

GSJ: Volume 9, Issue 2, February 2021, Online: ISSN 2320-9186 www.globalscientificjournal.com

CUSTOMER RETENTION STRATEGIES USING DATA MINING METHODS: A REVIEW

Aroosa Javaid, Malik Mubashir Hussain, Dr. Hamid Ghoos

Department of Computer Science, Institute of Southern Punjab Multan Pakistan Aroosa Javaid, Ph # 0300-6869927, Email: <u>arosajavaid1@gmail.com</u> Malik Mubashir Hussain

KeyWords

Retention, Data Mining, Supervised, Unsupervised, Machine Learning, loyality, features

ABSTRACT

Customer retention has become one of the main priorities for business these days. The more they retain customers the better they survive in competitive market. The goal of customer retention programs is to help companies retain many customers. To retain customer it's a tough job because customer data is so large and imbalanced that company can't handled it manually. There is a need for the system that can automatically predict leaving customers. Many researchers in past have tried to solve this problem by using data mining techniques. In this review paper we are reviewed the literature related to customer retention using data mining techniques. The structure of this paper is based on supervised, unsupervised, and hybrid models of data mining. At the end we have discussed the limitations of the previous work and future directions for the upcoming researchers.

INTRODUCTION:

According to the organization, customers are the foundation of its success and sales, which is why companies are becoming increasingly concerned about the importance of achieving the satisfaction of their customers. There are two aspects of marketing in the retail market: customer retention and acquisition. Customer acquisition is the process of attracting new customers to the product and retention is the process of keeping the customer continuously buys the product. The business holds different client details about their visit, purchasing experience, and so on. There is a need to process this huge customer's data to predict leaving customers. While using different strategies to determine the user's interest, they may miss some features and therefore the different techniques would face challenges because of their higher spatial property. The issue of customer's retention has been well discussed and studied. There are many methods has been discussed initially with regard to improving marketing. However, methods go beyond achieving high performance and require some strategic approach. So with the introduction of Machine Learning Techniques, the problem of high size and missing performance can be easily solved with big data. There are a number of machine learning algorithms available e.g. supervised, unsupervised, and Semi supervised machine learning.

SIGNIFICANCE:

The business holds different customers details about their visit, purchasing experience, and so on. In predicting user interest, huge data is used. While using different strategies to determine the user's interest, they might miss some features and therefore the different techniques would face challenges because of their higher spatial property. If the business already predicts which customer is about to leave then they try to offer great services or reward loyalty to their customers. There are many methods has been discussed

initially with regard to improving marketing. However, methods are difficult to achieve high performance and require some strategic approach. So by using machine learning algorithms and feature selection methods key features are selected from huge data that helps the business increase customer retention.

BACKGROUND:

Customer retention means the willingness of an organization to combine business with a particular customer or to continuously adjust to their requirement. Retention can also be characterized as the willingness of love, identity, commitment, trust and client to recommended and repeat purchases. It is important for the success of any businesses to establish successful customer's relationship. For the growth of businesses, customer satisfaction, retention, good words of mouth and loyalty are important. In addition, existing research on online customer retention is limited. Many businesses currently fail to attract new customers in order to maintain a suitable marketing department and assign mangers to pay attention to their existing customers. Most previous studies have concentrated on customer loyalty in restaurants, telecommunication firms, hotels and other services, although less attention has been given to customer retention in a retail industry.

DATA MINING

Data mining is a method used to predict customer retention. Data mining involves extracting information from a huge data and converting it into an easily interpretable system that enables organization to evaluate complex problems that result in customer loyalty and turn over to companies. Data mining provided many algorithms that we can use in the prediction of customer retention.

Supervised Machine Le	earning	Unsupervised Machine Le	earning
a. Class	ification	c. Clusteri	ng Algorithm
	i. Naïve Bayes		Hierarchal Clustering
	ii. Random Forest	ii.	Gaussian Mixture
i	ii. Nearest Neighbor	iii.	Neural Networks
i	v. Discriminant Analysis	iv.	C-Means
	v. Support vector Machine	v.	Fuzzy
b. Regro	ession	vi.	Hidden Markov Model
	i. Linear Regression GLM	vii.	K-medoids
	ii. SVR, GPR	viii.	K-Means
i	ii. Ensemble Methodology	ix.	KNN
i	v. Decision Tree		
	v. Neural Network		
\ \	<i>v</i> i. Boosting		
v	ii. Bagging		
vi	ii. Stacking		

Supervised

Naveen et.al (2020) predicted customer churn and relation model by using Machine Learning methods on cell2cell and IBM datasets.

They had predicted 30 variables and implemented Naïve Bayes, Decision Tree (DT) and Support Vector Machine (SVM). The outcome of their model was measured by using the Area under the curve (AUC) and gained 0.87, 0.82, and 0.77 for IBM and 0.98, 0.99 and 0.98 for the cell2cell dataset. On the other hand, **Hyun et.al (2020)** utilized the US-based telecommunication company dataset for the analysis of customer switching behavior and they implemented data cleaning and data balancing Pre-processing techniques. They had utilized some machine learning approaches Logistic regression (LR), Vector machine, Random Forest (RF) and DT. The outcomes of their model demonstrate that the predicted performance of their work is greater than 86% and the LR has the highest accuracy rate. They suggested using other data mining methods to create a predictable model similar to the ANN and Bayesian networks that were able to improve their research.

Hemlata et.al (2020) classifies the customer churn by using Supervised Machine Learning. They utilized some preprocessing strategies for data acquisition and cleaning. They grouped the dataset into training and testing that is acquired from an online source. They had implemented KNN (K-nearest Neighbors) and XGBoost Algorithms. Thus as that conclusion, KNN gains 83.85% accuracy with high error and low sensitivity and specificity while on the other hand, XGBoost gains accuracy of 86.85% with low error and high sensitivity and clarity. While, **Shreyas Rajesh (2020)** implemented machine learning algorithms, Extra Trees Classifier, XGBoosting Algorithm, SVM, SGD Classifier, AdaBoost Classifier, Gaussian Naive Bayes and LR to recognize the customer churn. He used BigML churn Telecom dataset and utilized some preprocessing strategies handling missing values, encoding categorical data, dividing the dataset into test and train sets, and feature scaling. The detection of their work was the Extra Tree classifier, XGBoosting Algorithm, and SVM was the finest with AUC score (0.8439, 0.78, and 0.735) respectively.

