International Journal of Intelligent Information Technologies
Volume 17 • Issue 3 • July-September 2021

Predicting Mobile Portability Across Telecommunication Networks Using the Integrated-KLR

Ayodeji Samuel Makinde, Edo State University, Uzairue, Nigeria

Abayomi O. Agbeyangi, Chrisland University, Abeokuta, Nigeria

Wilson Nwankwo, Edo University, Iyamho, Nigeria

ABSTRACT

Mobile number portability (MNP) across telecommunication networks entails the movement of a customer from one mobile service provider to another. This, often, is as a result of seeking better service delivery or personal choice. Churning prediction techniques seek to predict customers tending to churn and allow for improved customer sustenance campaigns and the cost therein through an improved service efficiency to customer. In this paper, MNP predicting model using integrated kernel logistic regression (integrated-KLR) is proposed. The Integrated-KLR is a combination of kernel logistic regression and expectation-maximization clustering which helps in proactively detecting potential customers before defection. The proposed approach was evaluated with five others, mostly used algorithms: SOM, MLP, Naïve Bayes, RF, J48. The proposed iKLR outperforms the other algorithms with ROC and PRC of 0.856 and 0.650, respectively.

KEYWORDS

Churn Prediction, Integrated-KLR, Mobile Number Portability, Sustainability, Telecommunication Sector

1. INTRODUCTION

MNP or otherwise known as “customer churn in disguise”, is the complete churning of a subscriber from one service provider (SP) to a supposedly better SP without changing the phone number (Lin, Chlamtac, & Yu, 2003). If this was not available, customers of any service provider would have to forgo their existing mobile number (MN) and take a new MN and it could come with the cost of a new business card or loss of valuable number if no backup exists. This is an excellent opportunity for many well-known business customers who may be reluctant to churn their SP and give up their mobile number even if the alternative SP provides better services (Gans, King, & Woodbridge, 2001). MNP reduces switching costs and increases telecom competition, thereby promoting sustainability by encouraging operators to offer innovative solutions (Nwankwo & Njoku, 2020).

However, telecom companies operating in saturated markets are continuously faced with the challenges of retaining and sustaining their potential customers due to the MNP techniques, which is not a blessing in disguise. This is because these companies expend enormous resources to attract and retain customers. They understand that customer retainership is the panacea for preventing customers from churning. Moreover, not only is anticipating churn difficult, but trying to predict how significant a problem it will turn out to be; how long it will last; and the ultimate consequences (in terms of loss of revenue, especially when a high potential customer churns to a better competitor) it will incur on

DOI: 10.4018/IJIIT.2021070104

the company is a challenge (Verbeke, Dejaeger, Martens, Hur, & Baesens, 2012). It is also noteworthy that mobile number churning can result in difficulties distinguishing different network operators' customers' number patterns when placing a call. Notably, doing the same was without difficulties prior to the introduction of churning.

Several machine learning techniques have been proposed in recent years to aid the discovery of hidden patterns in datasets, modelled to resolve the problems posed by churn across the telecom operators that have led to customers' decision to change their mobile phone numbers and their service providers respectively. Some of these machine learning techniques include hybrid firefly (Amin et al., 2017), ProfTree (Höppner, Stripling, Baesens, vanden Broucke, & Verdonck, 2018), rough set theory, RotBoost (De Bock, & Van den Poel, 2011), logit leaf model (De Caigny, Coussement, & De Bock, 2018), multilayer perceptron neural network (MLP), Decision Tree C4.5, support vector machines (SVM) (Huang et al., 2010; Huang, Kechadi, & Buckley, 2012), random forest and particle swarm optimization (Idris, Rizwan, & Khan, 2012), Bayesian network (Kisioglu, & Topcu, 2011), and SVM-Polynomial kernel (SVM-POLY) (Vafeiadis, Diamantaras, Sarigiannidis, & Chatzisavvas, 2015). However, little or no research work has been published on solving the real problem of mobile number portability, which could be regarded as another dimension of customer churn in disguise. Recently, there has been an increase in number portability, and in Nigeria, for example, several advertisements were promoted by mobile telecom operators on the issue of portability in late 2018 and 2019. One of the most cited benefits is the flexibility it offers to potential or prospective customers, particularly in terms of retaining their telephone number following porting to a new platform.

In this paper, a mobile number portability predicting model using integrated kernel logistic regression (Integrated-KLR) is proposed. The Integrated-KLR is a combination of kernel logistic regression and expectation-maximization clustering which helps in proactively detecting potential customers before defection. The significant contributions of this study over other similar studies in the domain of customer churn prediction are: (i) predicting and presenting mobile number portability as a threat to customer churn in the telecommunication industries. No research work that we know of has reported mobile number portability as a threat to customer churn in telecommunication industries and, (ii) determining the degree of customer churn levels. The degrees of churn as either low, medium or high for effective customer churn management have also not been reported in literature.

