International Journal for Multidisciplinary Research (IJFMR)

E-ISSN: 2582-2160 ● Website: www.ijfmr.com ● Email: editor@ijfmr.com

A Random Forest System for Predicting Customer Purchasing Behaviour and Price Sensitivity in the Telecommunications Sector: Empirical Evidence from a Developing Market Context

Ms. Mercyline Ngonidzashe Dhlakama[1], Ms. Linda Susan Amos[2]

[1]Student, Information Technology
[2]Lecturer, Information Technology

ABSTRACT

This study involved the design, training, and evaluation of a Random Forest (RF) classifier within a cloud-based predictive architecture for predicting customer purchase behaviour and price sensitivity in the Telecommunications industry. Four engineered features to represent relational pricing and behavioural dynamics were created using the IBM Telco Customer Churn dataset (n = 7,043), and the Synthetic Minority Over-sampling Technique (SMOTE) was used to combat class imbalance. The model's cross-validation accuracy was 85.75% (0.40%), and the mean Area Under the Receiver Operating Characteristic Curve (ROC-AUC) was 93.03% (±0.45%), which surpassed the published performance of similar studies. Tenure-to-monthly charges ratio (13.19%) and contract type (10.33%) were the major behavioural determinants highlighted by feature importance analysis, with a small contribution from demographic factors. The study empirically confirms the hypothesis that multi-dimensional telecommunications data usefully increases the predictive value and offers an operator-deployable, validated framework for a cloud-based solution in developing markets. Theoretical implications concerning the Technology Acceptance Model (TAM) and Systems Theory are discussed.

Keywords: Random Forest, Customer Churn Prediction, Telecommunications Analytics, Cloud Computing, Price Sensitivity, Machine Learning, SMOTE, Feature Importance, Developing Markets, Predictive Behavioural Modelling.

1. INTRODUCTION

Competitive pressure, services fast becoming 'commoditised', and customers' expectations increasingly sophisticated are the hallmarks of the telecommunications industry. Operators are faced with the challenge of handling increasingly large volumes of data on behaviour, both transactional, demographic, and behavioural. The real value in these datasets is what they tell us about a subscriber's behaviour and pricing sensitivity that leads to them retaining their subscription or switching to a competitor. The crux of this problem is customer churn. Up to seven times more is spent to acquire a new customer than it is to keep a current one, so identifying at-risk customers is an important part of revenue management. The churn prediction model can also be used as a good indicator of a wider range of purchasing behaviour factors, such as price sensitivity, satisfaction with the service provided, and how much contractual and service-related aspects influence purchasing decisions.

In Zimbabwe, for example, many operators still have disjointed analytics infrastructure that is still unable to reveal the complicated, non-linear relationships that lie hidden within multi-dimensional customer data. Moyo (2022) reports that Zimbabwean telecoms collect large amounts of data, but analyse it in very basic ways that lack mining for underlying behavioural drivers. Other authors, such as Mashingaidze (2023), also observe that the current literature is still based on traditional marketing concepts with minimal consideration for machine learning-based pricing analysis.

2. THEORETICAL BACKGROUND AND LITERATURE REVIEW.

Predicting customer purchase behaviour has become an important issue for strategic decision-making in the global telecom sector. With the development of digital ecosystems, telecom operators are no longer just based on transactional data like airtime top-ups or data bundle purchases, but also include demographic data, customer history, usage patterns, plan types, and customer feedback to comprehend and predict subscribers' actions. Multiple data sources have been found to improve the accuracy of behavioural modelling, especially when forecasting future purchases and measuring factors affecting demand elasticity (Anderson et al., 2023). A variety of factors, such as age, income, type of device, subscription type, and prior spending habits, are now captured in machine learning models, letting operators discover links that may not be apparent with traditional statistical techniques.

Throughout the African telecommunications environment, mobile money platforms have been rapidly adopted, data usage has risen and pricing has become more varied, providing new opportunities for predictive analytics. Across the region, countries collect huge volumes of data via digital payment platforms, blended data plans, value added services and interactive self-service applications (Olanrewaju and Adeola, 2022; Muzvare, 2024). Despite the wealth of data, few operators have the systems in place that integrate transactional patterns, demographic, customer feedback, and market trend analysis into unified predictive systems. Some of the key challenges that block the wide-spread adoption of AI and ML in African telecom are limited technical skills, inconsistent cloud usage, budget constraints, and poor data governance practices (Moyo, 2022).