Sahar F. Sabbe (2018) used a public dataset of customers in the Telecom industry, took 60% of training and 40% of the testing dataset. He implemented machine learning techniques, DT (CART), SVM, KNN, Adaboost, RF, Stochastic gradient boost, and MLP ANN and upholds some pre-processing techniques like data cleaning, transformation, and feature selection. Thus the outcome of their work that RF and AdaBoost output performance is the same with the accuracy of 96%. Multilayer Perceptron and SVM give an accuracy of 94%. DT gives 90%, Naive Bayesian 88%, and LR and LDA give 86.70%. Hence RF and Adaboost give the best output performance. This research is extended by including deep learning and a hybrid model. On the other hand, **Nagraj et.al (2018)** purposed customer retention by upholding some machine learning algorithms ANN, SVM, and Deep Neural Network DNN and they took two datasets one of German credit second of a bank customer. They had applied pre-processing technique Normalization in the customer bank dataset. They had achieved the accuracy of 98%, 92%, and 97% for ANN, SVM, and DNN respectively and for bank customer data and 72%, 76% and 72%, respectively for the credit dataset of German. ANN gives better accuracy for bank customer data while DNN gives better accuracy for the German credit dataset.

Machine learning algorithms KNN, CART, and SVM used by **Pushpa et.al (2019)** for customer retention. They had utilized Normalization on the customer dataset which is self-generated. The outcomes indicate that the KNN classifier is better than CART and SVM. The accuracy of KNN, CART, and SVM are 0.96, 0.95, and 0.94 respectively. They suggested that their work can be used to find applications in the 5th Generation of mobile communication where user retention is important. On the other hand, **Vijaya et.al (2019)** proposed feature selecting techniques by using two different datasets from French Telecommunication orange company KDD Cup (2009) and implemented some pre-processing data cleaning and random oversampling. They used DT, KNN, SVM, NB, and ANN. The outcome indicates that SVM gains higher accuracy of 91.66% with imbalanced data and KNN gains 93.9% with random oversampling. They had suggested in the future more advanced methodology is used for feature selection.

Zhen et.al (2012) Proposed Hierarchical Multiple Kernel Support Vector Machine (H-MK-SVM) for the customer churns using longitudinal behavioral data. They used three datasets extracted from AdventureWorks DW, real-world databases Foodmart 2000 and Telecom. They had implemented sampling and Normalization. The outcome of their Proposed was H-MK-SVM shows Superior performance on both imbalanced and balanced data as compare to SVM and MK-SVM. The accuracy of H-MK-SVM, SVM and MK-SVM, on balanced data is 76.31%, 70.0% and 70.65%, respectively. While XGBoost method is used by **Atallah et.al (2020)** for customer retention. They used the self-generated dataset of 5000 subscribers and perform some sampling methods oversampling, ADASYN, SMOTE, and Borderline –SMOTE. The outcomes of their research were oversampling method improved the performance of Gradient Boosted Trees 84% by SMOTE oversampling of ratio 20%. On the other hand, **Ahmet et.al (2020)** used a machine algorithm for user behavior prediction. They have used the self-generated dataset of 50 features and apply pre-processing feature encoding, feature extraction, pseudonymization, missing feature, and normalization. The outcomes of their work that the Gradient boosting algorithm performs better than other algorithms.

Machine Learning methods DT, Gradient boosted Machine Tree (GBM), RF, and Extreme Gradient boosting (XGBoost) used by **Abde-Irahim et.al (2019)** for prediction of churn in Telecom. They had used SyriaTel telecom company dataset and apply some preprocessing techniques oversampling and Undersampling. The result of their work is that XGBoost gives a better performance of 93%. On the other hand, **Afaq et.al (2010)** implementing data mining in ISP for customer churn. They used a dataset from Spenta Co. and utilized some pre-processing techniques feature extraction and modeling. They uphold some methods DT, LR and NN algorithm. The outcome indicated that their model achieved an accuracy of 89.08% in churn prediction rate by using feed-forward Neural Networks. while, **Vafeiadis et.al (2015)** analyzed the comparison of machine learning by using ANN, SVMs, DTs, Naïve Bayes, and LR and they used a public domain dataset. The outcomes indicate that support machine classifier poly with AdaBoost gain accuracy of 97% and Fmeasure over 84%.

Kiansing et.al (2001) proposed customer retention via data mining they used dataset from a transactional database residing in Oracle and performed deviation analysis and feature selection techniques on the dataset. They had utilized DT induction. This model gives the proposal that data mining gives a good consequence for customer retention. This Model further demonstrates that the context of particle application of data mining is much an art. On the other hand, **Ridwan et.al (2015)** operated NN Analysis, Multiple Regression Analysis, and LR Analysis for the customer churn prediction and implemented feature extraction based on nine-month of the bills and normalization on the dataset that is extracted from the warehouse of a Telecommunication company. The outcome of their experiment is that NN gain the best accuracy of 91.28%.

Data mining techniques DT and CART (Classification and regression Tree) used by **Chitra et.al (2011)** for the customer retention of the bank. They had used the bank dataset and implemented data reduction on the given dataset. The outcome of their work that CART (Classification and regression Tree) gives the overall greater classification rate. On the other hand, RF and Regression Forest used by **Bart et.al (2005)** for customer retention and profitability. They make use of a dataset from warehouse of a large European financial Service Company and applied data validation and sampling on it. They conclude their model provided a better estimate and validation sample as a match to the regression model.

