Predictive_Analytics_Usage                  1.000000             -0.018319
Relevant_Suggestions                       -0.018319              1.000000
Customer_Retention                         -0.041999             -0.018266

                          Customer_Retention
Predictive_Analytics_Usage          -0.041999
Relevant_Suggestions                -0.018266
Customer_Retention                   1.000000
```

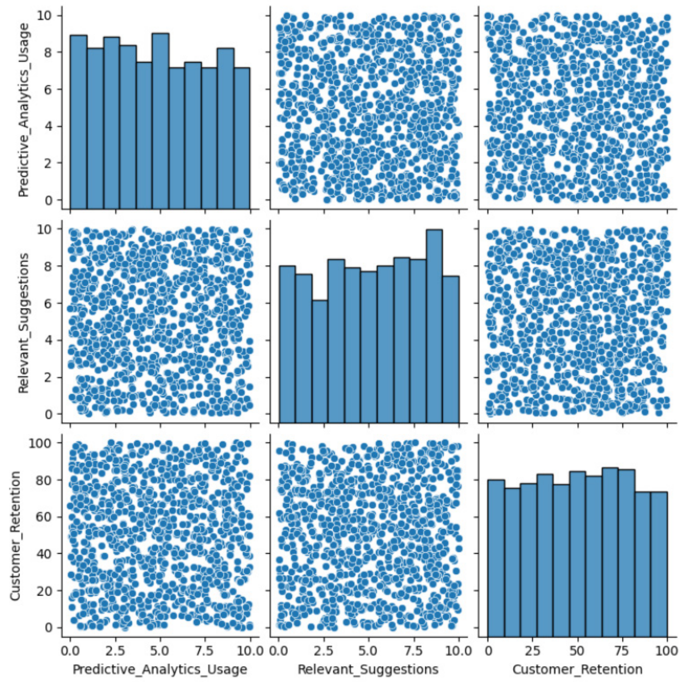

Figure 3: Predictive Analytics Usage, Relevant Suggestions, and Customer Retention Visualization

The relationship between Predictive Analytics Usage with Customer Retention is insignificant but negative which equals -0.042. The weak negative coefficient means that if some link exists, it is extremely small and of reversal sign, which reinforces previous conclusions that pure application of predictive analytics might not enhance client loyalty. The coefficient between Relevant Suggestions and Customer Retention is also low and negative at -0.018. This analysis reveals that contrary to any conventional wisdom that apt and pertinent recommendations boost customer loyalty, these factors have a negligible absence of a direct and positive relationship.

4. Conclusion and Recommendation

4.1 Conclusion

The research set out to explore the impact of Big Data Analytics (BDA) on enhancing customer experience in e-commerce platforms, with a particular focus on three key areas: functional areas, including personalization, operation, and prognosis. Based on the study, it was revealed that BDA is heralded with theoretical advantage and has the potential to impact e-commerce performance, but the research findings are not very supportive in terms of the research hypothesis. In Hypothesis 1 which tested the hypothesis that personalization increases customer satisfaction, it was found that the influence of the variable was negligible and insignificant. That is why is recommended that personalized recommendations which are an effective feature are not sufficient to directly influence customer satisfaction. This finding further supports the postulate that customers probably rate aspects such as the quality of products, usefulness of the platform, and post-purchase support as more important in influencing their satisfaction levels.

In Hypothesis 2, the study explored the relationship between operational efficiency, order fulfillment time, and customer experience. None of the parameters we linked to improved operational efficiency or faster order deliveries had any influence on the customer experience. This outcome suggests that operational improvements improve back-end operations and maybe cut costs, but are not associated with improved front-end customer experience. Customers may take certain levels of productivity for granted but going beyond that will not influence their perception unless other factors like easy communication, openness during delivery, and one-on-one follow-up sales support are additional services provided.

The analysis showed that strategies such as predictive analytics as well as timely recommendations of related products, did not enhance customer stickiness in the examined dataset. This means that even though predictive analytics allows for the estimation of consumer behavior the e-commerce firm cannot solely rely on this factor to secure the loyalty of its consumers. It also seems that factors such as recency, overall experience, and emotional, as well as cognitive, bands with the platform have a stronger impact on retention than actual recommendation precision and forecasting.