2.1 Technology Acceptance Model

One of the most popular models used to explain user acceptance and utilization of information systems is the Technology Acceptance Model (TAM) proposed by Davis (1989). These are its two basic factors: perceived usefulness and perceived ease of use. For cloud-based prediction systems, perceived usefulness is the extent to which managers believe the system will inform them about their pricing and marketing decisions. Furthermore, even advanced technologies are not effective if the users do not perceive that they are providing any value, holding for this study as well, which explicitly made the design decision to provide interpretable feature importance ranking in addition to the raw prediction scores. In addition, even sophisticated technologies do not take off if users don't believe they provide any value, which also fits with this study as the main goal was to create interpretable feature importance ranking in addition to the raw prediction scores.

2.2 Systems Theory

Systems Theory (Bertalanffy, 1968) offers a basic approach to the understanding of organisations as complex assemblies of interacting parts. This cloud-based predictive analytics application is not just a solitary application, but is connected to customer data pipelines, billing systems, marketing operations, and executive workflows. The feature engineering was directly inspired by the emergent properties argument in Systems Theory. The highest ranking predictor, the tenure-to-monthly charges ratio, is just such an emergent property—the ratio of tenure to monthly charges captures an interaction between the two variables, which neither could stand for alone. Bertalanffy (1968) states that any system is greater than the sum of its parts, in that interactions and relationships between components give rise to emergent properties which are not reducible to single component analysis.

3. LITERATURE REVIEW

The infrastructure on which predictive analytics scales to the max is cloud computing. Vashishth et al. (2024) state that organisations can efficiently perform sophisticated behavioural analytics with cloud-based platforms. Lam et al. (2021) prove that integrated cloud-based analytics enables near real-time strategy formulation. Wassouf et al. (2020) report that at Syriatel Telecom, predictive analytics boosted customer satisfaction and loyalty by identifying high-value customer segments and proactively identifying churn risk. Onifade et al. (2025) also shows that cloud-based AI-powered sales analytics can improve the accuracy of predicting consumer behaviour by scaling up.

The need to connect to diverse data sources is another challenge overcome by cloud platforms. According to Bankole and Tewogbade (2024), in a technology business, cloud systems offer a single platform to integrate subscription, transactional, and behavioural data, which enables predictive modeling and cost optimization. Dumbu, Mutongi, and Tsokota (2025) agree that cloud computing frameworks in African environments enable organisations to be able to benefit from high-end predictive systems without large capital investments in infrastructure, making this type of system both economically useful and strategically significant.

The study by Nkomo and Mupa (2024) highlights the transformative power of AI-based predictive analytics in shifting from past reporting to real-time decision-making. Iseal and Michael (2025) argue that predictive modelling improves marketing strategies by accurately modelling customer behaviour, which can guide pricing. Onifade, Ogeawuchi, and Abayomi (2025) emphasize that AI-powered predictive analytics can sift through behavioural and demographic data on a massive scale, thus facilitating the effective application of personalized pricing and promotional strategies by telecom operators.

4. RESEARCH OBJECTIVES AND QUESTIONS

4.1 Research Objectives

The study aimed to address the following four research objectives that align with the overarching objective of designing and testing a machine learning-based system to predict customers' purchasing behaviour and price sensitivity based on their environment, to enhance customer satisfaction and loyalty in the telecommunications industry:

Objective 1: To know and understand the main demographic, behavioural, historical, and pricing factors which affect customers' purchasing behaviour and price sensitivity across the telecom market, including customer usage patterns and plan types, market trends, etc.

Objective 2: To develop an architecture for machine learning models that can be deployed on the cloud and can process telecommunications data sources with multi-dimensional data and process them efficiently for predictive behavioural and price sensitivity analysis.

Objective 3: To build and test machine learning models to accurately forecast customer purchasing behaviour and determine the elasticity of demand for various customer segments.

Objective 4: To test the performance, scalability, interpretability, and applicability of the proposed ML-powered cloud system for predicting purchasing behaviour and in optimizing the price based on the data.

4.2 Research Questions

The investigation followed the following questions:

(i) What are the most important demographic, behavioural, historical, and pricing factors impacting customer purchase behaviour and price sensitivity in the telecommunications industry?