Preeti et.al (2016) used LR and DT in the Telecom industry for the customer churn prediction and they utilized the dataset from Telecom Company and applied data cleaning for making the system robust and feature extraction for the generation rules for the DT and estimation of parameters in LR The outcome of their work was by using both DT and LR is best to design customer retention and it will also easy to design or maintain the customer with a high probability to churn. On the other hand, UCI database of the University of California home telecommunication dataset was used and some preprocessing techniques data acquisition and oversampling also applied on it by **XIA et.al (2008)** for customer churn prediction. They had utilized SVM, ANN, DT, Naïve Bayesian Classifier, and LR, the outcome of their work were SVM give the better accuracy rat, strong generation ability, and good fitting Precision. Yu Zhao at. el (2005) presented improved one-class SVM for customer churn and used the subscriber database provided by the Oracle dataset. They compare their model with traditional methods DT, ANN, and Naïve Bayesian Classifier and the accuracy rate of their comparison SVM, ANN, DT, and Naïve Bayesian Classifier gain were 78.1%, 62%, 83.24%, and 87.15% respectively. Their Model shows a better accuracy rate than the other. They suggested that more research be done on how to select the appropriate kernel parameters and input features to get accurate results. While, **Mohammed al. el (2015)** used a dataset of UK mobile Telecommunication operator data of warehouse and they applied some data preparation processes discretisation of numerical variables, imputation of missing values, transformation from one set off discrete values to another, new variable derivation, and feature selection of the most informative variables. They had utilized LR and DT and the outcome of their work DT was Preferable for the customer churn and accuracy rate of DT was 70% and LR was 68%.

Abbas et.al (2016) used Electronic bank customer data from the bank's database and applied it to preprocess data cleaning, feature selection, and sampling on it. They used DT and the accuracy of DT was 99.70%. on the other hand, Applications of AdaBoost (Real, General, and Modest) utilized by **Shao et.al (2007)** for the churn prediction and they used a dataset provided by anonymous bank in China. They implemented some preprocessing handling of missing values and sampling. As a result of their model, these algorithms are proven improved for the predicting accuracy than the other algorithm.

Yaya et.al (2009) used improved balanced RF for churn prediction and they used the real bank customer dataset. They applied a sampling technique to the given dataset. The outcome of their research IBRF produced better accuracy than the other RF algorithm (balanced and weighted random forest). It offered great potential due to scalability, running speed, and faster training. On the other hand, Lonut et.al (2016) presented churn prediction for pre-paid mobile industry by using Neural Network, SVM, and Bayesian Network. They used the dataset of pre-paid mobile telecommunication industry and implemented data modeling on it. The outcome of their model was overall accuracy of 99.10% for Bayesian Network, 99.55% for NN, and 99.70% for SVM.

Suban et.al (2016) presented customer retention of MCDR by using data mining approaches Naïve Bayes, Radial Basis Function Network, Random Tree, and J48 Algorithm. They used two different datasets Dataset PAKDD 2006 available in "dat" and dataset of 3G network by an Asian Telco operator. They had applied some pre-processing handling missing values and Chi-square feature selection on the given datasets. The outcome of their research was Random Tree with three staged classifiers Chi2 gives the high accuracy rate of 87.67%. while, **LI-SHANG et.al (2006)** proposed knowledge discovery for customer churn by using DT and Logistic Regression. They utilized had customer dataset from the data warehouse and implemented sample extraction on it. The outcome of their model was DT performed better than the Logistic Regression. Similarly, **Ali et.al (2004)** proposed customer churn and retention by using Simple DT, DT with cost-sensitive DT, Boosting, and Logistic Regression. They used the customer dataset of major Australian online fastmoving consumer goods and implemented sampling. The outcome of their proposed model was AUC measurement shows different performance by the area under the ROC curve are 0.83, 0.91, 0.85, and 0.92 for Cost-Sensitive DT, LR, Simple DT and Boosting respectively.

Hongmei et.al (2008) Proposed GA- based Naïve Bayesian Classifier for the prediction of customer retention. They used dataset from Japanese credit debit and credit companies. They used a genetic algorithm for feature selection and the outcomes was compared with NB, TAN, and APRI classifier and indicate GA-based identified customer churn is better for a large number of customers shows higher classifying precision. On the other hand, Ali et.al (2015) predicted customer churn by using Echo State Network (ESN) and SVM. They used two different datasets one KDD dataset and the other from a Jordanian cellular telecommunication company. The outcome of their work was accuracy rate for the KDD dataset of ESN with SVM-readout was 93.7% and 87.9% and for the dataset from Jordain cellular telecommunication company was 99.2% and 96.8% respectively.

Zaho et.al (2008) presented a support vector machine for churn prediction and compares their result with the DT (C4.5), Logistic regression, ANN, and Naïve Bayesian classifier. They used a dataset of VIP customer's domestic branch of CCB as the core data and implemented sampling. The output of their work SVM with the higher accuracy rate of 0.5974, 0.5148 for C4.5, 0.5890 for LR, 0.5549 for Naïve Bayesian Classifier, and 0.5479 for ANN. On the other hand, **Shin et.al (2006)** proposed telecom churn management by using DT and Neural networks. They utilized random sampling on a dataset gained from wireless telecom companies in Taiwan. The output of their work was Both DT and NN give accurate churn prediction. Similarly, **Michael et.al (2000)** predicted improving retention and subscriber dissatisfaction by using LR, DT, NN, and Boosting. They used the dataset provided by the wireless carriers and utilized the outcome of their prediction was NN gives better nonlinear structure in the sophisticated representation than the DT, LR, and Boosting.