The rest of the paper is organized as follows: The next section presents the literature review of telecommunication customer churn prediction approached and critical discussion on existing techniques; the proposed methodology of this study is explored in Section 3. Section 4 presents the results, comparisons and discussion of the findings; the paper is concluded in Section 5.

2. LITERATURE REVIEW

2.1 Customer Chun Prediction Model (CCP)

Several researchers have examined and analyzed customer churn in various sectors, especially on mobile communication networks. Table 1 shows a summary of some related work to customer churn. Nwankwo and Njoku (2020) proposed a CCP model based on distanced factors to efficiently predict customer churns and also estimate the level of certainty of the classifier's decision in a given TCI dataset. It was claimed that if a mechanism could be defined to determine the classifier's certainty for different zones within the data, then the expected classifier's accuracy could be estimated even before the classification. It was reported that using the state-of-the-art evaluation measures, the distance factor is strongly correlated with the certainty of the classifier. The classifier obtained higher accuracy in the zone with a greater distance factor value than those placed in the zone with a smaller distance factor value.

A study undertaken by Ahmed and Maheswari (2017) presented a meta-heuristic-based churn prediction technique using hybrid firefly-based classification on massive telecom data. The Firefly algorithm was used to compare different components of the firefly with other Firefly and identify which firefly has the highest light intensity; then replaced the component using Simulated Annealing and the classification process was carried out. Furthermore, they used receiver operating characteristics (ROC) to compare the performance of firefly and hybrid Firefly. It was found that the Firefly algorithm works best on churn data and that the hybridized Firefly algorithm provides effective and faster results. On the other hand, the area under the ROC curve is often not recommended for profit maximization in the business (Maldonado, Flores, Verbraken, Baesens, & Weber, 2015).

Although the reviewed literature revealed that several works attempted to address the challenges of customer churn in the telecom industry by categorizing customers as churn or non-churn, no known study was associated with mobile number portability from a customer churn perspective. Surprisingly, no consideration has been given to segmenting churned customers into different levels of severity, such as low, medium, or high. Hence, this study is considered a novel contribution wherein the integrated KLR is deployed for classifying customers into porting or non-porting customers. Following such classification, the ported customers are further clustered into three categories using expectation maximization clustering based on their degree of portability.

3. PROPOSED METHOD

The schematic of the proposed method for the customer portability prediction process is shown in Figure 1. The experimental design of this study was chosen to assess the performance of the integrated kernel logistic regression (Integrated-KLR) against four other benchmark algorithms. The proposed mobile number portability prediction process comprises four stages which include data collection, pre-processing, classification, and clustering stage. The functionality of each stage is explained in the following subsection.

3.1 Development Stages

3.1.1 Data Collection

At the data collection stage, the problem domain was first identified to understand the type of data to collect. In this study, it may be restated that the objective is to identify porting and non-porting customers and to determine the degree of porting by clustering the ported customers into three clusters (low ‘weak’, medium, or high ‘strong’). Historical data was collected from a mobile telecom provider in Nigeria. The collected data consists of 5000 customer records. The proposed architecture provides room for new dataset supplied by the telecom users. The new data instances are first stored in a separate database and then immediately migrated to the historical data database.

3.1.2 Pre-Processing

At this stage, the errors and inconsistencies were detected and removed from the data to improve the quality of data (Rahm & Do, 2000). Having domain knowledge about the data is very important for effective features engineering. Domain knowledge helps in preventing over and underfitting to have a better understanding of the features to use. In order to improve the quality of the proposed model, new features from existing variables were created. This allows machine learning algorithms to comprehend data and discover patterns that might increase the performance of machine learning algorithms. To achieve this, a deep feature synthesis (DFS) algorithm was employed.

An iterative imputer developed by Scikit-Learn was used to model each feature with missing values as a function of other features. At each step, a feature is selected as an output, and all other features are treated as inputs. A regressor is then fitted on both the inputs and output and used to predict the missing values of the output. This is carried out for each feature and repeated for several imputation rounds.

Table 1. Summary of Related Work in Customer Churn Prediction Model

Author (year)	Method	Evaluation Criteria	Scope	Issues Addressed
Verbeke et al. (2012)	Decision Trees		Wireless telecom operator	Improveed the efficiency of customer retention campaigns and to reduce the costs associated with churn.
Amin et al. (2017)	Intelligent rule-based decision-making technique, based on rough set theory (RST)	Sensitivity/recall, Specificity, Precision, Accuracy, Misclassification error, F-Measure, Coverage	Telecommunication Sector	Lack of efficient, rule-based customer churn prediction approach in the telecommunication sector
De Caigny et al. (2018)	Logistic leaf model	AUC, and top decile lift (TDL)	Financial services, Newspaper, Telcommunication, Retail, DIY, Energy	Improved data classification
Huang et al. (2012)	Logistic Regressions, Naïve Bayes, Linear Classifiers, Decision Trees C4.5, Multilayer Perceptron Neural Networks, Support Vector Machines and the Evolutionary Data Mining Algorithm	True Positive, False Positive, Area under the receiver operating characteristics curve (AUC)	Telecommunication	Improved the accuracy of customer churn prediction in telecommunication service field.
Tsai and Chen (2010)	Decision Trees, Neural network, Association rules		Multimedia on demand (MOD) customer churn prediction	Retained MOD customers in the telecommunication industry.
De Caigny et al. (2020)	Convolutional neural network	AUC, and top decile lift (TDL)	European financial services provider	Incorporating textual data into customer churn prediction models
Amin et al. (2019)	Naïve Bayes	Precision, Recall, F-Measure, Accuracy	Telecommunication industry	Customer churn