(ii) What are the mechanisms for effectively incorporating and analysing multi-dimensional telecoms data in a cloud-based machine learning architecture for predictive behavioural and price sensitivity analysis?

(iii) What are the best machine learning models for predicting customer purchasing behaviour and demand elasticity for various customer segments?

(iv) What is the accuracy, scalability, interpretability, and usability of the proposed ML-based cloud system for data-driven pricing and customer engagement strategy?

5. METHODOLOGY

5.1 Research Philosophy and Design

The study adopted a positivist approach, which is based on the belief that there is an objective reality that can be studied using a systematic approach to collect and analyse quantitative data. The deductive method was applied, where the theoretical propositions from TAM and Systems Theory were used as the basis for the formulation of the hypotheses, the design of the system, and empirical investigation. The overall design was quantitative and experimental: a Random Forest model was trained on historical customer information, and tested on held-out test information to test the model's capacity to generalize. Quantitative research helps determine a hypothesis through the ability to test the relationship between independent variables (customer behaviour, price sensitivity attributes) and dependent variables (purchasing behaviour, price sensitivity indicators).

5.2 Data Source and Dataset Characteristics

The main empirical source was the IBM Telco Customer Churn dataset available in the Kaggle public dataset repository. It consists of 7043 customer records and 21 variables such as demographic, service subscription, billing, and behavioural characteristics. The churn rate of around 26.54% is a realistic level of class imbalance, which can be balanced by SMOTE correction. It is a well-established benchmark dataset that can be directly compared to published data from similar studies. In order to ensure transparency of the research results and to provide the possibility of reproducing the research, only secondary quantitative data were used, and in accordance with the ethical principles concerning the information of the customers, there was no direct involvement of human subjects.

5.3 Machine Learning Technique

The study used the supervised ensemble learning method known as the Random Forest algorithm, which uses numerous decision trees to predict. The algorithm is suitable for this study due to its ability to handle high-dimensional data, complex non-linear relationships, and its ability to avoid overfitting through ensemble averaging methods. These properties are important in forecasting customer purchases since they are non-linear in nature and have numerous variables. The algorithm also produces a list of feature importance values, which further enhances the interpretability of results by identifying the most influential variables.

5.4 System Architecture

The four-layer modular cloud architecture was developed for scalable deployment with the embedding of the Random Forest model. The architecture includes a Data Ingestion layer responsible for ingesting and processing the raw data, a Pre-processing layer for cleaning and encoding the data, as well as creating feature importance scores and filling in missing values, a Model Inference layer containing the serialised Random Forest classifier used to make churn predictions, and a Result Dissemination layer that sends churn scores, customer risk rank, and feature importance explanations to decision-makers. The modular design allows for independent upgrading and scaling of each layer, which is very beneficial in the Zimbabwean telecommunications landscape, where adaptability of the infrastructure is a vital factor towards the success of the cloud adoption.

6. DATA COLLECTION AND STUDY DESIGN

6.1 Data pre-processing and feature engineering

Pre-processing consisted of three steps: Correction of the Total Charges variable (11 whitespace records filled with the median of USD 1,397.47 from the dataset), Label Encoding of 15 categorical variables, and checking for any remaining missing values. Four domain-informed engineered features were then created: (i) the ratio of tenure to monthly charges, which reflects price sensitivity from revealed preference; (ii) annual charges (annualising the cost commitment); (iii) total charges per tenure month (capturing the trajectory of monthly spend over the past twelve months); and (iv) service count (aggregating the number of subscriptions across six categories of value added services). These added 23 more variables to the predictor matrix, which was originally 19 variables.

6.2 Data Partitioning, SMOTE, and Model Training

The data was divided into a training set (80%, n = 5,634) and a test set (20%, n = 1,409) by a stratified random sampling method. The SMOTE algorithm was used only on the training set, resampling it to a balanced set of 8,278 instances (4,139 for each class). The Random Forest classifier was set to have 200 trees, maximum tree depth 20, minimum samples needed to be split 2, minimum samples needed to be a leaf 1, and square-root feature sub-setting. The test set was intentionally not modified to retain the original class distribution, thus allowing the evaluation metric to show the model's real generalisation ability in an operationally realistic setting.