Jae et.al (2005) detected the change in customer behavior by using DT and used Dataset from a Korean online shopping mall. They had utilized data cleaning and discretization. The output of their detection was DT based methodology can be used to detect the changes of customer behavior from sales data and customer profile at different times. They suggested in future, this methodology can be extended to discover changes in customer behavior for three or more dataset. On the other hand, **Kristof et.al (2017)** proposed customer churn prediction by using LR and used Dataset provided by a large European mobile telecommunication provider. They implemented data cleaning, data reduction, sampling, missing value handling, and outlier. The outcome of their proposed was LR is a more advanced and assembled algorithm. It correctly prepares data generally less cumbersome than other algorithms.

Aurelie (2006) proposed bagging and boosting classification and used a dataset provided by Teradata Center at Duke University trees to predict customer churn. They used oversampling and the outcome of their proposed as boosting and bagging provided better classifiers than a binary logistic model. Predicted churn performance gain 16% of gini coefficient and 26% for top -docile lift. Both are provided good diagnostic measures, partial dependence plots, and variable importance. On the other hand, **Dudyala et.al (2008)** predicted credit card customer churn by using Multilayer Perceptron (MLP), DT (J48), LR, RF, SVM, and Radial Basis Function (RBF) Network. They performed oversampling, Undersampling, and synthetic minority oversampling on dataset from Latin American bank. The output of their prediction model gives the best result for under and oversampling and when original data is synthetic minority oversampling. Synthetic minority oversampling produced excellent results with 91.90% overall accuracy.

Kristof et.al (2006) predicted churn in subscription services by using the application of SVM and compare parameters with LR and random forest. They used dataset from a Belgian newspaper publishing company and implemented randomly Undersampling. The outcome of their predicted are SVM display fine performance when applied to a new, noisy marketing dataset and the accuracy for a real test set of SVM(SVMacc and SVMauc) are 84.90% and 85.14 respectively, LR is 84.60% and are 87.21%. They suggested deriving a solution to select the correct kernel and parameter value according to the problem is an interesting topic for future research.

Anuj et.al (2011) predicted customer churn in cellular network services by using NN and used dataset from the UCI repository database at the University of California. The output of their prediction was NN can predict the customer churn with an accuracy of 92.35%. They suggested we should propose Pre-processing and we also implemented deep methods. On the other hand, **Weiyun** et.al (2008) proposed a model for the prevention of customer churn and used dataset provided by Chines Bank. They implemented sampling on the given dataset. They had utilized improved balance RF and compare their result with DT, ANN, and CWC-SVM. The outcome of their proposal was improved balance RF gives highest accuracy of 93.4% and DT, ANN and CWC-SVM gives 62%, 78.12%, and 87.15% respectively.

Thomas et.al (2014) presented profit optimizing customer churn by using Bayesian network classifiers (Naïve Byes Classifier, Bayesian Network, Augmented Naive Bayes classifiers, Search-and-score algorithms, General Bayesian network classifiers, Constraint-based

algorithms, and Hybrid methods). They used four real-time datasets three datasets from the center for customer relationship management at Duke University, one from a European telecommunication operator, and one synthetic dataset. They used some preprocessing techniques feature selection (to limit the available attributes and Markov Blanket based algorithm for feature selection is used) and Undersampling (removing non-churn from the dataset). The outcome of their work was classification performance was measured area under the receiver operating characteristic curve (AUC) and the maximum profit (MP) criterion and both give a different ranking of the classification algorithm. The Naive Bayes methods do not lead to compatible networks, while the General Bayesian Network algorithms lead to simpler and more flexible networks and the Bayesian Network dividers cannot bypass traditional planning. On the other hand, **Koh et.al (2019)** predicted customer churn by using particle swarm optimization and extreme Learning Machine. They used the telecommunication dataset from Kaggle and implemented random oversampling and feature scaling. The outcome of their model retention plan can be built based on these features to reduce the churn rate from a company. The non-scaled feature gives training accuracy of 50.03% and testing accuracy of 49.95% similarly scaled feature gives training accuracy of 84.71% and testing accuracy of 81.16%. PSO and ELM determined high testing accuracy.

Author/ Year		Dataset		Pre-processing		Methods	Result	Future Work.
Year 2020	۶	cell2cell data-	٨	30 Predicted	۶	Naive Bayes	Outcome meas-	
Mr. V Naveen Kumar		set		variables	۶	SVM	ured by (AUC) and	
Mr. G Ramesh Babu	۶	IBM dataset			۶	decision tree	achieved 0.82, 0.87	
Mr. A Ravi Kishore						(DT)	and 0.77 for IBM	
			1				and 0.98, 0.99 and	
							0.98 for the	
							cell2cell dataset.	
Year (2020)		US-based tele-	۶	data cleaning		logistic regres-	Predicted perfor-	
Mohammed Al-		communication	٨	data balancing		sion (LR)	mance of their	
Mashraiea		company			≻	vector ma-	work is greater	
Hyun Woo Jeonb						chines	than 86% and the	
Sung Hoon Chunga					≻	random forest	LR has the highest	
						(RF)	accuracy rate.	
					۶	decision tree		
						(DT)		
Year 2020	≻	Dataset ac-		data Acquisition	>	XGBoost Boost-	KNN gains 83.85%	
Hemlata Dalmia		quired from an	\blacktriangleright	data Cleaning		ing	and XGBoost gains	
CH V S S Nikil		online source				XGBoost Algo-	86.85% accuracy	
Sandeep Kumar						rithm		
						K Nearest		
						Neighbors		
						(KNN)		

