European Journal of Sustainable Development (2025), **14**, 4, 528-552 ISSN: 2239-5938
Doi: 10.14207/ejsd.2025.v14n4p528

Digital Engagement and Gender-Specific Drivers of Customer Satisfaction in the Insurance Mining Industry

By Lassad Ben Dhiab[1], Bilal Louail[1,*], Siham Riache[1]

ABSTRACT:
This study investigates the determinants of insurance customer satisfaction with a focus on the role of digital engagement and their gender-specific effects. Since digital platforms now increasingly shape the field of determining customers' expectations, this study uses advanced statistical models like logistic regression, random forests, and gradient boosting to analyze the association between digital engagement, lowly paid premiums, and customer satisfaction. Findings indicate that web participation is a powerful predictor of satisfaction, where more participation correlates with increased satisfaction, particularly for male customers. Gender-based results indicate that female customers are highly sensitive to affordability for premiums, and thus the price strategy should be developed for this group. The study also identifies nonlinear interactions and interaction effects among such important variables as income, education, and the degree of engagement for which a customer is committed. This suggests the worth of tailored plans for customer engagement. The study contributes to the literature by introducing an advanced framework for the analysis of customer behavior among insurers, demonstrating the power of advanced analytics in uncovering underlying patterns and relationships. Therefore, insurers are required to invest in digital channels that provide differentiated experiences, gender-sensitive pricing, and segmenting customers into categories to provide more accurate solutions. Future studies are also called for in the research to examine the degree to which trust and privacy concerns underlie digital engagement, as well as examine other demographic variables that influence customer satisfaction.

Keywords: Digital Engagement, Customer Satisfaction, Insurance Industry, Gender-Specific Impacts.

1. Introduction

The insurance industry has been revolutionized at its core by digital technology, reconfiguring business models, and underwriter operating models at their core. Artificial intelligence, machine learning, and big data analytics-driven digital platforms have also opened new avenues for insurers to interact with policyholders, improve service, and increase levels of customer satisfaction (Volosovych et al., 2021). It is not just a technological shift but also a foundation shift in customers' expectations, behavior, and loyalty patterns within the era of digitization (Skaf et al., 2024). Over the centuries, customer satisfaction and loyalty have driven success in the insurance sector. Satisfied customers tend to renew policies, refer services, and build long-standing relationships with insurers. However, in today's competitive market, with customers readily comparing online offerings, the traditional determinants of satisfaction (such as affordability and claim

[1] Department of Finance and Insurance, College of Business Administration, Northern Border University, Arar, Saudi Arabia.
* Corresponding Author.

settlement) do not take effect. Online interaction, which can be defined as the degree and quality of interaction between insurance customers and insurance companies on digital platforms, has been one of the major drivers of satisfaction and loyalty (Eckert et al., 2022; Scaff et al., 2024).

In addition, gender differences in customer behavior and preference have awakened more research interest in insurance studies. Men and women were reported in the literature to possess unique digital behaviors, requirements, and drivers of satisfaction that require customized engagement strategies. Male customers are more interested in cost savings and efficiency (Adeoye et al., 2024), whereas female customers prefer personalized treatment and caring service (Bhattacharyya et al., 2024). These gendered relationships highlight the need for insurers to develop engagement strategies that meet both groups' expectations. There is so much digital transformation and satisfaction research being gathered that knowledge gaps regarding the role played by digital engagement in influencing satisfaction and loyalty continue to exist if observed through a gendered perspective. Past literature tends to address digital engagement and satisfaction individually, instead of considering the interdependence of these notions (Cappiello, 2020; Guan et al., 2020). Gender-based antecedents of digital engagement and their impact on customer loyalty are also not well researched (Skaf et al., 2024).

This study attempts to fill these gaps by investigating the impact of frequency of digital engagement on customer satisfaction and policy renewal behavior in insurance, with a focus on how gender differences hold. By gender-stratified analysis, the study would thus provide applied insight into the differential drivers of male and female customer satisfaction and loyalty.

The conclusions aim to guide the creation of even more precise digital engagement strategies that are more effective at supporting customer experience and loyalty in the insurance market. The study bridges the gap of knowledge regarding how the frequency of digital engagement impacts customer satisfaction and how gender variables play roles in these impacts. However, the conceptualization and measurement of digital engagement are still unclear. Subsequent research would have to control for depth vs. frequency of participation and determine whether passive browsing is equivalent to active transaction. That would improve model validity and enable insurers to better direct themselves in maximizing their online platforms.

And the primary objective of this study is to examine the extent of digital frequency of engagement against the extent of insurance customer satisfaction. And compare female and male differentiated drivers of customer satisfaction and customer loyalty. And provide actionable findings for insurers that can utilize in creating tailored digital engagement strategies that mirror the unique wants of male versus female policyholders.

This study is a contribution to literature since it presents novel evidence on digital engagement dynamics vis-à-vis gender-differentiated behavior within the insurance industry. The contributions of this research are twofold. In the first instance, the study combines digital engagement constructs with gender-sensitive analysis of behavior in the insurance industry to provide an overarching understanding of their nexus. In so doing, it contributes theoretical knowledge to the complex needs of male and female consumers as influenced by digital technology (Nimmagadda, 2022; Perumal Samy et al., 2022). Second, it addresses an important research gap by examining the extent of digital usage and how this makes male and female consumers satisfied and loyal differently. The study highlights gender differences in responses to digital communication channels and provides an in-depth insight into customer satisfaction and drivers of digital insurance services loyalty. Embracing a quantitative and qualitative research methodology, the study presents insurance companies with tangible recommendations for developing more successful digital engagement strategies that boost customer relationships and competitiveness. Further, the study gives guidelines for future studies for other service sectors to explore the impact of digitalization on customer experience, a contribution to the field of digital marketing and customer service.

The remainder of this paper reviews the literature on digital engagement and customer satisfaction, outlines data and methodology, reports empirical findings, and concludes with a discussion of their theoretical and practical implications.

2. Literature Review

The insurance business is undergoing a fundamental transformation, driven by the development of technology, customer modernization needs, and the growth in demand for personalized services (Ben Diab et al., 2024b). Digital engagement has turned into a considerable facilitator of customers in this fast-paced environment. The aim of this study was to examine the effect of electronic participation on client satisfaction by considering how gender, together with other demographic factors, affects client satisfaction. The aim of this analysis of the interactions was to continue exploring the determinants of the insurance company's satisfaction and making recommendations to insurance providers for the sake of improving customers' experiences.

2.1 Digital engagement and customers' satisfaction in the insurance sector:

The advent of digital technologies ranging from mobile applications to AI-powered chatbots has also significantly changed how insurance customers engage with insurance companies (Skaf et al., 2024). These technologies have concrete advantages, such as improved access to services, increased convenience, and the potential for insurers to provide more personalized experiences (Ben Dhiab et al., 2024b), and consistently demonstrate a strong link between digital engagement and customer satisfaction. Gebert-Persson et al. (2019) determined that perceived usefulness and ease of use of online claim services play a more significant role than trust in encouraging customers to use these tools. Specific to Méndez-Aparicio et al. (2020) who emphasized that efficient communication and functioning technology play an instrumental role in facilitating good experiences, and the same was verified by Eckert et al, (2022) with the inclusion that online channels for claims elevate customer satisfaction to a larger degree than traditional means. Lang and Riegel (2023) posited that the perceived utility of digital tools, their user-friendliness, and a customer's digital orientation influence their desire to utilize these services. The COVID-19 pandemic expedited the transition to digital platforms, necessitating insurers to have a deeper understanding of consumer behaviors and preferences inside these digital environments (Baranauskas and Raišienė, 2021). Agyei et al. (2021) emphasized that customer engagement is necessary to connect digital connectivity and loyalty, advocate for user-friendly platforms, streamline claim processes, and support fast customers to maintain competitiveness, and the study (Skaf et al., 2024) confirmed the extent to which digital engagement affects satisfaction. Facets of the digital experience, friendly customer care, can contribute significantly to the customer's sentiment towards such a service (Kim and Kim, 2024). Even though it is established that most customers surf with the digital channel in information-seeking contexts, nevertheless a vast majority of consumers are inclined to be extremely smitten with human interaction like verbal communication with the agent in the very buying activity itself, which is termed as webrooming (Hu and Tracogna, 2020). These results also imply that the insurers must be flexible and allocate their online strategies according to the different needs of their customers.