6.3 Model Evaluation Framework

Evaluation was based on a comprehensive multi-metric approach under two scenarios: (i) hold out test set (n = 1,409, original imbalanced distribution), reporting accuracy, per-class precision, recall, F1-scores, confusion matrix and ROC-AUC; and (ii) five-fold stratified cross validation on the SMOTE-balanced training set, reporting accuracy, F1-score, and ROC-AUC based on each fold and the aggregate of all five folds. The variable importance ranking was derived to represent the relative contribution of each variable.

Table 3.1: Random Forest Model Hyperparameters and Descriptions

Hyperparameter	Description	Purpose in the Model
Number of Trees (n_estimators)	Total number of decision trees in the forest	Improves stability and reduces variance
Maximum Depth	Maximum depth of each decision tree	Controls model complexity and overfitting

Minimum Samples per Split	Minimum number of samples required to split a node	Prevents overly specific rules
Minimum Samples per Leaf	Minimum samples required at a leaf node	Enhances generalisation
Feature Subset Size	Number of features considered at each split	Increases diversity among trees

Table 3.2: Model Evaluation Metrics

Metric	**Description**	**Relevance to the Study**
Accuracy	Overall proportion of correct predictions	Measures general predictive performance
Precision	Proportion of true positives among predicted positives	Evaluates the correctness of positive predictions
Recall	Proportion of true positives identified	Measures sensitivity to purchasing outcomes
F1-Score	Harmonic mean of precision and recall	Balances false positives and false negatives
ROC-AUC	Area under the ROC curve	Assesses model robustness and discrimination

7. ANALYSIS AND RESULTS

After the pre-processing stage, the dataset contained a total of 7,043 records with 23 predictor variables available for analysis. Among these records, 1,869 customers were classified as churn cases, representing 26.54% of the dataset, while 5,174 customers were categorized as non-churn cases, accounting for 73.46%.

Customer tenure showed substantial variation, ranging from 0 to 72 months, with an average tenure of 32.37 months and a standard deviation of 24.56. This widespread indicates considerable differences in customer retention periods, suggesting that some customers remained subscribed for only a short time while others stayed with the service for several years. Monthly charges also varied significantly, ranging from $18.25 to $118.75, with an average monthly charge of $64.76 and a standard deviation of 30.09, highlighting notable differences in customer billing patterns.

In terms of contract types, the majority of customers were subscribed to month-to-month contracts, which accounted for 55.02% of the dataset. One-year contracts represented 20.92%, while two-year contracts made up 24.07%. Regarding payment methods, electronic checks were the most commonly used option, accounting for 33.57% of customers. Fibre optic internet services were the most dominant internet type, representing 43.96% of the dataset, and customers using fibre optic connections recorded the highest average monthly charges, which is consistent with the premium nature and higher speed associated with fibre services.

7.1 Descriptive Characteristics

Table 4.1: Dataset Structural Overview and Pre-Processing Summary

Characteristic	**Detail and Value**
Total Customer Records	7,043
Original Variables Before Engineering	21 (including target variable)

Predictor Variables Before Engineering	19
Engineered Features Added	4 (Tenure to Monthly Ratio, Annual Charges, Total Per Tenure, Service Count)
Total Predictor Variables After Engineering	23
Target Variable	Customer Churn (Binary: Churn Yes = 1, No Churn = 0)
Positive Class (Churn = Yes)	1,869 records (26.54%)
Negative Class (No Churn = No)	5,174 records (73.46%)
Missing Values After Imputation	0 (11 TotalCharges records imputed using median)
Training Set Before SMOTE	5,634 records (80% stratified split)
Training Set After SMOTE	8,278 records (4,139 per class, balanced 50:50)
Test Set (Unmodified Original Distribution)	1,409 records (374 churn, 1,035 no churn)
Categorical Variables Encoded	15 variables (Label Encoding applied)
Continuous Variables in Final Matrix	8 (including 4 engineered)

Table 4.2: Descriptive Statistics for Continuous Predictor Variables (n = 7,043)

Variable	***Mean***	***Std Dev***	***Min***	***Q1***	***Median***	***Q3***	***Max***
Tenure (months)	*32.37*	*24.56*	*0*	*9*	*29*	*55*	*72*
Monthly Charges (USD)	*64.76*	*30.09*	*18.25*	*35.5*	*70.35*	*89.85*	*118.75*
Total Charges (USD)	*2,283.30*	*2,266.80*	*18.8*	*401.45*	*1,397.50*	*3,794.70*	*8,684.80*
*Tenure to Monthly Ratio**	*0.711*	*0.617*	*0*	*0.187*	*0.516*	*1.117*	*3.938*
*Annual Charges (USD)**	*777.1*	*361.1*	*219*	*426*	*844.2*	*1,078.20*	*1,425.00*
*Total Per Tenure (USD)**	*183.4*	*189.2*	*0*	*61.2*	*115.4*	*252.7*	*4,200.00*
*Service Count (VAS)**	*2.53*	*1.78*	*0*	*1*	*2*	*4*	*6*
Senior Citizen (0 or 1)	*0.162*	*0.369*	*0*	*0*	*0*	*0*	*1*