For feature selection, recursive feature elimination algorithm was executed. Recursive feature elimination is a covering feature determination strategy. It is an insatiable advancement algorithm which works to locate the best performing highlight subset (Rao & Rao, 2021). Table 2 presented an overview of selected features that was used in this study.

3.1.3 Classification

At this stage, the kernel logistic regression was used to classify the output from the pre-processing stage. The KLR algorithm classifies the output as either porting or non-porting. The network with the

Figure 1. Proposed mobile number portability prediction process using Integrated-KLR

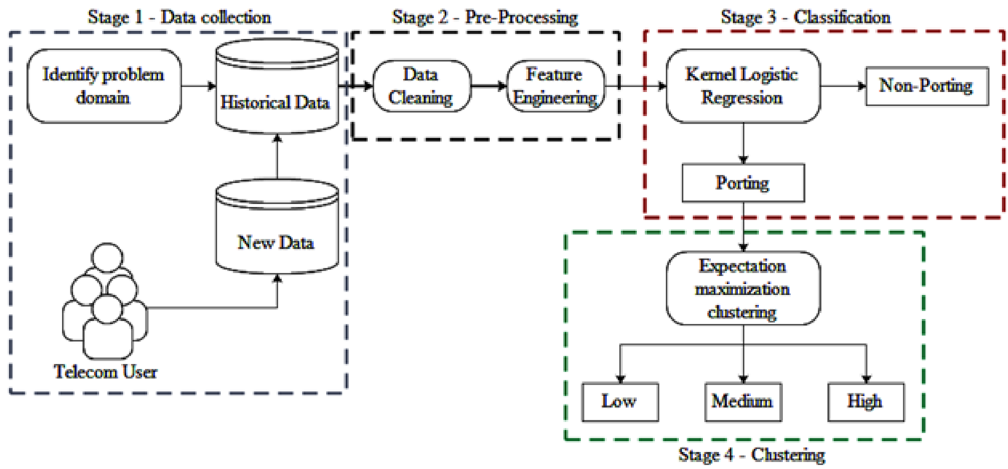


Table 2. Overview of selected variables in mobile telecommunication provider in Nigeria

Variable	Definition
Service quality	How users rate the output of the service delivery system by the telecommunication.
Satisfaction	Is the user assessment of the performance of a product or service.
Trust	How users rate the products and services reliability
Commitment	How often does the users use the services
Price	how users rate the variety of pricing scheme such as the attractiveness of promotions.
Reputation	overall quality as seen or judged by users in general
Ease of Use	How easily users can use a product.
Switching Costs	how users rate the fee in which customers pays for the current carrier in order to port out, and the fee in which the customers pays for his new carrier in order to port in.
Coverage	How users rate the quality of network connection
Value added Services	how users rate the variety of value added services such as voice mail, e-mail, and fax mail.
Customer Support	how users rate the range of services provided to assist customers in making cost effective and correct use of a product.
Loyalty program	how users rate the structured marketing efforts that reward, and therefore encourage, loyal buying behavior
Addition Service fee	how users rate the reasonability of the value added services prices.

most coverage, the network where the majority of calls are made, the network with the most innovative solutions, the network with the most bundled offerings, the network that caters to all technology needs, and the network with the most high-profile events and experience are among the criteria for port.

3.1.4 Clustering Stage

This is the final stage in the model development phase whereby the classified ported data is acted upon by a clustering algorithm, i.e. the expectation maximization clustering (EMC). The EMC algorithm creates three clusters from the ported data based on their level or degree of porting (i.e.

Table 3. Clustering values range

Clustering Variable	Value Range
Low	$0.0 \le x < 0.3$
Medium	$0.3 \le x < 0.6$
High	$0.6 \le x \le 1.0$

low, medium or high). Clustering is required in order to determine the severity of a user's porting. It will also assist in focusing on the most susceptible consumers and developing retention tactics. The clustering values range for users porting severity class was provided in Table 3.