2.2 Consumer experience and technological change:

New technologies such as artificial intelligence, blockchain, and Internet of Things have revolutionized the insurance sector by propelling and streamlining customers' experiences. AI-driven chatbots and virtual assistants deliver rapid, tailored responses to client inquiries (Gupta et al., 2022). Blockchain enhances transparency, reduces fraud, and establishes trust, thus making operations simpler (Shetty et al., 2022; Trivedi, 2023). The Internet of Things (IoT) facilitates usage-based insurance, enabling insurers to provide tailored pricing based on individual behavior or health parameters (Kawel and Saxena, 2024). But these technologies bring many challenges, such as concerns about data privacy, algorithmic bias, and digital exclusion (Grassi, 2024; Fiot et al., 2024). These problems need to be addressed to ensure universal access to these tools and a sense of confidence in their use. Additionally, gender differences influence how customers perceive and interact with these technologies, highlighting the need for digital strategies that are inclusive and cater to different customer needs.

2.3 Gender Disparity on Customer Satisfaction and Digital Interaction:

Gender plays an important role in digital customer interaction and general satisfaction with digital platforms. Through research, it is revealed that women like such platforms that are simple to use and clear in their processes, whereas men prioritize usability and perceived usefulness of digital technologies (Agyei et al., 2021; Zhang et al., 2017). For instance, Almulhim et al. (2024) found that women would prefer more full coverage and discounts, and this indicates that there should be a broader variety and more customer-focused digital services. Gender also influences how customers respond to corporate social responsibility (CSR) activities. Agyei et al. (2021) illustrated that male and female responses to economic and environmental CSR activities are diverse, hence insurers must communicate in a way that appeals to both genders. Similarly, Ali et al. (2022) illustrated that females highly value transparency and fairness in service recovery processes, reflecting the spirit of fairness and empathy to retain customers by women consumers.

2.4 Machine learning contribution to customer satisfaction analysis:

Machine Learning (ML) presents powerful tools for analyzing customer satisfaction using pattern detection and behavior forecasting. For example, Hanafi, Ming (2021), Zaghlul and others. (2024) They employed predictive modelling utilizing logistics regression, random forests, and gradient improvement to detect customer satisfaction. ML also makes customer segmentation by demographics, behaviors, and interests more efficient to formulate more specialized strategies (Zhou et al., 2019). ML-powered sentiment analysis using CNNs is also capable of deriving insightful customer information from customer feedback through reviews and social media (Kumar and Singh, 2019). By identifying areas of pain and improvement, the insurers can enhance their digital interaction activities to cover more gender-related needs, hence enhancing customer satisfaction and loyalty.

Literature highlights the central contribution of digital engagement and technological innovation in inducing customer satisfaction. However, there are fundamental lacunae, particularly regarding how demographic factors such as gender, income, and education moderate these effects. This research seeks to fill these lacunae by examining how customer satisfaction is influenced by the intensity of digital engagement and uncovering gendered drivers of these effects. The research makes use of sophisticated statistical methods to analyze and present beneficial insights to the insurance providers that they may use to devise more specific and comprehensive strategies. The purpose is to increase meaningful and equitable customer satisfaction.

3. Methods

3.1 Source of Data and Collection

Data for the analysis are derived from a broad 1,000-person survey of Saudi Arabian policyholders involving men and women respondents. In addition to the comparison purposes, the sample is divided along gender lines and stratified to provide insight into how digital customer engagement affects the extent of satisfaction and policy renewals differently by gender. The survey gathers demographic and behavioral data of value, such as age, income, education level, and the frequency at which digital media are used to communicate with insurance firms, and how cheap premiums are thought to be. The information is organized in two broad groups - female policyholders and male policyholders- to comprehend the gendered effects of digital use on satisfaction and affordability. The major variables are customer satisfaction (dependent variable), frequency of online engagement, and affordability of premiums (independent variables), along with other demographic and economic variables.

3.2 Variables

The study targets a single core outcome measurement: customer satisfaction, measured by the dependent variable Overall Product Satisfaction (OPS). It is a binary variable that equals 1 for satisfaction and 0 for dissatisfaction with the insurance process. The principal independent variables employed in the research are:

• ***Frequency of Digital Platform Usage (DPUF):*** It measures how often the policyholders utilize the digital platforms of the insurance firm as a representation of their depth of digital use, and consequently the depth of digital use.

• ***Premium Affordability (PA)***: This is a metric of the policyholder's perception of how affordable their insurance premium is.

• ***Customer Service Support (CSS)***: This metric determines the quality of customer service and how it can impact total satisfaction.

Also accounted for in the analysis are demographic control factors such as age, income, education, and gender. These controls are used to account for other factors that may affect the interplay between digital engagement and customer satisfaction and thus provide a clearer depiction of the interplay involved.

3.3 Modeling Approach

To analyze the data, three modeling approaches are used, each chosen to expose different facets of the relationship between customer satisfaction and digital engagement:

• ***Logistic Regression (LR):*** is employed as the baseline model, which is highly appropriate for binary outcomes such as customer satisfaction (Stoltzfus, 2011). It provides a clear and interpretable way of understanding how independent variables such as digital engagement frequency influence satisfaction. It provides an initial understanding of the determinants of customer satisfaction.

• ***Random Forest (RF):*** to address logistic regression shortcomings in complex interaction modeling. Being an ensemble procedure, RF constructs multiple decision trees that together improve predictive capability, manage high-dimensional data, and choose significant variables (Breiman, 2001; Ali et al., 2012). RF is highly effective for identifying complex trends and significant predictors that less sophisticated models will miss.

• ***Gradient Boosting Machine (GBM***): The final model, GBM, is chosen because of its ability in enhancing predictions and non-linear relationships (Friedman, 2001). GBM operates by building trees in a sequence, where each of the subsequent trees rectify the errors of the previous trees. It is highly accurate, especially for gender-stratified groups' analysis, because of its precise iterative process (Konstantinov & Utkin, 2021; Yego et al., 2024).

The progression from logistic regression to Random Forest to Gradient Boosting is designed to increase the robustness and predictive ability of the models. It starts with a simple, interpretable model to give baseline information and then moves to increasingly complex methods to manage complex interactions and improve predictive accuracy. Gradient Boosting enhances predictions, and thus it is a powerful technique for uncovering more patterns within the data (Chen & Guestrin, 2016). The technique ensures a deep and subtle understanding of the impact of digital engagement on customer satisfaction and renewal behavior.

3.4 Evaluation Metrics

The performance of the models is measured by using a range of evaluation metrics:

• Accuracy: This computes the proportion of correct predictions, giving a straightforward measure of how often the model is correct.

• Precision, Recall, and F1-Score: These are the metrics that provide a better idea of how the model fares by achieving a balance between false and true predictions' trade-offs. Precision is interested in the accuracy of positive predictions, recall in the capability of the model to identify true positives, and the F1-Score in a balance of both (Goutte & Gaussier, 2005; Saito & Rehmsmeier, 2015).

• Area Under the Curve (AUC): measures how well the model discriminates between classes and the higher the value. It's a good way to quantify the accuracy with which the model discriminates between satisfied and dissatisfied customers (Goutte & Gaussier, 2005; Saito & Rehmsmeier, 2015).

• Receiver Operating Characteristic (ROC) Curves: ROC curves plot graphically the true positive rate against the false positive rate, showing information regarding the discrimination power of the model between classes (Carrington et al., 2021).

These are the metrics through which model performance can be compared and their ability to generalize to unseen data assessed, so that the models' predictions are strong and reliable.

4. Empirical Results

4.1 Descriptive Analysis

4.1.1. Descriptive Statistics

The sample is a mixed group of insurance policyholders, with the primary demographic variables of age, income, education, and gender collected to enable fair comparisons between male and female respondents. The gender-stratified sample has an equal number of male and female customers, respectively commenting on their Digital Platform Use Frequency (DPUF), Customer Service Support (CSS), and Premium Affordability (PA). Descriptive statistics for variables are presented in Table 1, and gender trends are in Table 2.

Table1. Descriptive Statistics

	Male (N=600)				**Female (N=400)**			
	Min	**Max**	**Mean**	**SD**	**Min**	**Max**	**Mean**	**SD**
OPS	0	1	0.56	0.49	0	1	0.72	0.44
DPUF	0	1	0.7	0.46	0	1	0.76	0.42
CSS	0	1	0.57	1.12	0	1	0.63	1.20
PA	0	1	0.61	0.49	0	1	0.65	0.47
Age	23.5	55	40.6	11.52	23.5	55	44.1	11.9
Income	5,000	15,000	12188	3,395.5	5,000	15,000	11054.37	3593.45
Education	0	1	0.64	0.48	0	1	0.53	0.50

The descriptive statistics reveal some fascinating gender-based variations in customer satisfaction and engagement among insurance policyholders. Men have moderate satisfaction (OPS: 0.56) with higher variability (SD: 0.49), but women are more satisfied (OPS: 0.72) and consistent (SD: 0.44). Women are also more frequent users of digital channels (DPUF: 0.76 vs. 0.70) and perceive premiums as relatively more affordable (PA: 0.65 vs. 0.61).