Shreyas Rajesh I Telecom data- ing Values sifler er, KGBoosting Algorithm, and Labhsetwar set > Encoding > XGBoosting Algorithm, and Subsetwar Set > Facture Scaling > SVM with AUC SVM with AUC Subsetwar > Feature Scaling > SVD Classifier 0.78, and 0.735) > Adaboost Classifier Sate - Value - Value - Value - Value - Value - Value Sahar F. Sabbe P Public dataset - Data transformation - Decision Tree Rf and Adaboost In future this Telecom indus- - Data transformation - Decision Tree Rf and Adaboost research is Sahar F. Sabbe - Goustomer in - Telecom indus- - Data cleaning - Stornastic grad - Or gof SK. Multi- extended by Tree - Stornastic grad - Categorization - Stornastic grad - Or gof SK. Multi- extended by Adabost - Telecom indus- - Eature selec- - KNN and SVM givea - Including deep Adapter - Stornastic grad - Telecom indus- - Stornastic grad - Dagive So%, Multi- - Van	year 2020	\blacktriangleright	BigML churn	\checkmark	Handling Miss-		Extra Trees Clas-	Extra Tree classifi-	
Year 2018 Sahar F. Sabbe> Public dataset of customer in Telecom indus- try> Data transfor- mation> Decision Tree (CART)RF and Adaboost ustance (CART)In future this respectively.Year 2018 Sahar F. Sabbe> Public dataset of customer in Telecom indus- try> Data transfor- mation> Decision Tree (CART)RF and Adaboost ustance (CART)In future this respectively.Year 2018 Sahar F. Sabbe> Public dataset of customer in Telecom indus- try> Data transfor- mation> Decision Tree (CART)RF and Adaboost ustance (CART)In future this respectively.Year 2018 Nagraj V. Dharwadka and Priyanka S. Patil> German Credit dataset> Normalization of ustance> Support Vector Machine (SNN)Normalization p2%, and 97% for and for bank cus- tomer data and 72%, 76% and 72%, respectively tor the credit dataset of German.> Normalization Pating and p2%, and 97% for and for bank cus- tomer data and 72%, 76% and 72%, respectively for the credit dataset of German.> Normalization Pating and for bank cus- tomer data and 72%, 76% and 72%, respectively for the credit dataset of German.> Normalization Pating and SVM the and tomer data and 72%, 76% and 72%, respectively for the credit dataset of German.> Normalization Pating and SVM the and tomer data and 72%, 76% and 72%, respectively for the credit dataset of German.> KNNKNN dassifier is data for tomer data and 72%, 76% and 72%, respectively for the credit dataset of German.> KNNKNN dassifier is data for tomer data	Shreyas Rajesh		Telecom data-		ing Values		sifier	er, XGBoosting	
Year 2018 > Public dataset > Data transfor- mation > Gaussian Naive Bayes RF and Adaboost In future this Year 2018 > Public dataset > Data transfor- mation > Decision Tree RF and Adaboost In future this Sahar F. Sabbe > Public dataset > Data transfor- mation > Decision Tree RF and Adaboost In future this Year 2018 > Public dataset > Data cleaning > SVM cy of 9%. Multi- extended by extended by Year 2018 > 60% of training 8.40% testing dataset > Normalization > Support Vector Mich the sacuracy of 94%. DT gives 90%. Year 2018 > German Credit dataset > Normalization > Support Vector Machine (SVM) accuracy of 98%., Machine (SVM) Public dataset Year 2018 > German Credit dataset > Normalization > Support Vector Machine (SVM) accuracy of 98%., Machine (SVM) Public dataset Year 2018 > German Credit dataset > Normalization > Support Vector Machine (SVM)	Labhsetwar		set	≻	Encoding		XGBoosting	Algorithm, and	
Year 2018 > Public dataset of customer in Sahar F. Sabbe > Public dataset of customer in Telecon indus- try > Data transfor- of customering > Decision Tree (CART) RF and Adaboost with same accura- cy of 96%. Multi- extended by In future this research is cov of 96%. Multi- extended by Year 2018 > Public dataset of customering > Data transfor- pota cleaning > Decision Tree (CART) RF and Adaboost with same accura- cy of 96%. Multi- extended by Year 2018 > 60% of training at 40% testing dataset > Data cleaning tion > KNN layer perceptron accuracy of 94%. Nalve Bayesian Nalve Bayesian accuracy of 98%, and IR and tub Ag we 86.70%. In future this research is customer data and Priyanka S. Patil Year 2018 > German Credit dataset > Normalization > Support Vector Machine (SVM) accuracy of 98%, accuracy of 98%, respectively for the credit dataset of German. Year 2018 > German Credit dataset > Normalization > Support Vector Machine (SVM) accuracy of 98%, respectively for the credit dataset of German. Year 2019 > Self generated pushpa Singh Vish- was Agrawal > Normalization > KNN KNN classifier is Normalization In future this vork can be vork can be Year 2019 > Self generated pushpa Singh Vish- was Agrawal <td< td=""><td></td><td></td><td></td><td>≻</td><td>Categorical Data</td><td></td><td>Algorithm</td><td>SVM with AUC</td><td></td></td<>				≻	Categorical Data		Algorithm	SVM with AUC	
Year 2018 > Public dataset > Data transformation > Decision Tree RF and Adaboost In future this Sahar F. Sabbe > Public dataset > Data transformation > Decision Tree RF and Adaboost In future this Sahar F. Sabbe > Public dataset > Data cleaning > SVM Cy of 96%. Multi- extended by research is Year 2018 > Of training & 40% testing dataset > Data cleaning > KNN Iaver proceptron including deep & 60% of training adtaset > Mormalization > Support Vector accuracy of 94%. DT gives 90%. Network (DNN) Year 2018 > German Credit dataset > Normalization > Support Vector accuracy of 98%. DT gives 90%. Proceeding Year 2018 > German Credit dataset > Normalization > Support Vector accuracy of 98%. Proceeding Year 2018 > German Credit dataset > Normalization > Support Vector accuracy of 98%. Proceeding Year 2018 > German Credit dataset > Normalization > Support Vector accuracy of 98%. Proceeding Year 2019 > Self generated usa Agrawal > Normalization > KNN CART <td></td> <td></td> <td></td> <td>≻</td> <td>Feature Scaling</td> <td></td> <td>SVM</td> <td>Score (0.8439,</td> <td></td>				≻	Feature Scaling		SVM	Score (0.8439,	
Year 2018 > Public dataset > Data transformation > Decision Tree RF and Adaboost In future this Sahar F. Sabbe > Public dataset > Data cleaning > SVM Cy of 96%. Multi- extended by Telecon indus- > Data cleaning > SVM Cy of 96%. Multi- extended by Year 2018 > 60% of training > Feature selec- > Adaboost and SVM and learning and Adayse > 60% of training > Stochastic gradient for the second of the secon							SGD Classifier	0.78, and 0.735)	
Year 2018 > Public dataset > Data transfor- Solution R and Adaboost In future this Sahar F. Sabbe of customer in mation (CART) with same accura- research is Sahar F. Sabbe of customer in mation (CART) with same accura- research is Year 2018 > 60% of training > Data cleaning > SVM cy of 96%. Multi- extended by Year 2018 60% of training tru > Feature selec- > KNN lager perceptron including deep dataset dataset Normalization > Machon forest accuracy of 94%. hybrid model. Nagraj V. Dharwadka of German Credit Normalization > Support Vector accuracy of 98%. Nagraj V. Dharwadka dataset Machine (SVM) 92%, and 97% for secure of 98%. Nagraj V. Dharwadka German Credit Normalization > Support Vector accuracy of 98%. Nagraj V. Dharwadka German Credit Ataset Network (DNN) DNN respectively receit dataset receret data and							Adaboost Clas-	respectively.	