3.2 Integrated Kernel Logistic Regression (iKLR)

The kernel logistic regression is a powerful classification technique that has been applied successfully in many classification problems. Most of the developed current customer porting (churn) prediction methods use algorithms such as support vector machines (SVMs) and neural networks. However, KLR has not been used in the classification of customer porting, perhaps due to its high computational requirements. In this study, iKLR is used to obtain sparse customer porting predictions within a short time. The fitted function of the KLR performs similarly to the SVM for two-classification because of the similarity between the two-loss functions, as shown in Equation 1:

$$\left(1 - yf\right)_+ \Rightarrow \ln\left(1 + e^{-yf}\right) \tag{1}$$

This replacement in the loss function gives the KLR unmatched advantages of offering a natural estimate of the probability given in Equation 2. Also, KLR can naturally be generalized to the multiclass case through kernel multi-logit regression (Karsmakers, Pelckmans, & Suykens, 2007). Therefore, the new KLR is as given in Equation 3:

$$p\left(x\right) = e^{f(x)} / \left(1 + e^{f(x)}\right) \tag{2}$$

$$\min_{f \in H_K} \frac{1}{n} \sum_{i=1}^{n} ln\left(1 + e^{-y_i f(x_i)}\right) + \frac{\lambda}{2} f_{H_k}^2 \tag{3}$$

3.3 Expectation-Maximization (EM) Clustering

What makes EM clustering so appealing in the context of this work is its ability to optimize a large number of variables simultaneously. Also, EM clustering can find reasonable estimates for any missing information in the data and cluster multidimensional data that lends to modelling using Gaussian mixture; EM clustering can create both traditional hard clusters and non-traditional soft clusters.

3.4. Performance Metrics

True positive (TP) rate, false positive (FP) rate, false discovery rate (FDR), precision, Matthews correlation coefficient (MCC), f-measure, recall, ROC area, and PRC area are examples of cutting-edge performance metrics. The mathematical equations of these performance metrics are given in Equation (4) - (9), respectively.

3.4.1 False Discovery Rate (FDR)

The false discovery rate is the inverse of the positive predictive value (PPV), which indicates the likelihood that a 'significant' outcome actually indicates the presence of an impact. Thus, if the false discovery rate is 70%, the positive predictive value is 30%:

$$FDR = \frac{FP}{FP + TP} \tag{4}$$

3.4.2 Precision

Precision refers to how precise/accurate the model is in predicting positive outcomes; specifically, how many of those expected positive outcomes are real positive outcomes. When the cost of False Positives is large, precision is an important metric to consider:

$$Precision = 1 - FDR \tag{5}$$

3.4.3 Matthews Correlation Coefficient (MCC)

The Matthews correlation coefficient is an alternate metric that is untouched by the problem of unbalanced datasets. It is a contingency matrix approach for measuring the Pearson product-moment correlation coefficient between real and expected values. MCC is the only binary classification rate that produces a high score if the binary predictor accurately predicted the plurality of positive and negative data instances:

$$MCC = \frac{TP \times TN - FP \times FN}{\sqrt{\left(TP + FP\right)\left(TP + FN\right)\left(TN + FP\right)\left(TN + FN\right)}} \tag{6}$$

3.4.4 F-Measure

F-Measure enables the simultaneous measurement of precision and recall in a single metric that weights all properties equally:

$$F - Measure = 2 \cdot \frac{precision \cdot recall}{precision + recall} \tag{7}$$

3.4.5 Recall

Recall determines the proportion of individual positives that the model captures by marking them as TP as shown in equation 8. When there is a large expense involved with FN, recall shall be the model metric used to find the correct model. Recall is the proportion of examples expected to belong to a class that actually belong to the class divided by the total number of examples. When the cost of false positives is high, recall aids in the decision-making process:

$$Recall = \frac{TP}{\left(TP + FN\right)} \tag{8}$$

3.4.6 ROC Area

A Receiver Operator Characteristic (ROC) area is used to illustrate binary classifiers' diagnostic capability. The ROC is not dependent on the distribution of the classes. This makes it particularly valuable for testing classifiers that forecast unusual incidents like consumer churn. Therefore, the ROC value is the area at which the true positive rate (TPR) or recall as shown in equation 8 is plotted against the false positive rate (FPR) in equation 9:

$$FPR = \frac{FP}{FP + TN} \tag{9}$$

3.4.7 PRC Area

The field under the precision-recall curve (PRC) area represents the relationship between precision and recall over all possible cut-off values for a measurement. When assessing binary classifiers on unbalanced datasets, the PRC region is more descriptive than the ROC area. PRC values are shown alongside corresponding recall values in the PRC region. Similar to the ROC region, the PRC area evaluates the whole model. Therefore, the PRC value is the area at which the precision is plotted against the recall.