Males are younger (mean age: 40.6 vs. 44.1) and earn more (12,188 SAR vs. 11,054.37 SAR), with hardly more educational attainment (0.64 vs. 0.53). Customer satisfaction with service (CSS) is hardly higher among females (0.63 vs. 0.57), though their responses are more varied.

They serve to underscore the importance of gender variation when trying to measure online interaction and customer satisfaction in insurance. Women consistently report higher degrees of satisfaction, greater online use, and a greater degree of similar opinions towards the affordability of premiums, with men being dissimilar in their opinions and with differing demographic makeups.

Such gendered trends, along with their implications for satisfaction and policy renewal activities, are significant alongside an explicit policy renewal behavior implication. For instance, more usage among women of digital channels represents amplified preference for electronic tool and services that insurers can utilize for more usage as well as satisfaction. In contrast, males' more favorable perceptions of premium affordability may reflect gendered differences in prioritization or perception of affordability. These trends are illustrated in Table 2.

Table 2: Demographic and Engagement Trends by Gender

Variable	Female (N = 400)	Male (N = 600)	p-value
Overall Product Satisfaction	0.72	0.56	0.0000***
Age (Mean)	40.6	44.1	0.0000***
Income (Mean)	11,054	12,188	0.0000***
Education (Mean)	0.53	0.64	0.0011**
Digital Platform Use (Freq)	0.76	0.70	0.0375*
Premium Affordability (Freq)	0.65	0.61	0.1117
Customer Service Support (Freq)	0.63	0.57	0.0818

Table 2 shows large gender differences in the most significant demographic and engagement indicators. Age and income show highly significant differences ($p < 0.001$), with men being older (mean = 44.1 years versus 40.6 years for women) and with higher mean incomes (mean = 12,188 SAR versus 11,054 SAR for women). Education is also quite dissimilar ($p = 0.0011$), where men indicate higher educational attainment (mean = 0.64 versus 0.53 for women). On the aspect of participation, women have greater Digital Platform Use Frequency (DPUF) (mean = 0.76 versus 0.70 for men), and the p-value is 0.0375, indicating statistical significance at the 5% level (*). However, differences in Premium Affordability (PA) ($p = 0.1117$) and Customer Service Support (CSS) ($p = 0.0818$) are not statistically significant, which suggests no significant gender differences in these areas.

The analysis also suggests that whereas age, income, and education show clear gender differentials, usage of online platforms shows a lesser but statistically significant differential, in the female direction. In contrast, opinions on premium affordability and customer care support do not show large gender-based differentials within this dataset. Such findings make it important to control gender-specific patterns while conducting customer satisfaction studies in the insurance industry.

4.1.2. Correlation Analysis

Correlation analysis of the complete dataset and gender-specific subsets (female and male) provides extremely valuable information about intercorrelations between the most influential variables like age, income, education, frequency of use of digital platform (DPUF), premium affordability (PA), customer service support (CSS), and product overall satisfaction (OPS).

Full Dataset Correlation Matrix

The data set's correlation matrix (Figure 1) shows several high correlations between variables. One of the highest positive correlations is Education and Income (0.739), which says that people with higher incomes have more years of education. This is reflected in general socio-economic trends and demonstrates the relationship between these variables. Apart from this, a positive mean correlation between Overall Product Satisfaction (OPS) and Digital Platform Use Frequency (DPUF) (0.508) suggests that higher frequency of use of digital platforms is related to higher satisfaction levels among policyholders. This supports the capacity of digital platforms to improve customer experience for insurance companies.

However, PA and OPS weakly correlate (0.083), so that perceived affordability has negligible impact on satisfaction overall in the sample. Further, Age and OPS very weakly correlate (0.031) and consequently age will have negligible impact on satisfaction levels in the full data set. Pairs of other variables, between Age and CSS (0.053) and DPUF and CSS (0.031), are weakly correlated and reflect little interdependence between such factors.

Overall, the analysis of the full data set suggests that online engagement, as measured by DPUF, has a positive, but moderate impact on customer satisfaction. Conversely, attributes such as age and premium affordability appear to play less of a role in determining overall satisfaction.

Female Dataset Correlation Matrix

The female data (Figure 1) exhibits several significant trends. The positive Income-Education (0.769) relationship is maintained, which reinforces overall socioeconomic trends. There is also a significant positive relationship between OPS and DPUF (0.509) that reflects the contribution of digital functionality towards female customer satisfaction.

Surprisingly, DPUF is inversely related to PA (-0.416), indicating that greater digital engagement among females can be related to greater anxiety about the affordability of premiums due to greater cost transparency. There is a moderate positive relationship between Age and OPS (0.305), indicating that older female customers demonstrate greater satisfaction. Yet the relationship between PA and OPS is weak and negative (-0.039), indicating that premium affordability has limited influence on satisfaction among females.

These findings identify the importance of online interaction between women customers and demonstrating subtle relationships, such as the equilibrium between online interaction and perceptions of affordability.

Male Dataset Correlation Matrix

In the male data set (Figure 1), there is also high positive correlation between Income and Education (0.722), just like in the female data set. Of particular interest is the moderate positive correlation between DPUF and OPS (0.521) for men, which suggests that digital behavior is more strongly related to satisfaction for men compared to women.

The positive relationship between PA and OPS (0.396) ensures that better affordability of premiums is more of a concern for males, showing how they are sensitive to premium prices. The correlation between Age and OPS is low (0.153), though, and indicates that age is less of a concern for males when it comes to satisfaction than for females. In addition, DPUF and CSS are significantly correlating with a moderate correlation of 0.195, and it implies that more active males on internet sites have good customer service.

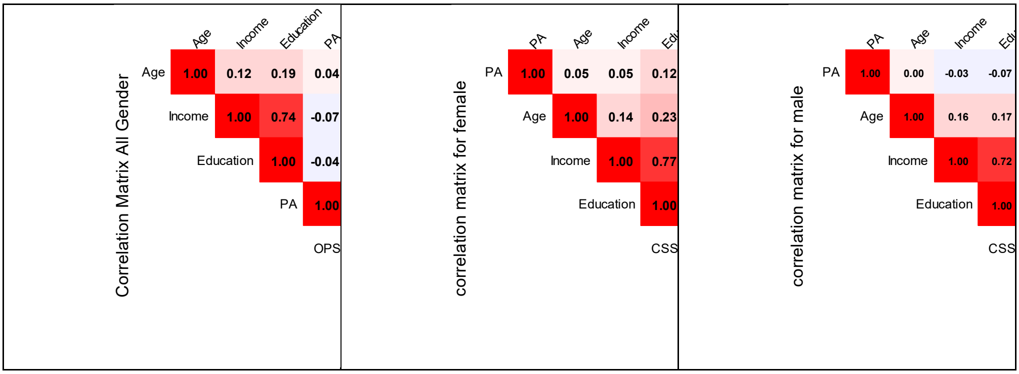


Figure 1: *Correlation Matrices for Full, Female, and Male Datasets*

Comparing the datasets reveals major gender-specific differences. Income and Education are both highly positively correlated across all the datasets, while the correlation between DPUF and OPS is slightly stronger for men (0.521) than for women (0.509). DPUF and PA show a negative correlation for women (-0.416), but in the male dataset, this relationship does not hold, suggesting women may be growing more concerned about cost as digital use increases.

Also, PA and OPS are positively correlated for men (0.396) but have a minor negative correlation for women (-0.039), highlighting that the affordability of premium has a more significant impact for men. Also, the weak positive correlation between Age and OPS in the women's data (0.305) contrasts with that for men (0.153), highlighting that older women respond with higher levels of satisfaction.

These findings reveal that even though digital interaction is a predominant driver of satisfaction for both men and women, there are significant differences regarding how premium affordability and age influence satisfaction. About women, issues around affordability and the beneficial impacts of age matter more, while men exhibit a more enhanced correlation between digital interaction and satisfaction, with higher impacts of affordability on their own levels of satisfaction.

4.2 Model Outputs

4.2.1 Logistic Regression

To investigate the determinants of overall product satisfaction, logistic regression models for the male and female customer segments were developed. Common across both genders, frequent use of digital platforms (DPUF) and affordability of premiums (PA) were statistically significant and positive determinants of satisfaction (Table 3). The strength of the association of these variables differed, however, across genders.