Year 2018 > Public dataset > Data transfor- mation > Decision Tree RF and Adaboost In future this Sahar F. Sabbe of customer in Telecom indus > Data cleaning > SVM cy of 96%. Multi- extended by Viry > Feature selec- > KNN layer perceptron including deep 8<40% testing							sifier		
Year 2018 Sahar F. Sabbe> Public dataset of customer in Telecom indus- try> Data transfor- mation> Decision Tree (CART)RF and Adaboost with same accura- cy of 96%. Multi- extended by layer perceptronIn future this research is extended by layer perceptron> 60% of training & 40% testing dataset> Data cleaning to > 60% of training dataset> Mathematical accuracy of 94%. by feature selec- to adataset> KNNlayer perceptron accuracy of 94%. DT gives 90%, Naive Bayesian 88%, and LR and UDA give 86.70%.hybrid model. hybrid model.Year 2018 Nagraj V. Dharwadka and Priyanka S. Patil> German Credit dataset> Normalization to were accuracy of 98%, Pachema accuracy of 98%, Machine (SVM)3ccuracy of 98%, 92%, and 97% for 3ccuracy of 98%, 92%, and 97% for Anny. SVM, and Network (DNN)92%, and 97% for tomer data accuracy of 98%, 92%, and 60% for the credit dataset of credit dataset of credit dataset of set> Normalization Pachema accuracy of 98%, Pachema accuracy accuracy of 98%, Pachema accuracy of 98%, Pachema accuracy accuracy of 98%, Pachema accuracy accuracy of 98%, Pachema accuracy accura							Gaussian Naive		
Year 2018 > Public dataset > Data transformation > Decision Tree RF and Adaboost In future this Sahar F. Sabbe of customer in Telecom indus- > Data cleaning > SVM cy of 96%. Multi- extended by Telecom indus- > Data cleaning > SVM cy of 96%. Multi- extended by try > Feature selec- KNN layer perceptron including deep & 40% testing tion > Adaboost and SVM give an learning and & 40% testing tion > Adaboost accuracy of 94%. hybrid model. & dataset V Normalization > Support Vector Nalve Bayesian Nagraj V. Dharwadka dataset > Normalization > Support Vector accuracy of 98%, Nagraj V. Dharwadka dataset > Normalization > Support Vector accuracy of 98%, Nagraj V. Dharwadka dataset > Normalization > Support Vector accuracy of 98%, Nagrai Y. Dharwadka dataset <td< td=""><td></td><td></td><td></td><td></td><td></td><td></td><td>Bayes</td><td></td><td></td></td<>							Bayes		
Sahar F. Sabbeof customer in Telecom indus- trymation(CART)with same accura- cy of 96%. Multi- including deep including deep to 60% of training datasetmation(CART)with same accura- cy of 96%. Multi- including deep accuracy of 94%.research is extended by including deep accuracy of 94%.Year 2018>German Credit dataset>Normalization accuracy of 98%.Naive Bayesian B8%, and LR and tDA give 86.70%.Naive Bayesian B8%, and LR and tDA give 86.70%.Year 2018>German Credit dataset>Normalization accuracy of 98%.Support Vector ber participationaccuracy of 98%. Naive Bayesian B8%, and LR and tDA give 86.70%.Year 2018>German Credit dataset>Normalization accuracy of 98%.Support Vector ber participationaccuracy of 98%. tomer data and tomer data and tomer data and tomer data and tomer data and tomer data and tomer data and tower (ANN)Non respectively tomer data and tomer data and tomer data and tower (ANN)In future this work can be tower (ANN)Year 2019>Self generated set>Normalization tower data- set>KNNKNN classifier is tower (ANN)In future this tower (ANN)Year 2019>Self generated set>Normalization tower (ANN)>KNN classifier is tower (ANN)In future this tower (ANN)Year 2019>Self generated set>Normalization tower (ART tower (ART, and SVM areIn future this tower (ART, an							LR		
Year 2018>Sef german Credit>Normalization>Support Vectoraccuracy of 97%. Multi- including deep iarning and bybrid model.Year 2018>German Credit>Normalization>Support Vectoraccuracy of 97%. accuracy of 97%.Hore including deep iarning and bybrid model.Year 2018>German Credit>Normalization>Support Vectoraccuracy of 97%. accuracy of 97%.Feature selec- icer 1000000000000000000000000000000000000	Year 2018	>	Public dataset	>	Data transfor-		Decision Tree	RF and Adaboost	In future this
try> Feature selec- tion> KNNlayer perceptronincluding deep learning and hybrid model.& 40% testing dataset& 40% testing dataset> Random forest Stochastic gra- dient boostaccuracy of 94%.hybrid model.Year 2018> German Credit dataset> Normalization> Support Vectoraccuracy of 98%, dient boostaccuracy of 98%, dient boostNagraj V. Dharwadka and Priyanka S. Patil> German Credit dataset> Normalization> Support Vectoraccuracy of 98%, dient 60XM)Year 2018> German Credit dataset> Normalization> Support Vectoraccuracy of 98%, dient 60XM)AnN, SVM, and datasetNagraj V. Dharwadka and Priyanka S. Patil> German> Normalization> Support Vectoraccuracy of 98%, dient 60XM)Year 2018> German Credit dataset> Normalization> Support Vectoraccuracy of 98%, datasetActificial Neural respectivelyYear 2019> Self generated customer data was Agrawal> Self generated set> Normalization> KNNKNN classifier is better than CART quot caracy of KNN, applications in CaRT, and SVM are fifth Genera-	Sahar F. Sabbe		of customer in		mation		(CART)	with same accura-	research is
Year 2018 Nagraj V. Dharwadka and Priyanka S. Patil> German Credit A self generated a set> Normalization A set> Support Vector Adatasetaccuracy of 98%, DT gives 90%, Naive Bayesian B8%, and LR and LDA give 86.70%.Iearning and hybrid model.Year 2018 Nagraj V. Dharwadka and Priyanka S. Patil> German Credit dataset> Normalization Adataset> Support Vector Accuracy of 98%, Machine (SVM)accuracy of 98%, 92%, and 97% for > Deep Neural ANN, SVM, and Network (DNN)92%, and 97% for DNN respectively and for bank cus- tomer data and 72%, 76% and 72%, respectively for the credit dataset of German.Year 2019 Pushpa Singh Vish- was Agrawal> Self generated set> Normalization A Normalization> KNNKNN classifier is better than CART usd th Genera- fifth Genera-			Telecom indus-	≻	Data cleaning		SVM	cy of 96%. Multi-	extended by
& 40% testing dataset& Random forest Stochastic gra- dient boostaccuracy of 94%. DT gives 90%, Naive Bayesian 88%, and LR and LDA give 86.70%.hybrid model.Year 2018>German Credit dataset>Normalization Adataset>Support Vector Deep Neural Network (DNN)accuracy of 98%, D2%, and 97% for D2%, and 97			try	≻	Feature selec-		KNN	layer perceptron	including deep