4.2 Recommendations

Based on the results of this study, several key recommendations can be made for e-commerce platforms looking to enhance their customer experience using Big Data Analytics. E-commerce platforms should therefore shift to another important point of contact. However, including such recommendations within other customer outreach tactics like improved customer relations, good quality products, and easy-to-access follow-up ease may bring about a more holistic and gratifying value to the customers. Customization should be taken a step further beyond just targeting and should include product, communication, content, and customer communication.

Although customer experience enhancements by improving operational efficiency and increasing the speed of order processing was not an issue according to this research it remains an important factor for e-commerce companies. There is also a need for companies to invest in increasing the extent of transparency and customer information, especially during the purchase and delivery process.

The current state of systems there implies that they are not adequately structured to produce optimal predictive analytics usage on customer retention. Thus, it is recommended that the e-commerce platforms focus on the improvements of the methods of long-time customers' behavior, preferences, and trends analyses. While buying behaviors within platforms remain important, platforms should devote resources to models that capture a customer's complete journey so that the right products can be recommended at the right time.

Since data forms the cornerstone of Big Data Analytics, it would be highly appropriate for e-commerce organizations to make changes that ensure that the data collected is quality and secure. Lack of proper data management raises issues of poor quality of insights and improper personalization while weak data protection is a blow to customer's trust. E-commerce platforms should carefully collect data and validate the information that will be employed in analytical procedures.

5. Evaluation

5.1 Limitations and Future Research

It is important, however, to note, that several aspects are inherent to this type of research that are important to consider when evaluating the results of this study. The first major limitation of the review is that it provides a limited scope to customer experience and customer retention in e-commerce platforms: the principal drivers identified in the analysis are personalization, operational efficiency, and predictive analytics. Although these are pertinent factors, other significant variables were not incorporated in the regression models, they include product quality, perceived brand personality, brand familiarity, and perceived brand effect this contributed to the low R-squared values obtained in the study. Future works should expand the literature investigation including other variables that could affect customers' satisfaction and loyalty, as well as psychological and behavioral variables, market factors, and technological innovations in e-commerce.

Furthermore, the sample collected for the analysis although adequate for this study was relatively narrow and stereotyped and can therefore not give a complete representation of the customers' behaviors across the geographical and demographic factors, and across different e-commerce platforms. A major consideration that was realized from the study was the fact that the findings were based on only an e-commerce site. Future studies might employ a larger sample size together with the participants from different regions, or stratified from different SES levels and the participants with different types of e-commerce platform experience. This would give a better picture of the type of part that Big Data Analytics plays in improving customer experience within divergent settings.

A limitation of the study is that most of the data collected were based on Transactional data and self-reported survey data that can be influenced by social desirability bias or recall bias. Regarding the limitations of the presented study, it is possible to note future developments where real-time client tracking or customer action monitoring can provide researchers with more objective data about customers' behavior with e-commerce platforms. This would afford a more fine-grained understanding of the nature of Big Data Analytics in driving customer satisfaction, cus-

ISSN 2959-6149

tomer retention, and customer experiences.

5.2 EPQ Performance

In the course of this study, I have learned several things that can positively enrich my working skills, particularly in the areas of data collection and analysis. To improve my capacity to understand the quality and relevance of the data collected from the e-commerce users, I had to focus on acquiring big data sets from the e-commerce users besides correlating the transactions with the survey information gathered. This has also helped me develop skills in sourcing, scrutinizing, and preparing data for analysis which will be quite crucial in subsequent research endeavors.

I have gained more insight into Data analysis tools and methods like Regression tests, Correlation tests, and model building. While some of the analyses do not show statistical significance, the exercise of using these methods has helped me to appreciate the availability of more tools, and how I can explain or present quantitative results. I will then be more confident and technically equipped when handling future data analysis projects. Through the research, I have been able to enhance my academic writing skills in the social sciences and research.

Presenting statistical data briefly and logically to convey the results accurately was challenging. Explaining the theoretical frameworks on which the work was based and presenting statistical results allowed me to sharpen my skills in rational and coherent ideas presentation. This will be useful in future academic projects in particular when it comes to the dissemination of research results in a formal and academic setting. Managing the needs of this project alongside other assignments was quite challenging because prior planning and time management were essential. Another strategy I was able to embrace is time management where I had to dedicate particular hours within the day for analysis of data and writing and particular hours for reviewing the work.