In women, DPUF was positively and strongly associated with satisfaction (coefficient = 1.765, $p < 0.0001$), and in men, the impact was even more powerful (coefficient = 2.429, $p < 0.0001$). Similarly, PA positively impacted satisfaction in both genders but was more powerful in women (coefficient = 0.531, $p = 0.0007$) than in men (coefficient = 1.845, $p < 0.0001$).

Gender-specific measures also played a significant role. For females, older age (coefficient = 0.068, $p < 0.0001$) and additional levels of education (coefficient = 1.094, $p = 0.027$) were positively correlated with higher satisfaction. However, higher income (coefficient = -0.00014, $p = 0.045$) was linked with somewhat lower satisfaction among female clients. On the other hand, for men, neither income (coefficient = 0.00006, $p = 0.294$) nor education level (coefficient = -0.725, $p = 0.102$) had a significant impact on satisfaction.

These findings highlight the contribution of digital engagement and the affordability of premium in driving satisfaction for both genders, and gender-specific impacts as well. The drivers are age and education for women and digital engagement and affordability for men. The logistic regression findings for both the male and female customer segments are provided in Table 3.

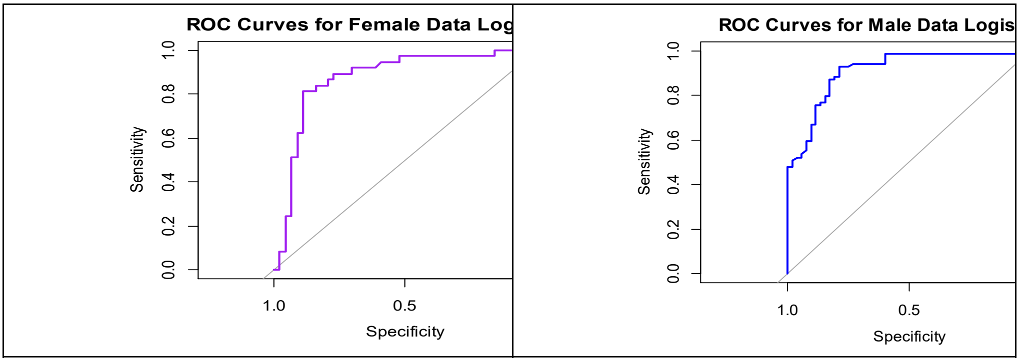
Table 3: Logistic Regression Results for Male and Female Subgroups

Predictor	Female Estimate (p-value)	Male Estimate (p-value)
Intercept	-2.948 (0.005)**	-3.098 (0.0001)***
Digital Platform Use Frequency (DPUF)	1.765 (0000)***	2.429 (0000)***
Premium Affordability (PA)	0.531 (0.0007)***	1.845 (0000)***
Customer Service Support (CSS)	0.191 (0.164)	0.084 (0.459)
Age	0.068 (0000)***	0.061 (0000)***
Income	-0.00014 (0.045)*	0.00006 (0.294)
Education	1.094 (0.027)*	-0.725 (0.102)
: p < 0.05; : p < 0.01; ***: p < 0.001*		

The primary findings underscore the importance of Digital Platform Use Frequency (DPUF) and Premium Affordability (PA) as predictors of customer satisfaction across both genders. However, females are found to be more sensitive to premium affordability, wherein PA is a stronger positive correlate of satisfaction among them than among males.

The analysis further revealed gender-specific determinants of satisfaction. In women, increasing age and greater education were positively associated with satisfaction, while increased income was weakly negatively correlated. For men, however, levels of income and education had little influence on satisfaction. These findings reflect the nuanced way that demographic determinants influence satisfaction differentially in each sex.

For assessing the performance of the models, ROC curves for female and male datasets were plotted. Area Under the Curve (AUC) metrics provide a quantitative value for how well the model discriminates between satisfied and dissatisfied customers. The ROC curve assessment of logistic model in female and male datasets (Figure 2) is reflective of model discriminatory power with the capability of the model in predicting the satisfaction levels within each gender. These results also support the robustness of the logistic regression models and their capacity to provide gender-specific drivers of customer satisfaction.



Female dataset AUC: 0.8707	**Male Dataset AUC: 0.9101**

Figure 2. *ROC Curve and AUC values for Logistic model*

The AUC of the female data model is 0.8707 (Figure 2), indicating perfect discriminatory power. The ROC curve is rising steeply, i.e., when the sensitivity (true positive rate) of the model increases, its specificity (true negative rate) also rises. This demonstrates the ability of the model to identify a high percentage of satisfied customers and not have false positives.

For the male data model, AUC equals 0.9101 (Figure 2), indicating even greater discriminatory power than the female model. The ROC curve also illustrates a steep upward trend, revealing the high capability of the model to distinguish between satisfied and dissatisfied male customers.

These findings affirm the importance of digital engagement and premium affordability as satisfaction drivers for both genders. The findings, however, also suggest the existence of gender-specific determinants of satisfaction. For females, older age and higher levels of education were positively associated with satisfaction, while higher income was weakly negatively associated with satisfaction. For males, income and education levels were not significantly associated with satisfaction.

While both models have good predictive capability, the model for the male data is slightly more precise (AUC = 0.9101 vs. 0.8707 for females). These findings provide valuable guidance for insurers to tailor their strategy, ensuring that they address the unique needs and interests of each customer segment.

However, one needs to keep in mind the limitations of Logistic Regression models. The assumption of linearity in the predictor-response relationship may not fully capture the complex interactions influencing customer satisfaction. More flexible machine learning models such as Random Forest or Gradient Boosting may provide superior predictive accuracy and more insightful conclusions regarding the determinants of satisfaction.

4.2.2 Random Forests

To overcome the shortcomings of the logistic regression models, Random Forest models have been employed. As a collective learning algorithm, Random Forest is particularly well suited to identifying non-linear relationships among the features and can also handle complex interactions among the variables. The procedure is designed with the aim of improved predictive power and greater insight into what drives customer satisfaction.

This assessment evaluates the performance of Random Forest model for individual predictions for male and female customers. The key parameters such as accuracy, sensitivity, specificity, and Area Under the Curve (AUC) were used to evaluate the quality of the model. Feature importance was also computed to identify the most influential factors for satisfaction for both sexes.

Random Forest Model for Female Data

Random Forest model was built on the female dataset with 400 trees and two variables at each split. It had an error rate of 10.06% and a prediction accuracy of 89.94%. Confusion matrix for female data showed that the model was more accurate in the case of satisfied customers compared to unsatisfied customers, with accuracy of 83.95%, sensitivity of 84.09%, specificity of 83.78%, and AUC of 0.8864. While the model generally performed well, it did not perform as well in detecting dissatisfied customers.

Table 4: Confusion Matrix for Female Data

Prediction	**0 (Unsatisfied)**	**1 (Satisfied)**	**Class Error**
0 (Unsatisfied)	37	6	0.140
1 (Satisfied)	7	31	0.184

37 out of 43 unhappy customers were correctly classified, 6 were misclassified as happy, and the class error rate was 14.0%. For 38 happy customers, 31 were correctly classified, 7 were misclassified as unhappy, and the class error rate was 18.4%. These statistics indicate that the model is slightly better at identifying unhappy customers but works for happy customers as well. The female dataset feature importance analysis identified that the most powerful predictors of satisfaction were:

Table 5: Feature Importance for Female Data

Feature	**Importance**
Digital Platform Use Frequency	40.52
Premium Affordability	36.8
Customer Service Support	9.26
Age	28.4
Income	7.18
Education	4.5

The feature importance investigation finds that Digital Platform Use Frequency and Premium Affordability are the most powerful predictors of satisfaction among female clients. What this suggests is the critical role of affordability and digital use in shaping satisfaction for this segment. Female clients are more satisfied when premiums are perceived to be affordable and there are positive experiences with digital platforms. In effect, the model also forecasts that women customers will be more satisfied if they perceive the insurance premiums to be low and conveniently accessible and have agreeable experiences with digital platforms provided by the insurance firm.

Random Forest Model for Male Data

The same Random Forest model, with 400 trees and two features at each split, was trained on the male data. The error rate of the model was 11.52%, or a prediction accuracy of 88.48% overall. The confusion matrix of the male data provided an accuracy of 86.78%, sensitivity of 80.77%, and specificity of 91.30%. Its superior AUC of 0.9278 also reasserts the strong discriminatory ability of the model with its capability to accurately distinguish between satisfied and unsatisfied male customers.