4. RESULTS

Understanding the effectiveness of a new model would ultimately demand a comparative evaluation of any related existing model. Table 4 describes the level of accuracy of the non-portability class. From the table, the random forest has the highest TP rate (0.992) with Naïve Bayes having the lowest TP rate (0.682). Likewise, the random forest shows the highest FP rate of 0.925, while Naïve Bayes has the lowest FP rate (0.172). However, having a very high positive rate will worsen the expected output for the Non-portability class prediction. This is evidence in the MCC, ROC area, and PRC area columns, where the iKLR outperforms SMO, MLP, Naïve Bayes, RF, and J48 with 0.456, 0.856, and 0.945, respectively, as shown in Table 3. Although the precision level of Naïve Bayes is the highest (0.918) however, with the lowest recall (0.682) as compared to other models.

The accuracy level of the MNP class is significant for the proposed iKLR, as shown in Table 5 when compared with other models in terms of ROC and PRC area. Compared to the non-portability class prediction in Table 5, Naïve Bayes tends to have the highest values for both TP rate and FP rate with 0.828 and 0.318, respectively. However, the FP rate for the Naïve Bayes (Olayinka, Nwankwo, & Olayinka, 2020; Nwankwo & Ukhurebor, 2020) is relatively low compared to what was observed in Table 5 for the case of random forest. However, checking for the overall performance accuracy, the iKLR has proven more effective compared to other models with ROC and PRC area of 0.856 and 0.650, respectively.

Table 4. Accuracy by class value 'Non-portability'

_TP Rate	FP Rate	Precision	Recall	F-Measure	MCC	ROC Area	PRC Area	
iKLR	0.886	0.450	0.848	0.886	0.867	0.456	0.856	0.945
SMO	0.897	0.467	0.845	0.897	0.870	0.460	0.715	0.834
MLP	0.877	0.459	0.844	0.877	0.860	0.435	0.836	0.937
Naïve Bayes	0.682	0.172	0.918	0.682	0.783	0.450	0.828	0.927
RF	0.992	0.925	0.753	0.992	0.856	0.190	0.852	0.943
J48	0.875	0.492	0.835	0.875	0.854	0.403	0.761	0.853

Table 5. Accuracy by class value 'Portability'

_TP Rate	FP Rate	Precision	Recall	F-Measure	MCC	ROC Area	PRC Area	
iKLR	0.550	0.114	0.629	0.550	0.587	0.456	0.856	0.650
SMO	0.533	0.103	0.646	0.533	0.584	0.460	0.715	0.466
MLP	0.541	0.123	0.608	0.541	0.572	0.435	0.836	0.604
Naïve Bayes	0.828	0.318	0.478	0.828	0.606	0.450	0.828	0.616
RF	0.075	0.008	0.778	0.075	0.137	0.190	0.852	0.638
J48	0.508	0.125	0.589	0.508	0.546	0.403	0.761	0.502

Computing the accuracy of the weighted average for both the MNP and non-portability as presented in Table 6. It was also recorded that the proposed iKLR outperformed other benchmark models with respect to their ROC and PRC area with 0.856 and 0.868, respectively.

From the stratified cross-validation described in Table 7, iKLR has the lowest root means square error (RMSE) of 0.3631. Whereas, the values of iKLR follow that of SMO in most of the stratified cross-validation analysis. Figure 2 shows a high iKLR margin curve, which describes how confident the classifier is in predicting the true class (portability and non-portability).

Figure 3 illustrates the degree at which telecom customers are likely to engage in MNP. It is important to note that classifying telecom customers into porting or non-porting is not sufficient to know the degree of portability of such customer. From Figure 3, only the MNP customers were considered for their degree of portability to other telecom platforms. It was discovered that 31% of

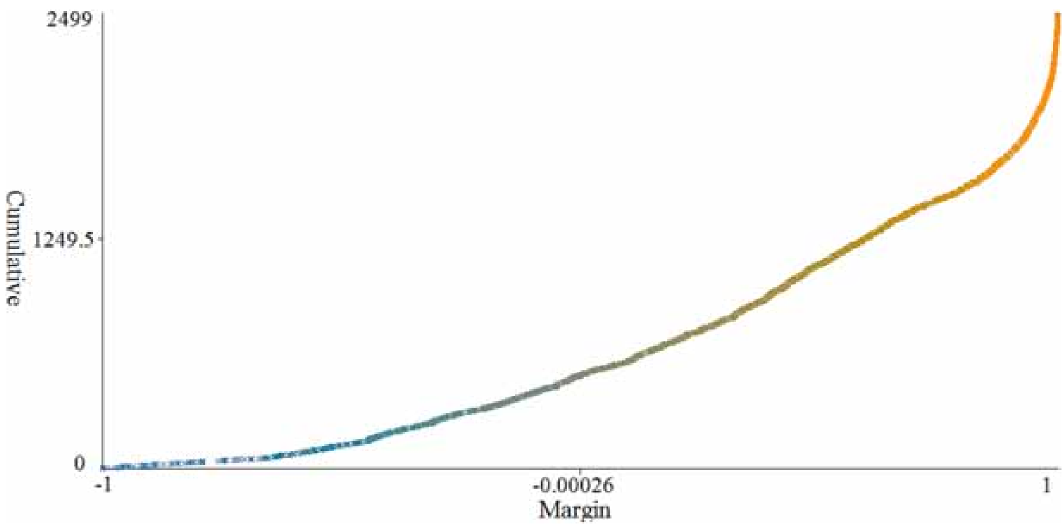
Table 6. Class Accuracy by Weighted Average