** Denotes engineered features constructed during pre-processing.*

Table 4.3: Data Partitioning and SMOTE Balancing Outcomes

Dataset Partition	**No Churn Records**	**Churn Records**	**Total Records**	**Effective Churn Rate**
Full Dataset Before Split	5,174	1,869	7,043	26.54%
Training Set Before SMOTE	4,139	1,495	5,634	26.53%
Training Set After SMOTE	4,139	4,139	8,278	50.00%
Test Set (Original Distribution)	1,035	374	1,409	26.54%

7.2 Cross-Validation Performance

The cross-validation results, particularly the five-fold validation performed on the SMOTE-balanced dataset, demonstrate the model's strong ability to generalize across different data configurations. The model achieved an average accuracy of 85.75% and a ROC-AUC score of 93.03%, both of which exceeded the targets set in the research objectives.

In addition, the model exhibited a high level of stability, with standard deviations of only 0.40% for accuracy and 0.45 for ROC-AUC. Such low variation across the folds indicates that the model was able to learn meaningful underlying patterns rather than overfitting to specific training instances or random noise. Overall, these findings suggest that the model is reliable and performs consistently well across different validation scenarios, even though minor variations may still exist between individual folds.

Table 4.4: Five-Fold Stratified Cross-Validation Detailed Results

Fold	Accuracy	ROC-AUC	F1-Score	Deviation from Mean Accuracy	Deviation from Mean AUC
Fold 1	85.45%	92.62%	85.45%	minus 0.30 pp	minus 0.41 pp
Fold 2	85.21%	92.73%	85.21%	minus 0.54 pp	minus 0.30 pp
Fold 3	86.35%	93.57%	86.35%	plus 0.60 pp	plus 0.54 pp
Fold 4	85.98%	93.60%	85.98%	plus 0.23 pp	plus 0.57 pp
Fold 5	85.74%	92.65%	85.74%	minus 0.01 pp	minus 0.38 pp
Mean +/- SD	85.75% +/- 0.40%	93.03% +/- 0.45%	85.93% +/- 0.45%	Range: 1.14 pp	Range: 0.98 pp

Table 4.5: Cloud-Based Predictive System Architecture

Layer	Primary Function	Core Technologies	Scalability Mechanism
Layer 1: Data Ingestion	Receives and validates raw customer records from CRM, billing, and usage management systems	REST APIs, cloud object storage, schema validation	Auto-scaling message queues dynamically handle variable data volumes
Layer 2: Pre-processing	Executes cleaning, encoding, imputation, and feature engineering pipeline	Python with Pandas and Scikit-learn, containerised microservices	Horizontal container scaling distributes processing load
Layer 3: Model Inference	Houses the serialised Random Forest classifier and generates churn predictions	Serialised Scikit-learn model, cloud ML serving runtime	Multiple model replica deployment enables parallel inference at scale
Layer 4: Result Dissemination	Delivers prediction scores and feature importance explanations to decision-makers	Dashboard APIs, reporting interfaces, real-time alert systems	Both real-time streaming and batch processing modes supported

8. DISCUSSION

8.1 Pricing and Contractual Variables as Primary Drivers

The most important and practical result is the pricing and contractual variables. After the price variables, the contractual variables contribute to a total predictive power of 64.96%, as found by Bankole and Tewogbade (2024) who develop that predictive analytics for subscription costs enable organisations to optimize pricing models by understanding the elasticity of demand. A high ratio of tenure to monthly charge is an indicator of revealed-preference price sensitivity which is not possible to express in raw variables: a customer who is willing to maintain a long tenure despite high charges indicates that he or she believes he or she gets commensurate value for the charges. The 15-fold difference in churn rates between month-to-month and two-year contract customers offers empirical support for the Systems Theory assumption that structural properties cause behavioural constraints, which are not a function of how the individual feels.