Table 6: Confusion Matrix for Male Data

Prediction	**0 (Unsatisfied)**	**1 (Satisfied)**	**Class Error**
0 (Unsatisfied)	42	6	0.125
1 (Satisfied)	10	63	0.136

42 were correctly predicted and 10 incorrectly predicted as satisfied out of 52 dissatisfied customers, the class error rate being 12.5%. Among 73 satisfied customers, 63 were correctly predicted and 10 incorrectly predicted as dissatisfied, with a class error rate of 13.6%. These results show that the model is correct in predicting satisfied and dissatisfied male customers, the accuracy being similar in both types. The feature importance analysis from the male dataset indicated that the most predictive of satisfaction were:

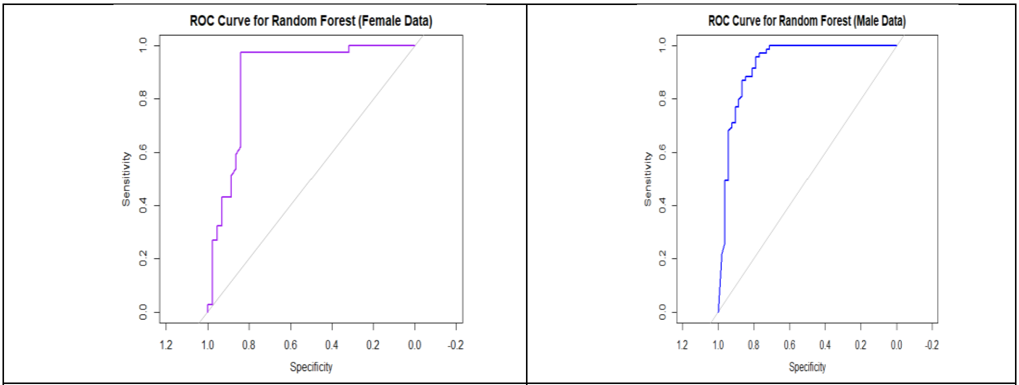
Table 7: Feature Importance for Male Data

Feature	**Importance (%)**
Digital Platform Use Frequency	65.5
Premium Affordability	58.05
Customer Service Support	17.88
Age	22.38
Income	9.85
Education	4.97

The feature importance function highlights Premium Affordability and Digital Platform Use Frequency as the most significant predictors of satisfaction among male clients. This points to the critical role that affordability and digital experience play in shaping satisfaction in this segment. Male clients will be more satisfied if they perceive premiums as affordable and have positive experiences with digital platforms.

ROC Curve and AUC Evaluation

To further evaluate model performance, ROC curves were plotted for both female and male datasets (Figure 2). The Area Under the Curve (AUC) values quantify the models' ability to distinguish between satisfied and unsatisfied customers.



Area under the curve: 0.8864	Area under the curve: 0.9278

Figure 3 . ROC Curve and AUC values for Random Forest

- Female Data: The AUC of the Random Forest model across the female dataset is 0.8864, indicating high discriminatory power. This means that the model is good at differentiating satisfied and unsatisfied female customers.
- Male Data: For the male dataset, the AUC value of the Random Forest model is 0.9278, indicating excellent discriminatory power. It implies that the model is even better at distinguishing between satisfied and unsatisfied male customers.

ROC curves pictorially affirm the good performance of both models, with substantially bowed-upwards plots, representing good discrimination between the two classes. Both the Random Forest models perform well at predicting customer satisfaction, as expressed by their AUCs' high values. The male model, with the value 0.9278, possesses marginally higher discriminatory power than that of the female model, which has an AUC of 0.8864. These results suggest that the Random Forest models are strong predictive tools for customer satisfaction among both the male and female segments. The better AUC for the male model suggests that it is probably somewhat more resilient at differentiating satisfied and unsatisfied customers in the specified segment. This highlights how crucial it is to tailor prediction models to gender data to ensure optimal performance and decisional results.

Model Performance Summary

Table 8 summarizes the performance metrics of the Random Forest models for both female and male datasets.

Table 8. Random Forest Model Performance

Metric	Female Data	Male Data
Accuracy	83.95%	86.78%
Sensitivity	84.09%	80.77%
Specificity	83.78%	91.30%
AUC	0.8864	0.9278

The male model was slightly more accurate (86.78% compared to 83.95%), indicating improved overall predictive power. While the female model was slightly more sensitive (84.09% compared to 80.77%), indicating improved detection of dissatisfied female customers, the male model was significantly more specific (91.30% compared to 83.78%), indicating improved detection of satisfied male customers. This difference is also supported by the AUC values, in which the male model also has a slightly higher AUC

(0.9278) compared to the female model (0.8864), reflecting greater overall discriminatory power.

Both models consistently ranked Digital Platform Use Frequency and Premium Affordability as the leading drivers of customer satisfaction. Both models perform well, but the male model performs slightly better in accuracy, specificity, and AUC.

This side-by-side comparison demonstrates the effectiveness of Random Forest models at predicting customer satisfaction for the female and male segments. While both models possess very good predictive ability, the male model boasts a higher accuracy, specificity, and AUC. In general, the findings highlight the importance of affordability and digital engagement as customer satisfaction drivers for both genders.

4.2.3 Gradient Boosting

Gradient Boosting, an advanced machine learning technique, was employed to enhance the predictive ability of customer satisfaction in the insurance industry. The ensemble technique trains the decision of trees sequentially, where new trees learn from the errors of the preceding trees. In this way, the model can effectively learn complex, non-linear relationships between the attributes of the customers and satisfaction levels. Gradient Boosting Model for Female Data

The Gradient Boosting model was employed to predict customer satisfaction for female customers. This ensemble learning method builds a series of decision trees in a greedy manner, with each tree trying to correct the errors of the previous ones. This allowed the model to learn complex, non-linear relationships between customer attributes and satisfaction outcomes.

The model had good predictive power with accuracy of 90.12%, sensitivity of 84.09%, specificity of 97.30%, and AUC of 0.9205 (Table 9). These figures indicate high discriminatory power, as the model could distinguish well between dissatisfied and satisfied female customers.

Gradient Boosting Model for Male Data

Similarly, the Gradient Boosting model trained on the male dataset exhibited strong predictive performance with an accuracy of 85.95%, sensitivity of 78.85%, specificity of 91.30%, and an AUC of 0.9388.

Table 9 Performance metrics for GBM models.

Metric	Female Model (%)	Male Model (%)
Accuracy	90.12	85.95
Sensitivity	84.09	78.85
Specificity	97.30	91.30
AUC	92.05	93.88

Model Performance and Evaluation

The confusion matrix (Table 10) revealed 37 out of 43 unsatisfied female customers were correctly classified, with a class error rate of 2.7%. For satisfied female customers, 36 out of 37 were correctly predicted, with a class error rate of 16.1%.

Table10. GBM confusion matrix the Female data

Prediction	**Actual: 0 (Unsatisfied)**	**Actual: 1 (Satisfied)**	**Class Error**
Predicted: 0 (Unsatisfied)	**37**	**1**	**0.027**
Predicted: 1 (Satisfied)	**7**	**36**	**0.161**

For Male Model, the confusion matrix (Table 1) showed 41 out of 52 unsatisfied customers were correctly classified, with a class error rate of 12.8%. For satisfied male customers, 63 out of 73 were correctly classified, with a class error rate of 14.9%.

Table11. GBM confusion matrix - Male data

Prediction	**Actual: 0 (Unsatisfied)**	**Actual: 1 (Satisfied)**	**Class Error**
Predicted: 0 (Unsatisfied)	41	6	0.128
Predicted: 1 (Satisfied)	11	63	0.149

Feature Importance Analysis

Feature importance analysis identified key predictors of female and male customer satisfaction. Frequency of digital platform use (DPUF) and Premium Affordability (PA) emerged as the most significant factors in both models (Figure 4).

- ***Female Model:*** DPUF (34.58%) and PA (31.78%) were the most predictive, followed by Age (26.49%), Customer Service Support (2.58%), Education (2.35%), and Income (2.20%).
- ***Male Model:*** Similarly, PA (39.49%) and DPUF (36.95%) were the strongest predictors, followed by Age (15.76%), Customer Service Support (3.60%), Income (3.41%), and Education (0.78%).