_TP Rate	FP Rate	Precision	Recall	F-Measure	MCC	ROC Area	PRC Area	
iKLR	0.798	0.363	0.791	0.798	0.794	0.456	0.856	0.868
SMO	0.802	0.372	0.793	0.802	0.796	0.460	0.715	0.738
MLP	0.790	0.372	0.783	0.790	0.785	0.435	0.836	0.850
Naïve Bayes	0.720	0.210	0.804	0.720	0.737	0.450	0.828	0.846
RF	0.754	0.686	0.759	0.754	0.669	0.190	0.852	0.864
J48	0.780	0.396	0.771	0.780	0.774	0.403	0.761	0.761

Table 7. Stratified cross-validation

_iKLR	SMO	MLP	Naïve Bayes	RF	J48	
Total Number of Instances	**2499**	2499	2499	2499	2499	2499
Kappa	**0.4543**	0.4561	0.4334	0.4122	0.0957	0.4012
MAE	**0.261**	0.1977	0.2657	0.2801	0.3223	0.2648
RMSE	**0.3631**	0.4446	0.3764	0.4811	0.3829	0.4133
RAE	**67.72%**	51.29%	68.94%	72.67%	83.64%	68.72%

Figure 2. Integrated-KLR Margin Curve

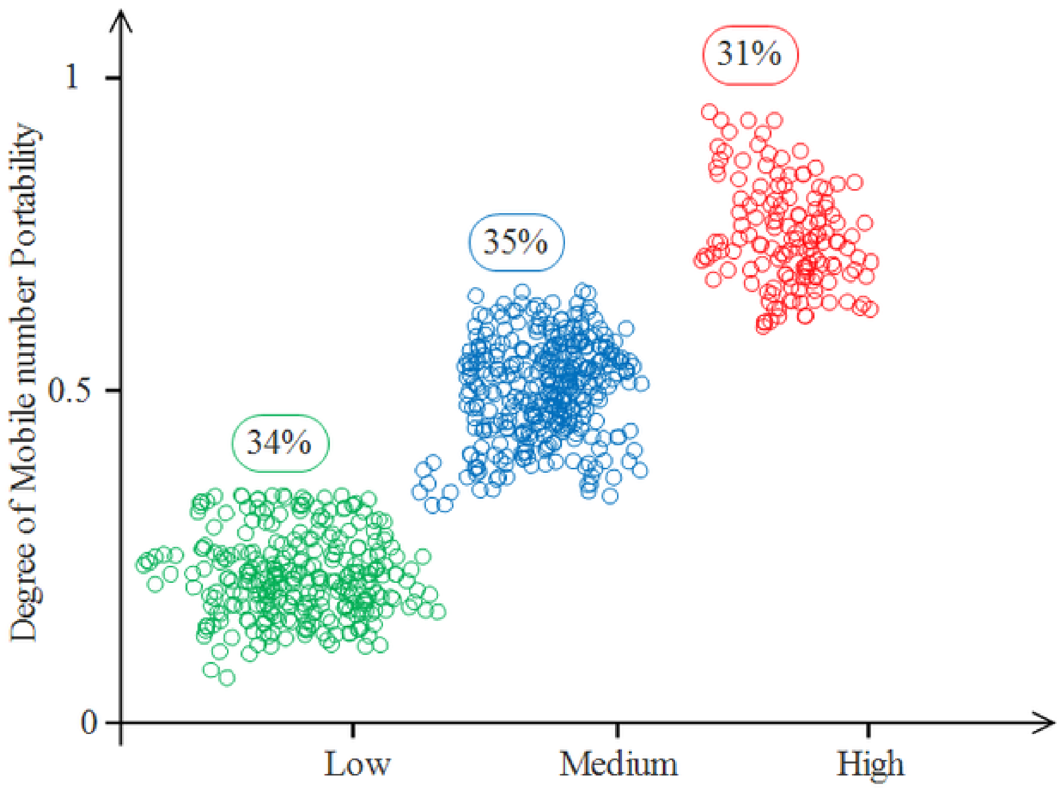


customers are likely to switch to another service provider due to a variety of factors such as poor network service, high call/message rate, and so on. Whereas 34% of the customers have a low portability level. 35% of the customers find themselves between a low and high degree of portability. Understanding the degree of portability will significantly reduce the cost spent on retaining porting customers by focusing more on most likely porting customers rather than wasting resources on the less vulnerable customers to mobile number portability as presented in Figure 3.