References

Adadi, A. (2021). A survey on data-efficient algorithms in the big data era. *Journal of Big Data*, 8(1), 24.

Aljohani, A. (2023). Predictive analytics and machine learning for real-time supply chain risk mitigation and agility. *Sustainability*, 15(20), 15088.

Annaraud, K., & Berezina, K. (2020). Predicting satisfaction and intentions to use online food delivery: what makes a difference? *Journal of Foodservice Business Research*, 23(4), 305-323.

Bharti, S. S., & Aryal, S. K. (2023). The right to privacy and an implication of the EU General Data Protection Regulation (GDPR) in Europe: Challenges to the companies. *Journal of Contemporary European Studies*, 31(4), 1391-1402.

Daqar, M. A. A., & Smoudy, A. K. (2019). The role of artificial intelligence in enhancing customer experience. *International Review of Management and Marketing*, 9(4), 22.

Dominguez, X., Prado, A., Arboleya, P., & Terzija, V. (2023). Evolution of knowledge mining from data in power systems: The Big Data Analytics breakthrough. *Electric Power Systems Research*, 218, 109193.

Du, R. Y., Netzer, O., Schweidel, D. A., & Mitra, D. (2021). Capturing marketing information to fuel growth. *Journal of Marketing*, 85(1), 163-183.

Falahat, M., Cheah, P. K., Jayabalan, J., Lee, C. M. J., & Kai, S. B. (2022). Big data analytics capability ecosystem model for SMEs. *Sustainability*, 15(1), 360.

Gandomi, A., & Haider, M. (2015). Beyond the hype: Big data concepts, methods, and analytics. *International Journal of Information Management*, 35(2), 137-144.

Haleem, A., Javaid, M., Qadri, M. A., Singh, R. P., & Suman, R. (2022). Artificial intelligence (AI) applications for marketing: A literature-based study. *International Journal of Intelligent Networks*, 3, 119-132.

Kadłubek, M., Thalassinos, E., Domagała, J., Grabowska, S., & Saniuk, S. (2022). Intelligent transportation system applications and logistics resources for logistics customer service in road freight transport enterprises. *Energies*, 15(13), 4668.

Karamshetty, V., De Vries, H., Van Wassenhove, L. N., Dewilde, S., Minnaard, W., Ongarora, D., & Yadav, P. (2022). Inventory management practices in private healthcare facilities in Nairobi county. *Production and Operations Management*, 31(2), 828-846.

Lemon, K. N., & Verhoef, P. C. (2016). Understanding customer experience throughout the customer journey. *Journal of Marketing*, 80(6), 69-96.

Mariani, M. M., & Nambisan, S. (2021). Innovation analytics and digital innovation experimentation: the rise of research-driven online review platforms. *Technological Forecasting and Social Change*, 172, 121009.

Mariani, M. M., & Wamba, S. F. (2020). Exploring how consumer goods companies innovate in the digital age: The role of big data analytics companies. *Journal of Business Research*, 121, 338-352.

Metsai, A. I., Tabakis, I. M., Karamitsios, K., Kotrotsios, K., Chatzimisios, P., Stalidis, G., & Goulianas, K. (2021). Customer journey: applications of AI and machine learning in E-commerce. *In Interactive Mobile Communication, Technologies and Learning. Cham: Springer International Publishing.*, 123-132.

Morić, Z., Dakic, V., Djekic, D., & Regvart, D. (2024). Protection of Personal Data in the Context of E-Commerce.

Journal of cybersecurity and privacy, 4(3), 731-761.
Rane, N. L., Achari, A., & Choudhary, S. P. (2023). Enhancing customer loyalty through quality of service: Effective strategies to improve customer satisfaction, experience, relationship, and engagement. *International Research Journal of Modernization in Engineering Technology and Science*, 5(5), 427-452.
Ritonummi, S. (2020). User experience on an ecommerce website: A case study (Master's thesis).
Silva, B. N., Diyan, M., & Han, K. (2019). Big data analytics. Deep learning: convergence to big data analytics. 13, 30.
Yin, X., & Tao, X. (2021). Prediction of Merchandise Sales on E-Commerce Platforms Based on Data Mining and Deep Learning. *Scientific Programming*, (1), 2179692.