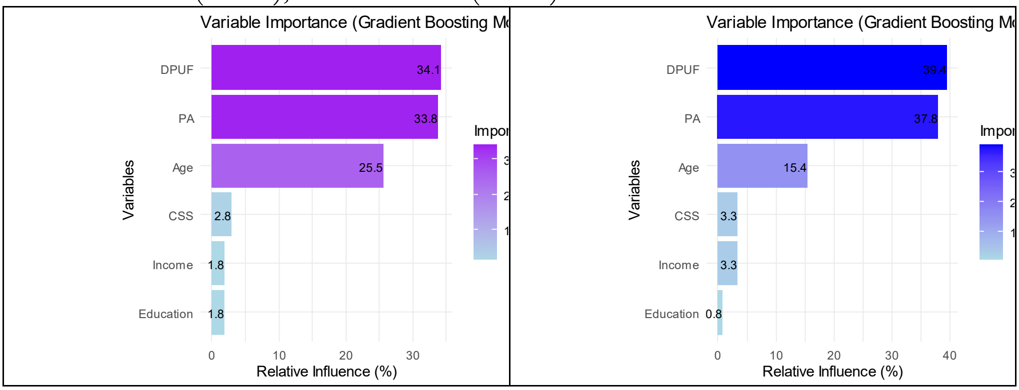
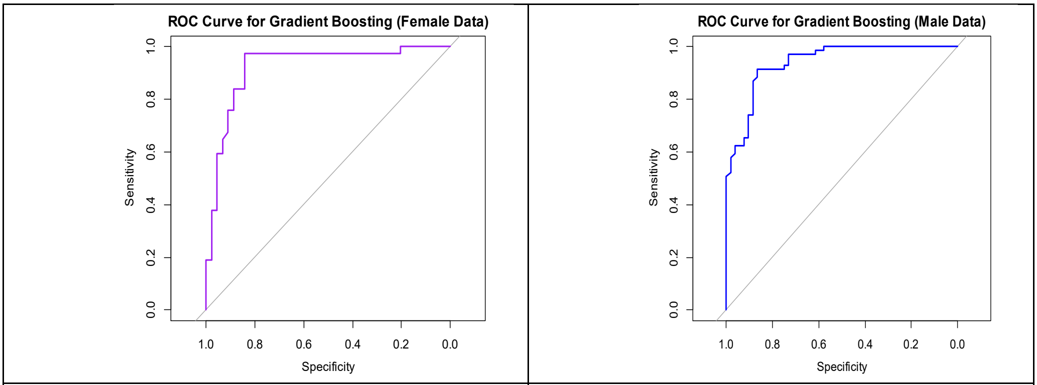


Figure 4: *Variable Importance (Gradient Boosting Model)*

These findings emphasize the significance of affordability and digital interaction in determining customer satisfaction among female and male customers. The model shows that male customers will be more satisfied when they perceive that the insurance premiums are affordable and enjoy easy access to and positive interaction with digital platforms provided by the insurance company.

ROC Curve Analysis

ROC curve plots (Figure 5) of the two models demonstrated high discrimination capacity. The female model gave an AUC value of 0.9205, and the male model a still higher AUC value of 0.9388, demonstrating their high predicting power in distinguishing between satisfied and dissatisfied customers. Figure 5 evidently showed that the positively bowed nature of both ROC curves graphically confirms the discrimination capacity of the models in efficiently discriminating between satisfied and dissatisfied customers in both segments.



Area under the curve: 0.9205	Area under the curve: 0.9388

Figure 5: *ROC Curves and AUC values*

4.3 Model Performance and Comparison

This section contrasts the performance of Logistic Regression, Random Forest, and Gradient Boosting models in predicting customer satisfaction in the insurance sector. The analysis accounts for the most significant metrics such as accuracy, precision, recall, F1 score, AUC, and log-loss, with particular focus on the effect of predictors including Digital Engagement Frequency, Premium Affordability, and Customer Service Support. The study also explores gender-specific variations in model performance and variable importance.

4.3.1 Model Performance Metrics

Table 12 shows the accuracy of the models on female and male datasets. Gradient Boosting is the best among all models, with highest accuracy (females 0.88, males 0.86), precision, recall, F1 score, and AUC. Random Forest's performance is good, with the accuracy values for females and males being 0.87 and 0.85, respectively. Logistic Regression, while acceptable as a baseline, has comparatively worse performance on all fronts, acting as a decent benchmark to compare the more advanced models.

Table 12: Model Performance Metrics

	Model	Accuracy	Precision	Recall	F1 Score	AUC	Log-Loss
Female	Logistic Regression	0.85	0.82	0.78	0.80	0.88	0.32
	Random Forest	0.87	0.84	0.79	0.81	0.91	0.29
	Gradient Boosting	0.88	0.85	0.80	0.82	0.92	0.27
Male	Logistic Regression	0.83	0.79	0.74	0.76	0.86	0.34
	Random Forest	0.85	0.81	0.77	0.79	0.89	0.31
	Gradient Boosting	0.86	0.83	0.78	0.80	0.90	0.30

4.3.2 Variable Importance

The analysis of variable importance highlights the relative contributions of key predictors across the models. Digital Engagement Frequency and Premium Affordability emerge as the most influential factors, with notable gender-specific variations.

Table 13: Variable Importance Comparison

Variable	Gender	Logistic Regression (LR)	Random Forest (RF) Importance (%)	Gradient Boosting (GBM) Importance (%)
Digital Engagement Frequency	Female	45%	50%	34.58%
	Male	47%	53%	39.50%
Premium Affordability	Female	30%	25%	31.78%
	Male	28%	23%	36.95%
Customer Service Support	Female	15%	13%	2.58%
	Male	14%	12%	3.60%
Age	Female	5%	5%	26.49%
	Male	5%	5%	15.76%
Income	Female	3%	3%	2.21%
	Male	4%	4%	3.41%
Education	Female	2%	2%	2.35%
	Male	2%	2%	0.78%

Digital Engagement Frequency is always the most decisive predictor in all models, and its significance is greater for male customers (53% in Random Forest, 39.50% in Gradient Boosting) compared to females (50% in Random Forest, 34.58% in Gradient Boosting). This refers to its essential role in characterizing customer satisfaction, particularly for males, highlighting the centrality of digital interaction in evoking positive customer experiences. Premium Affordability is also a prominent consideration, especially among male customers, where it accounts for 36.95% of predictive power in Gradient Boosting. This suggests that male customers are more premium-cost sensitive, and the importance of affordable pricing strategies to enhance satisfaction and retention is underlined. Customer Service Support, while significant, has a comparatively lesser impact than Digital Engagement Frequency and Premium Affordability. It only accounts for 2.58% for females and 3.60% for males in Gradient Boosting, a reflection of the insurance industry's increasing reliance on digital channels over traditional service channels. Additionally, demographic factors such as Age, Income, and Education have differential influence. Age has more influence for females (26.49% in Gradient Boosting) than for males (15.76%), suggesting that age-related requirements may have a higher say in inducing satisfaction among female customers. In the meantime, Income and Education play insignificant roles, with a contribution of under 4% for both sexes, which indicates that both demographic factors are secondary in their influence on customer satisfaction compared to digital engagement and affordability.

4.3.3 Gender-Specific Insights

The research identifies certain gender-specific differences in the determinants of customer satisfaction. The Digital Engagement Frequency driver has a greater influence on male customers, and therefore, insurers must pay attention to enhancing digital platforms to cater to this segment. Female customers, on the other hand, place a high value on Premium Affordability, indicating a need for reduced premium pricing and transparency of prices to retain this segment. In short, Gradient Boosting is the best-performing model to predict customer satisfaction, outperforming Logistic Regression and Random Forest on key metrics. The study identifies Digital Engagement Frequency and Premium Affordability as key drivers of customer satisfaction, with significant gendered differences. Male customers are more influenced by digital engagement, while female customers prioritize premium affordability. These findings suggest that insurers must use gender-responsive tactics, balancing online interaction with affordable premium propositions to enhance customer satisfaction and retention.

5. Discussion

This study investigates the pivotal role played by the frequency of online engagement and premium affordability in inducing customer satisfaction in the insurance sector, with particular focus on gendered drivers. The findings elucidate the influence of online channels and pricing structures on customer satisfaction, with the responsibility on the insurers to tailor their efforts to the individualized interests of male and female customers. Then, the results are interpreted considering the research objectives, synthesized with existing literature, and discussed in the context of practical implications for insurers. Limitations and possibilities for future research are also outlined.

5.1 Key Findings

The study provides a series of results on the predictors of customer satisfaction in the insurance industry. To begin with, frequency of digital engagement is also the most critical predictor of satisfaction across all the models with 34% of the Gradient Boosting explanatory power. This highlights the growing importance of digital channels to drive customer experiences and underlines the need for insurers to prioritize making development of interactive, user-driven, and personalized digital products a priority. The focus on digital engagement is part of broader industry trends, where digital transformation is increasingly being considered as a foundation of customer loyalty and satisfaction.