5. CONCLUSION

In this study, the application of iKLR is explored in predicting customer mobile number portability in the mobile telecommunication sector. This MNP technique in telecommunications is considered 'customer churn in disguise' because no precise study has previously dealt with the problem in this perspective other than discussions about the benefits of MNP to mobile telecommunications operators' customers. As a result, such previous studies may be said to have taken a customer-centric rather than provider-centric approach. Nevertheless, the focus of this study was the determination of the degree of churning using the iKLR clustering algorithm. To evaluate the results of the proposed approach, a benchmarking study was applied to five differently used algorithms (SOM. MLP, Naïve Bayes, RF, J48) in predicting customer churning. It was found that the proposed iKLR outperforms

Figure 3. Degree of mobile number portability cases using iKLR



the other algorithms in terms of ROC and PRC area. It is important to note that this study employs a dataset from a specific mobile telecommunications operator in Nigeria; thus, the results may differ when compared to datasets from other operators.

REFERENCES

Ahmed, A. A. Q., & Maheswari, D. (2017). Churn prediction on huge telecom data using hybrid firefly based classification Churn prediction on huge telecom data. *Egyptian Informatics Journal*, *18*(3), 215–220. doi:10.1016/j.eij.2017.02.002

Amin, A., Al-Obeidat, F., Shah, B., Adnan, A., Loo, J., & Anwar, S. (2019). Customer Churn Prediction in Telecommunication Industry under Uncertain Situation. *Journal of Business Research*, *94*, 290–301. doi:10.1016/j.jbusres.2018.03.003

Amin, A., Anwar, S., Adnan, A., Nawaz, M., Alawfi, K., Hussain, A., & Huang, K. (2017). Customer churn prediction in the telecommunication sector using a rough set approach. *Neurocomputing, 237*(May), 242–254. 10.1016/j.neucom.2016.12.009

Crawford, B., Khayyam, H., Milani, A. S., & Jazar, R. (2019). Big data modelling approaches for engineering applications. In R. N. Jazar (Ed.), *Nonlinear approaches in engineering applications*. Springer.

De Bock, K. W., & Van Den Poel, D. (2011). An empirical evaluation of rotation-based ensemble classifiers for customer churn prediction. *Expert Systems with Applications*, *38*(10), 12293–12301. doi:10.1016/j.eswa.2011.04.007

De Caigny, A., Coussement, K., & De Bock, K. W. (2018). A new hybrid classification algorithm for customer churn prediction based on logistic regression and decision trees. *European Journal of Operational Research*, *269*(2), 760–772. doi:10.1016/j.ejor.2018.02.009

Francois-Lavet, V., Henderson, P., Islam, R., Bellemare, M. G., & Pineau, J. (2018). An introduction to deep reinforcement learning. *Foundations and Trends in Machine Learning*, *11*(3–4), 219–354. doi:10.1561/2200000071

Gans, J. S., King, S. P., & Woodbridge, G. (2001). Number to the people: Regulation, ownership and local number portability. *Information Economics and Policy*, *13*(2), 167–180. doi:10.1016/S0167-6245(00)00039-1

Hamid, K., Bahman, J., Mahdi, J., & Reza, N. J. (2020). *Artificial Intelligence and Internet of Things for Autonomous Vehicles*. https://www.researchgate.net/publication/33502181310.1007/978-3-030-18963-1_2

Höppner, S., Stripling, E., Baesens, B., Broucke, S., & Verdonck, T. (2020). Profit driven decision trees for churn prediction. *European Journal of Operational Research*, *284*(3), 920–933. doi:10.1016/j.ejor.2018.11.072

Huang, B., Kechadi, M. T., & Buckley, B. (2012). Customer churn prediction in telecommunications. *Expert Systems with Applications*, *39*(1), 1414–1425. doi:10.1016/j.eswa.2011.08.024

Huang, B. Q., Kechadi, T. M., Buckley, B., Kiernan, G., Keogh, E., & Rashid, T. (2010). A new feature set with new window techniques for customer churn prediction in land-line telecommunications. *Expert Systems with Applications*, *37*(5), 3657–3665. doi:10.1016/j.eswa.2009.10.025

Idris, A., Rizwan, M., & Khan, A. (2012). Churn prediction in telecom using Random Forest and PSO based data balancing in combination with various feature selection strategies. *Computers & Electrical Engineering*, *38*(6), 1808–1819. doi:10.1016/j.compeleceng.2012.09.001

Karsmakers, P., Pelckmans, K., & Suykens, J. A. K. (n.d.). *Multi-class kernel logistic regression: A fixed-size implementation.* Academic Press.