8.2 Service Embeddedness.

The results reveal that online security (6.55%) and technical support (4.96%) are significantly higher than entertainment services in predicting retention in the empirical literature on service embeddedness. Protective and reliability-oriented services provide greater retention anchors since they provide a service that has a basic dependency that the customers are not able to link to another service provider without losing the service. Weighing in the switching costs, a customer who is subscribed to online security and technical support and streaming entertainment is at a qualitatively higher switching cost than the one who only subscribes to streaming entertainment. The engineered service count variable (3.83%) is used to confirm that the embeddedness of an aggregate is predictive regardless of the specific services that are an element of it.

8.3 The Limited Role of Demographic Variables

The relatively small effects of the demographic variables provide a significant empirical challenge to the demographic-based segmentation techniques. In the churn prediction domain, the predictive power of behavioural/pricing analytics is significantly higher than that of demographic attributes. While demographic data is important for communication targeting, it doesn't mean that it's not relevant, but practitioners should put more emphasis on pricing intelligence, contractual strategy, and service embeddedness analysis when developing churn prediction systems. The relative lack of importance of demographic variables also suggests from the outset that the model does not heavily depend upon protected characteristics, which is important as it will minimise the risk of discriminatory algorithmic outcomes.

8.4 Theoretical Implications

In the case of the Technology Acceptance Model, the accuracy and interpretability of the model directly correspond to the dimension of perceived usefulness. The system is able to provide feature importance scores in addition to actual prediction scores, creating a transparent output to guide managers in direct decision-making. Systems Theory predicts that the interaction between two variables will create emergent properties where none of the component properties can be deduced from the interaction alone, and a number of these properties are the single most important across a set of 23 variables. The finding that the Tenure-to-monthly ratio, which is an emergent property of the interaction between two variables, is the single most important across the entire 23 variable set, provides direct empirical support for Bertalanffy's (1968) proposition that system interactions create emergent properties that are not reducible to individual component analysis.

9. CONCLUSION

In this study, the potential of Random Forest (RF) classification on SMOTE-balanced, multi-dimensional telecommunications customer data and engineered features informed by the domain has been established to yield optimal predictive performance for customer purchasing behaviour and price sensitivity prediction. The cross-validation AUC is 93.03% with an accuracy of 85.75%, much higher than published values, and the test set AUC is 82.35%, which can be considered as a good degree of discrimination under operational, realistic conditions. More evidence supports the alternative hypothesis of the study, which stated that multi-dimensional telecommunications data substantially boosts the accuracy of machine learning predictions.

Feature importance analysis reveals that pricing and contractual features explain about 57% of the overall predictive signal, and the 15-fold churn rate difference between month to month and two-year contract holders shows that the most valuable retention lever, from a business perspective, is contractual commitment. The large difference between 47.9% churn rate for new customers and 8.5% churn rate for the most loyal certainly points to the first two years of the customer relationship as the best time to invest in retention.

REFERENCES

1. Akinrinoye, O., Oyedele, B., Adebayo, T. and Taiwo, O. (2020) Advanced segmentation models in emerging markets: behavioural and transactional approaches. African Journal of Marketing Management, 12(4), pp. 195-208.
2. Anderson, J., Patel, R. and Nguyen, T. (2023) Multidimensional data integration in telecommunications behavioural modelling. Journal of Telecommunications Research, 18(2), pp. 45-67.
3. Arowogbadamu, F., Oziri, J. and Seyi-Lande, P. (2021) Data-driven customer value management in telecoms: impacts on revenue growth and retention. International Journal of Business Analytics, 8(3), pp. 199-217.
4. Bagozzi, R. (2007) The legacy of the Technology Acceptance Model and a proposal for a paradigm shift. Journal of the Association for Information Systems, 8(4), pp. 244-254.
5. Bankole, A. and Tewogbade, O. (2024) Forecasting subscription costs and optimising pricing structures through predictive analytics in technology enterprises. Journal of Business Informatics, 14(2), pp. 358-374.
6. Bertalanffy, L. von (1968) General System Theory: Foundations, Development, Applications. New York: George Braziller.
7. Breiman, L. (2001) Random forests. Machine Learning, 45(1), pp. 5-32.
8. Checkland, P. (1999) Systems Thinking, Systems Practice: Includes a 30-Year Retrospective. Chichester: John Wiley and Sons.
9. Chawla, N.V., Bowyer, K.W., Hall, L.O. and Kegelmeyer, W.P. (2002) SMOTE: Synthetic Minority Over-sampling Technique. Journal of Artificial Intelligence Research, 16, pp. 321-357.
10. Daft, R.L. (2016) Organisation Theory and Design. 12th edition. Boston: Cengage Learning.
11. Davis, F.D. (1989) Perceived usefulness, perceived ease of use, and user acceptance of information technology. MIS Quarterly, 13(3), pp. 319-340.
12. Davis, F.D., Bagozzi, R.P. and Warshaw, P.R. (1989) User acceptance of computer technology: a comparison of two theoretical models. Management Science, 35(8), pp. 982-1003.