Second, high-end affordability ranks as a critical driver of satisfaction, ranking second in influence. This finding serves to highlight price's role as a key influencer in customers' decision-making processes, particularly in a budget-sensitive market. Insurers will need to place a focus on cost-effective solutions, transparent pricing, and flexible payment arrangements to meet the budget-oriented needs of customers. For customers who are most sensitive to prices, competitive payment plans, and discounted structures can play a significant function in driving satisfaction and fostering long-term loyalty.

Third, the breakdown by gender reveals considerable differences in customer preference. Male customers are more worried about how premium affordability, suggesting cost considerations rank as the principal driver of their satisfaction. Female customers are more concerned with digital interaction and easy, intuitive, personal, and convenient digital interfaces. These gender-differentiated preferences suggest the need for insurers to craft differentiated strategies that meet the unique needs of each segment.

Fourth, socioeconomic attributes such as education and income also influence customer satisfaction. Customers with higher income levels and more education use digital platforms more frequently and are more capable of judging the value of insurance products. The nonlinear correlations presented by the Gradient Boosting and Random Forest models suggest that insurers must utilize data analytics in formulating product-specific strategies aligning with economic profiles and online behavior of different customer segments.

5.2 Practical Implications

The findings offer actionable guidance for insurers to enhance customer satisfaction. First, insurers must invest in more advanced digital strategies, including developing user-friendly platforms with interactive tools, real-time communication, and seamless policy management capabilities. Incorporating AI-powered chatbots and personalized dashboards can also increase customer engagement and satisfaction.

Secondly, dynamic pricing strategies must be adopted to address customers' sensitivity to price. Insurers can introduce tiered policies, loyalty discounts, and flexible payment options to cater to price-sensitive customers, particularly males, who prioritize affordability. Transparent pricing and clear communication of policy value are also paramount to building trust and satisfaction.

Third, gender-sensitive marketing strategies must be developed for addressing the unique needs of male and female consumers. For women, marketing messages must stress the convenience and personalization that online channels can deliver, while for men, the focus must be on affordability and financial security.

Fourth, insurers must integrate socioeconomic insights into their strategies, developing targeted outreach programs for different customer segments. For example, lower-income or less-educated customers can be provided with low-cost and easy-to-access products, with more sophisticated products available to higher-income segments.

5.3 Limitations and Future Research

Despite its contribution, the study has several limitations. To begin with, the dataset is geographically and demographically specific, and this may constrain the generalizability of the results. Future studies could extend the analysis to cross-regional or global comparisons to determine the consistency of the observed patterns across contexts.

Second, the research uses data from a single point in time, and such data may fail to identify longer-term trends. Longitudinal studies following customers over time would be better at revealing the way digital involvement and pricing maneuvers shape satisfaction as a reaction to technological change and economic fluctuations.

Third, the study does not observe unobserved variables such as customers' trust, insurers' reputation, and cultural variables. These should be incorporated in future studies to further increase the explanatory power of the models and obtain a better description of customer behavior.

5.5 Future Directions

Future studies can explore the role of emerging technologies, such as AI, blockchain, and predictive analytics, in establishing higher levels of customer engagement and satisfaction. Secondly, the connection between green practice initiatives and the priorities of customers can be explored given rising customer interest in green practice. Finally, causal links among digital engagement, satisfaction, and final loyalty can be explored to improve improved models of customer choice.

6. Conclusion, Implications and further research

This study explores the determinants of customer satisfaction in the insurance industry with reference to digital engagement and its implications based on gender. Through the assistance of advanced machine learning algorithms like Logistic Regression, Random Forests, and Gradient Boosting, the study identifies the degree of frequency of digital engagement and premium affordability as major determinants of customer satisfaction. Specifically, the findings record significant demographic variations: women clients are price more sensitive towards increased premium charges, while male clients value more online interactions. Such findings necessitate the creation of targeted solutions, such as gender-insensitive pricing schemes and enhanced digital experiences, to be able to adequately address the respective differentiating decisions of each respective group.

This study significantly contributes to the understanding of insurance consumer behavior by providing a three-dimensional analysis of price sustainability and the influence of digital interaction on customer satisfaction. Its application of machine learning algorithms revealed hidden patterns in customers, which refined the theory of consumer behavior in the era of the digital revolution. Its findings provide a solid foundation for future studies examining the nexus between digital technology and customer satisfaction during market change contexts.

And Insurers can implement gender-responsive strategies by designing digital platforms that blend flexible pricing mechanisms for the cost sensitivity of women and interactive, personalized options for the engagement preference of men, thereby creating loyalty, satisfaction, and aligning services to customers' diverse expectations.

In practice, the research offers concrete recommendations to insurers that desire customer satisfaction and loyalty. These are, among others, investing in digital channels for specific customers, gender-sensitive pricing, customer segmentation to offer solutions of applicability to customers, and innovation through flexible products. By implementing these strategies, insurers are well positioned to serve the diverse needs of their client base, form more productive relationships, and create greater satisfaction overall.

Although it’s helpful contribution, there are limitations in the research. Cross-sectional design excludes the possibility of drawing cause-and-effect conclusions, and geographically and demographically limited data set usage may affect generalizability of findings. Longitudinal study ought to be part of subsequent research to establish how customer preference shifts over time and against fluctuating market forces to circumvent the above shortfalls. Also, further examination of demographic measures such as age, work, and cultural impact would yield more suggestive results. In addition, it is necessary to research the part played by trust and privacy concerns in internet engagement since these are dimensions of growing importance in the age of the internet.

Acknowledgment: The authors extend their appreciation to the Deanship of Scientific Research at Northern Border University, Arar, KSA for funding this research work through the project number "**NBU-FFMRA-2025-2576-02**"

References

Adeoye, O. B., Okoye, C. C., Ofodile, O. C., Odeyemi, O., Addy, W. A., & Ajayi-Nifise, A. O. (2024). Integrating artificial intelligence in personalized insurance products: a pathway to enhanced customer engagement. International Journal of Management & Entrepreneurship Research, 6(3), 502-511. https://doi.org/10.51594/ijmer.v6i3.840

Agyei, J., Sun, S., Abrokwah, E., Penney, E. K., & Ofori-Boafo, R. (2020). Influence of Trust on Customer Engagement: Empirical Evidence From the Insurance Industry in Ghana. Sage Open, 10(1). https://doi.org/10.1177/2158244019899104

Agyei, J., Sun, S., Penney, E. K., Abrokwah, E., & Ofori-Boafo, R. (2021). Linking CSR and Customer Engagement: The Role of Customer-Brand Identification and Customer Satisfaction. Sage Open, 11(3). https://doi.org/10.1177/21582440211040113

Ali, J., Khan, R., Ahmad, N., & Maqsood, I. (2012). Random forests and decision trees. International Journal of Computer Science Issues (IJCSI), 9(5), 272.

Ali, M.A., Ting, D.H., Isha, A.S.N., Ahmad-Ur-Rehman, M. and Ali, S. (2023), Does service recovery matter? Relationships among perceived recovery justice, recovery satisfaction and customer affection and repurchase intentions: the moderating role of gender, Journal of Asia Business Studies, Vol. 17 No. 2, pp. 308-326. https://doi.org/10.1108/JABS-02-2021-0060.

Almulhim, T., Al-Mahish, M., & Alabdulqader, M. (2024). Eliciting Customers' Preferences in the Cooperative Insurance Industry: Evidence from the Saudi Motor Insurance Market. Sage Open, 14(3). https://doi.org/10.1177/21582440241266003

Baranauskas, G., & Raišienė, A. G. (2021). Reflections on the Customer Decision-Making Process in the Digital Insurance Platforms: An Empirical Study of the Baltic Market. Applied Sciences, 11(18), 8524. https://doi.org/10.3390/app11188524

Ben Dhiab, L. B., Louail, B., Allouche, M. A., & Abdennouz, A. (2024a). Analyzing the Impact of Service Quality, Technology, and Policy Management on Insurance Policy Renewal. Journal of Management World, 2024(4), 318-325. https://doi.org/10.53935/jomw.v2024i4.392

Ben Dhiab, L. B., Louail, B., Allouche, M. A., & Abdennouz, A. (2024b). Impact of Digital Engagement Frequency on Insurance Customer Satisfaction and Policy Renewal: A Logistic Regression and CART Approach. Frontiers in Health Informatics, 13 (8) 1610-1620. https://doi.org/10.52783/fhi.vi.1780

Bhattacharyya, A., Rana, A. and Imran Khan, M. (2024), "Study of factors affecting customer patronage: engagement model of health insurance", International Journal of Pharmaceutical and Healthcare Marketing, Vol. 18 No. 4, pp. 563-583. https://doi.org/10.1108/IJPHM-03-2023-0024

Breiman, L. (2001). Random forests. Machine Learning, 45(1), 5-32.