Year 2018 > German Credit > Normalization > Support Vector accuracy of 98%, B8%, and LR and LDAgive 86.70%. Year 2018 > German Credit > Normalization > Support Vector accuracy of 98%, B8%, and LR and LDAgive 86.70%. Year 2018 > German Credit > Normalization > Support Vector accuracy of 98%, B8%, and LR and LDAgive 86.70%. and Priyanka S. Patil Ataset Puep Neural ANN, SVM, and Network (DNN) DNN respectively Antificial Neural and for bank cus- Network (ANN) tomer data and 72%, 76% and 72%, respectively for the credit dataset of German. 72%, 76% and 72%, respectively for the credit dataset of German. Year 2019 > Self generated > Normalization > KNN KNN classifier is In future this Year Agrawal set Normalization > KNN knn classifier is In future this Year 2019 > Self generated > Normalization > KNN knn classifier is In future this Year 2019 > Self generated		۶	60% of training		tion		Adaboost	and SVM give an	learning and
Year 2018>German Credit>Normalization>Support Vectoraccuracy of 98%, Machine (SVM)92%, and 97% forand Priyanka S. Patil>Marking (SVM)92%, and 97% forPeep NeuralANN, SVM, andAnd Priyanka S. PatilAnd Priyanka S. PatilAnd Priyanka S. PatilAnd Priyanka S. PatilAnd Priyanka S. Patil <td></td> <td></td> <td>& 40% testing</td> <td></td> <td></td> <td></td> <td>Random forest</td> <td>accuracy of 94%.</td> <td>hybrid model.</td>			& 40% testing				Random forest	accuracy of 94%.	hybrid model.
Year 2018 > German Credit > Normalization Support Vector accuracy of 98%, (Machine (SVM) 92%, and 97% for and Priyanka S. Patil - - Deep Neural ANN, SVM, and Network (DNN) DNN respectively - Antificial Neural and for bank cus- Network (DNN) - Network (ANN) tomer data and 72%, 76% and 72%, respectively for the Year 2019 > Self generated > Normalization > KNN Pushpa Singh Vish- was Agrawal - Self generated > Normalization > KNN Vector - - - - - - CART - - - - - - Pushpa Singh Vish- was Agrawal -			dataset				Stochastic gra-	DT gives 90%,	
Year 2018>German Credit>Normalization>Support Vectoraccuracy of 98%, Machine (SVM)92%, and 97% forand Priyanka S. Patildataset-Poep NeuralANN, SVM, and Network (DNN)92%, and 97% forAnd Priyanka S. PatilPoep NeuralANN, SVM, and Network (DNN)DNN respectively DNN respectivelyArtificial Neuraland for bank cus- Network (ANN)and for bank cus- tomer data and 72%, 76% and 72%, respectively for the credit dataset of German.Year 2019>Self generated>Normalization>KNNKNN classifier is better than CARTIn future this work can be 							dient boost	Naive Bayesian	
Year 2018> German Credit dataset> Normalization> Support Vector Machine (SVM)accuracy of 98%, 92%, and 97% forand Priyanka S. Patildataset> Deep Neural Network (DNN)ANN, SVM, and DNN respectively Artificial Neural Network (ANN)DNN respectively tomer data and 72%, 76% and 72%, respectively for the credit dataset of German.Year 2019> Self generated uss Agrawal> Normalization> KNNKNNKNN classifier is and SVM. The ac- used to find curacy of KNN, applications in CART, and SVM areIn future this mork of KNN, applications in fifth Genera-							MLP ANN	88%, and LR and	
Nagraj V. Dharwadka and Priyanka S. PatildatasetMachine (SVM)92%, and 97% for ANN, SVM, andand Priyanka S. PatilAdtasetPeep Neural Network (DNN)ANN, SVM, and DNN respectively and for bank cus- tomer data and 72%, 76% and 72%, respectively for the credit dataset of German.Year 2019>Self generated set>Normalization>KNNKNN classifier is and SVM. The ac- used to find and SVM. The ac- used to find applications in CART, and SVM areIn future this applications in fifth Genera-								LDA give 86.70%.	
and Priyanka S. Patil and Priyanka S. Patil Pushpa Singh Vish- Was Agrawal Metwork (DNN) Set Metwork (DNN) Metwork (DNN) Artificial Neural Network (ANN) Metwork	Year 2018	≻	German Credit	≻	Normalization	≻	Support Vector	accuracy of 98%,	
Year 2019> Self generated> Normalization> KNNKNNKNN classifier is better than CARTIn future this work can be better than CARTYear 2019> Self generated> Normalization> KNNKNNKNN classifier is better than CARTIn future this work can be better than CARTYear 2019> Self generated> Normalization> KNNKNN classifier is better than CARTIn future this work can be better than CARTYear 2019> Self generated> Normalization> KNNKNN classifier is better than CARTIn future this work can be is of the curacy of KNN, applications in CART, and SVM. The ac- is of the Genera-	Nagraj V. Dharwadka		dataset				Machine (SVM)	92%, and 97% for	
Year 2019> Self generated> Normalization> KNNKNN classifier isIn future thisPushpa Singh Vish- was AgrawalSet> Normalization> KNNKNN classifier isIn future thisGerman.Start of ind indication> SVMindication in indicationindication in indicationindication in indicationindication in indicationYear 2019> Self generated> Normalization> KNNindication in in indicationindication in indicationindication in in ifith Genera-	and Priyanka S. Patil						Deep Neural	ANN, SVM, and	
Year 2019 Pushpa Singh Vish- was Agrawal> Self generated set> Normalization (and set and be							Network (DNN)	DNN respectively	
Year 2019> Self generated> Normalization> KNNKNN classifier isIn future thisPushpa Singh Vish- was Agrawal< Set							Artificial Neural	and for bank cus-	
Year 2019> Self generated> Normalization> KNNKNN classifier isIn future thisPushpa Singh Vish- was Agrawal< set							Network (ANN)	tomer data and	
Year 2019> Self generated> Normalization> KNNKNN classifier isIn future thisPushpa Singh Vish- was Agrawalcustomer data- set> Normalization> KNNbetter than CARTwork can be used to find curacy of KNN, CART, and SVM areused to find ifth Genera-								72%, 76% and 72%,	
Year 2019> Self generated> Normalization> KNNKNN classifier isIn future thisPushpa Singh Vish-customer data-> CARTbetter than CARTwork can bewas Agrawalset- A> SVMand SVM. The ac-used to findLLLLLCARTfith Genera-								respectively for the	
Year 2019> Self generated> Normalization> KNNKNN classifier isIn future thisPushpa Singh Vish- was Agrawalcustomer data- set> CARTbetter than CARTwork can be used to findLset- LSVMand SVM. The ac- curacy of KNN,used to find applications in CART, and SVM arefifth Genera-								credit dataset of	
Pushpa Singh Vish- customer data- > CART better than CART work can be was Agrawal set > SVM and SVM. The ac- used to find L L Curacy of KNN, applications in CART, and SVM are fifth Genera-								German.	
was Agrawal set SVM and SVM. The ac- used to find curacy of KNN, applications in CART, and SVM are fifth Genera-	Year 2019		Self generated	≻	Normalization		KNN	KNN classifier is	In future this
curacy of KNN, applications in CART, and SVM are fifth Genera-	Pushpa Singh Vish-		customer data-				CART	better than CART	work can be
CART, and SVM are fifth Genera-	was Agrawal		set				SVM	and SVM. The ac-	used to find
								curacy of KNN,	applications in
0.96, 0.95, and tion (5G) of								CART, and SVM are	fifth Genera-
								0.96, 0.95, and	tion (5G) of