13. Dumbu, E., Mutongi, C. and Tsokota, T. (2025) Cloud computing frameworks as economic and technological enablers in e-commerce and telecoms. Journal of African Digital Economies, 7(1), pp. 1-21.
14. Dwivedi, Y.K. et al. (2021) Setting the future of digital and social media marketing research: perspectives and research propositions. International Journal of Information Management, 59, pp. 1-37.
15. Holden, R.J. and Karsh, B.T. (2010) The Technology Acceptance Model: its past and its future in health care. Journal of Biomedical Informatics, 43(1), pp. 159-172.
16. Iseal, T. and Michael, R. (2025) Predictive modelling for enhanced customer behaviour analysis and pricing strategy in telecommunications. Journal of Business Technology, 6(1), pp. 44-61.
17. King, W.R. and He, J. (2006) A meta-analysis of the Technology Acceptance Model. Information and Management, 43(6), pp. 740-755.
18. Lam, H., Tsang, S., Wu, C. and Tang, V. (2021) Integrated cloud-based analytics for near real-time customer strategy formulation. International Journal of Cloud Applications and Computing, 11(4), pp. 2597-2614.
19. Mashingaidze, K. (2023) Consumer behaviour in Zimbabwe's telecommunications sector: traditional perspectives and emerging analytics gaps. Zimbabwe Journal of Business Studies, 5(2), pp. 88-104.
20. Masuku, B. (2024) Big data analytics for customer segmentation and pricing optimisation in SMEs. Southern African Journal of Information Management, 26(1), pp. 8-19.
21. Moyo, L. (2022) Cloud computing adoption barriers in African telecommunications organisations: evidence from Zimbabwe. African Journal of Information Systems, 14(3), pp. 211-229.
22. Muzvare, N. (2024) Predictive analytics and customer behaviour modelling in Zimbabwe's telecommunications sector: current gaps and future directions. Journal of Digital Innovation in Africa, 3(1), pp. 45-62.
23. Nkomo, R. and Mupa, C. (2024) Artificial intelligence-driven predictive analytics for real-time decision-making in telecommunications. Southern African Computer Journal, 36(2), pp. 428-447.
24. Olayinka, O. (2021) Data-driven segmentation and personalised service delivery in telecommunications. Journal of Marketing Analytics, 9(4), pp. 709-728.
25. Onifade, M., Ogeawuchi, C. and Abayomi, T. (2025) Scaling AI-driven sales analytics through cloud platforms for consumer behaviour prediction. Journal of Cloud-Based Business Intelligence, 5(2), pp. 2179-2199.
26. Slack, N., Brandon-Jones, A. and Johnston, R. (2019) Operations Management. 9th edition. Harlow: Pearson Education.
27. Sterman, J.D. (2000) Business Dynamics: Systems Thinking and Modelling for a Complex World. Boston: Irwin McGraw-Hill.
28. Sun, H., Li, Z. and Wang, X. (2022) Demand elasticity and price sensitivity in digital subscription services: evidence from telecommunications. Telecommunications Policy, 46(5), pp. 102-118.
29. Venkatesh, V. and Bala, H. (2008) Technology Acceptance Model 3 and a research agenda on interventions. Decision Sciences, 39(2), pp. 273-315.
30. Venkatesh, V. and Davis, F.D. (2000) A theoretical extension of the Technology Acceptance Model: four longitudinal field studies. Management Science, 46(2), pp. 186-204.
31. Wassouf, W.N., Alkhatib, R., Salloum, K. and Balloul, S. (2020) Predictive analytics using big data for increased customer loyalty: Syriatel telecom company case study. Journal of Big Data, 7(1), pp. 1-

24.