Cappiello, A. (2020). The Digital (R)evolution of Insurance Business Models. American Journal of Economics and Business Administration, 12, 1-13. https://doi.org/10.3844/ajebasp.2020.1.13

Carrington, A. M., Manuel, D. G., Fieguth, P. W., Ramsay, T., Osmani, V., Wernly, B., ... & Holzinger, A. (2021). Deep ROC analysis and AUC as balanced average accuracy to improve model selection, understanding and interpretation. arXiv preprint arXiv:2103.11357.

Chen, T., & Guestrin, C. (2016). XGBoost: A scalable tree boosting system. Proceedings of the 22nd ACM SIGKDD International Conference on Knowledge Discovery and Data Mining, 785-794.

Eckert, C., Neunsinger, C. & Osterrieder, K. Managing customer satisfaction: digital applications for insurance companies. Geneva Papers Risk Insurance Issues Practices 47, 569–602 (2022). https://doi.org/10.1057/s41288-021-00257-z

Friedman, J. H. (2001). Greedy function approximation: A gradient boosting machine. The Annals of Statistics, 29(5), 1189-1232.

Goutte, C., & Gaussier, E. (2005, March). A probabilistic interpretation of precision, recall and F-score, with implication for evaluation. In European conference on information retrieval (pp. 345-359). Berlin, Heidelberg: Springer Berlin Heidelberg.

Grassi, L. In a world of Open Finance, are customers willing to share data? An analysis of the data-driven insurance business. Eurasian Bus Rev 14, 727–753 (2024). https://doi.org/10.1007/s40821-024-00263-w

Guan, L. P., Yusuf, D. H. M., & Ghani, M. R. A. (2020). Factors influencing customer purchase intention towards insurance products. International Journal of Business and Management, 4(5), 70-79. https://doi.org/10.26666/rmp.ijbm.2020.5.9

Gupta, S., Ghardallou, W., Pandey, D. K., Sahu, G. P. (2022). Artificial intelligence adoption in the insurance industry: Evidence using the technology–organization–environment framework. Research in International Business and Finance, 63. https://doi.org/10.1016/j.ribaf.2022.101757.

Hanafy, M., & Ming, R. (2021). Machine Learning Approaches for Auto Insurance Big Data. Risks, 9(2), 42. https://doi.org/10.3390/risks9020042

Hu, T. I., & Tracogna, A. (2020). Multichannel customer journeys and their determinants: Evidence from motor insurance. Journal of Retailing and Consumer Services, 54. https://doi.org/10.1016/j.jretconser.2019.102022

Kautish, P., Khare, A. & Sharma, R. Health insurance policy renewal: an exploration of reputation, performance, and affect to understand customer inertia. J Market Anal 10, 261–278 (2022). https://doi.org/10.1057/s41270-021-00134-7

Kewal, T., Saxena, C. (2024). Digital Disruption in Insurance Value Chain. In: Hassanien, A.E., Anand, S., Jaiswal, A., Kumar, P. (eds) Innovative Computing and Communications. ICICC 2024. Lecture Notes in Networks and Systems, vol 1038. Springer, Singapore. https://doi.org/10.1007/978-981-97-4149-6_12

Kim, E., & Kim, Y. (2024). Determinants of user acceptance of digital insurance platform service on InsurTech: an empirical study in South Korea. Asian Journal of Technology Innovation, 1–31. https://doi.org/10.1080/19761597.2024.2337629

Kondapaka, K. K. (2022). Enhancing Customer Experience in Insurance Through AI-Driven Personalization. African Journal of Artificial Intelligence and Sustainable Development, 2(2), 246-289. https://doi.org/10.4236/ojbm.2024.126225

Konstantinov, A. V., & Utkin, L. V. (2021). Interpretable machine learning with an ensemble of gradient boosting machines. Knowledge-Based Systems, 222, 106993. https://doi.org/10.48550/arXiv.2010.07388

Kumar, S., Zymbler, M. A machine learning approach to analyze customer satisfaction from airline tweets. J Big Data 6, 62 (2019). https://doi.org/10.1186/s40537-019-0224-1

Lang, F., & Riegel, L. (2023). Acceptance of online customer channels for damage claims in Germany. Information Technology and Management, 1-16. https://doi.org/10.1007/s10799-023-00404-z

Méndez-Aparicio MD, Izquierdo-Yusta A and Jiménez-Zarco AI (2017) Consumer Expectations of Online Services in the Insurance Industry: An Exploratory Study of Drivers and Outcomes. Frontiers in Psychology, 8:1254. https://doi.org/10.3389/fpsyg.2017.01254

Nimmagadda, V. S. P. (2022). Artificial intelligence for customer behavior analysis in insurance: Advanced models, techniques, and real-world applications. Journal of AI in Healthcare and Medicine, 1(1), 1-20.

Perumalsamy, J., Krothapalli, B., & Althati, C. (2022). Machine Learning Algorithms for Customer Segmentation and Personalized Marketing in Life Insurance: A Comprehensive Analysis. Journal of Artificial Intelligence Research, 2(2), 83–123.

Saito, T., & Rehmsmeier, M. (2015). The precision-recall plot is more informative than the ROC plot when evaluating binary classifiers on imbalanced datasets. PLOS ONE, 10(3), e0118432.

Shetty, A., Shetty, A. D., Pai, R. Y., Rao, R. R., Bhandary, R., Shetty, J., ... & Dsouza, K. J. (2022). Block chain application in insurance services: A systematic review of the evidence. SAGE Open, 12(1), https://doi.org/10.1177/21582440221079877

Skaf, Y., Eid, C., Thrassou, A., & El Nemar, S. (2024). Technology and service quality: Achieving insurance industry customer satisfaction and loyalty under crisis conditions. EuroMed Journal of Business, 19(1), 1-20.

Stoltzfus, J. C. (2011). Logistic regression: a brief primer. Academic emergency medicine, 18(10), 1099-1104.

Stricker, L., Wagner, J., & Zeier Röschmann, A. (2023). The Future of Insurance Intermediation in the Age of the Digital Platform Economy. Journal of Risk and Financial Management, 16(9), 381.

Trivedi, S. (2023). Blockchain framework for insurance industry. International Journal of Innovation and Technology Management, 20(06), https://doi.org/10.1142/S0219877023500347

Vargo, S. L., Maglio, P. P., & Akaka, M. A. (2008). On value and value co-creation: A service systems and service logic perspective. European Management Journal, 26(3), 145-152.

Viot, C., Lecuyer, C., Bayart, C. and Lancini, A. (2024), "The adoption of smart services: do privacy concerns, trust in benevolence and usage experience matter?", Journal of Consumer Marketing, Vol. 41 No. 6, pp. 648-657. https://doi.org/10.1108/JCM-04-2022-5299

Volosovych, S., Zelenitsa, I., Kondratenko, D., Szymla, W., & Mamchur, R. (2021). Transformation of insurance technologies in the context of a pandemic. Insurance Markets and Companies, 12(1), 1-13. https://doi.org/10.21511/INS.12(1).2021.01

Yego, N. K., Nkurunziza, J., & Kasozi, J. (2023). Drivers of potential policyholders' uptake of insurance in Kenya using Random Forest. Insurance Markets and Companies, 14(1), 22. doi:10.21511/ins.14(1).2023.03

Zaghloul, M., Barakat, S., & Rezk, A. (2024). Predicting E-commerce customer satisfaction: Traditional machine learning vs. deep learning approaches. Journal of Retailing and Consumer Services, 79, https://doi.org/10.1016/j.jretconser.2024.103865.

Zhang, C., Lei, X., Strauss, J., and Zhao, Y. (2017) Health Insurance and Health Care among the Mid-Aged and Older Chinese: Evidence from the National Baseline Survey of CHARLS. *Health Econ.*, 26: 431–449. https://doi.org/10.1002/hec.3322 .

Zhou, F., Ayoub, J., Xu, Q., and Jessie Yang, X. (October 3, 2019). "A Machine Learning Approach to Customer Needs Analysis for Product Ecosystems." ASME. J. Mech. Des. January 2020; 142(1): 011101. https://doi.org/10.1115/1.4044435