Kisioglu, P., & Topcu, Y. I. (2011). Applying Bayesian Belief Network approach to customer churn analysis: A case study on the telecom industry of Turkey. *Expert Systems with Applications*, *38*(6), 7151–7157. doi:10.1016/j.eswa.2010.12.045

Kusy, M., & Zajdel, R. (2015). Application of reinforcement learning algorithms for the adaptive computation of the smoothing parameter for probabilistic neural network. *IEEE Transactions on Neural Networks and Learning Systems*, *26*(9), 2163–2175. doi:10.1109/TNNLS.2014.2376703 PMID:25532211

Lakshmana, P. M., & Ragupathy, R. (2019). A New sentiment score based improved Bayesian networks for real-time intraday stock trend classification. *International Journal of Advanced Trends in Computer Science and Engineering*, *8*(4), 1045–1055. Advance online publication. doi:10.30534/ijatcse/2019/10842019

Lin, Y. B., Chlamtac, I., & Yu, H. C. (2003). Mobile Number Portability. *IEEE Network*, *17*(5), 8–16. doi:10.1109/MNET.2003.1233913

Lomet, D. B. (2001). Bulletin of the Technical Committee on Data Engineering. *Bulletin of the Technical Committee on Data Engineering, 24*(4), 1–56. papers2://publication/uuid/30073F7F-1B7C-4496-ADA4-94FF4E6EE8F7

Mahdavinejad, K., & Mohammad, S. (2018). Machine learning for Internet of Things data analysis: A survey. *Digital Communications and Networks*, *4*(3), 161–175. doi:10.1016/j.dcan.2017.10.002

Maldonado, S., Flores, Á., Verbraken, T., Baesens, B., & Weber, R. (2015). Profit-based feature selection using support vector machines - General framework and an application for customer retention. *Applied Soft Computing*, *35*, 240–248. doi:10.1016/j.asoc.2015.05.058

Maria, C. A. A., Rosemarie, T. B., Jocelyn, O. T., & Daniel, E. M. Jr. (2020a). Analysis and System Architecture Design for Road Traffic Incidents. *International Journal of Advanced Trends in Computer Science and Engineering*, *9*(12). Advance online publication. doi:10.30534/ijatcse/2020/3391.22020

Maria, T. V., Eric, M., Riyanto, J., & Tuga, M. (2020b). Heart Disease Prediction System using Data Mining Classification Techniques: Naïve Bayes, KNN, and Decision Tree. *International Journal of Advanced Trends in Computer Science and Engineering*, *9*(3), 3028–3035. Advance online publication. doi:10.30534/ijatcse/2020/82932020

Nwankwo, W., & Ukhurebor, K. E. (2020a). Data Centres: A Prescriptive Model for Green and Eco-Friendly Environment in the Cement Industry in Nigeria. *International Journal of Scientific and Technology Research*, *9*(5), 239–244.

Nwankwo, W., & Ukhurebor, K. E. (2020b). Web forum and social media: A model for automatic removal of fake media using multilayered neural networks. *International Journal of Scientific and Technology Research*, *9*(1), 4371–4377.

Olayinka, A., Nwankwo, W., & Olayinka, T. C. (2020). Model based Machine Learning Approach to Predict Thermoelectric Figure of Merit. *Archive of Science and Technology*, *1*(1).

Rajeswari, R. P., Juliet, K., & Aradhana, N. (2017). Text Classification for Student Data Set using Naive Bayes Classifier and KNN Classifier. *International Journal of Computer Trends and Technology*, *43*(1), 8–12. doi:10.14445/22312803/IJCTT-V43P103

Vafeiadis, T., Diamantaras, K. I., Sarigiannidis, G., & Chatzisavvas, K. C. (2015). A comparison of machine learning techniques for customer churn prediction. *Simulation Modelling Practice and Theory*, *55*, 1–9. doi:10.1016/j.simpat.2015.03.003

Verbeke, W., Dejaeger, K., Martens, D., Hur, J., & Baesens, B. (2012). New insights into churn prediction in the telecommunication sector: A profit driven data mining approach. *European Journal of Operational Research*, *218*(1), 211–229. doi:10.1016/j.ejor.2011.09.031

Ayodeji S. Makinde holds a PhD in Computer Science with bias in Artificial Intelligence and Software systems. He is currently a lecturer in the Department of Computer Science, Edo University Iyamho Nigeria. He is a member of several learned societies.

Abayomi O. Agbeyangi is currently working towards the completion of a PhD in Computer Engineering with bias in machine learning and artificial intelligence. He is a Lecturer in the Department of Computer Sciences at Chrisland University Abeokuta Nigeria. He has authored several research papers published in local and international journals. He is a member of several learned professional societies.

Wilson Nwankwo is a multidisciplinary professional with over 17 years professional/industrial experience cutting across Computing, Information Technology, System Engineering, Risk management, Healthcare, Portfolio management, Project Finance and Governance, Project management, and Budgeting. He currently works as the Director of Open Distance Learning at Edo University. He is also a Professor in the Department of Computer Science with bias in Cyber physical systems and Software engineering. Academically, He holds a PhD in Information Technology, MSc in Computer Science, MBA, Doctor of Optometry, BSc in Business Computing and LLB in Law. His research interests include Intelligent enterprise systems, AI, Cyber Security, Software engineering and Nanotechnology.