							0.94 respectively.	Mobile com-
								munication
								where user
								retention is
								importance.
Year 2019	≻	Dataset from		Data cleaning		DT	SVM gains higher	In future more
E.Sivasankar		French Tele-	≻	Random Over-		KNN	accuracy of 91.66%	advanced me-
J. Vijaya		communica-		sampling		SVM	with imbalanced	thods are used
		tion orange				NB	data and KNN gains	for feature
		company KDD				ANN	93.9% with ran-	selection tech-
		Cup (2009)					dom oversampling.	niques.
Year 2012	Thr	ree real-word	۶	Sampling	\triangleright	Hierarchical	Accuracy of H-MK-	
Zhen-Yu Chen	dat	abases datasets	۶	Normalization		Multiple Kernel	SVM, SVM, and	
Minghe Sun	≻	Foodmart 2000				Support Vector	MK-SVM on ba-	
Zhi-Ping Fan	≻	Adventure				Machine (H-MK-	lanced data are	
		Works				SVM)	76.31%, 70.0% ,	
	≻	DW Telecom			≻	SVM	and 70.65% re-	
					>	Multiple Kernel	spectively.	
						Support Vector		
						Machine (MK-		
						SVM)		
Year 2020	≻	Self generated		Over Sampling		XGBoost (Gra-	Performance of	
Atallah M. AL-		dataset consist	≻	Random -		dient Boosted	Gradient Boosted	
Shatnwai		the informa-		oversampling		Trees Algorithm)	Trees 84% by	
Mohammad Faris		tion of 5000	≻	SMOTE			SMOTE oversam-	
		subscribers	≻	ADASYN			pling of ratio 20%.	
			≻	Border line -				
				SMOTE				
Year 2020	>	Self generated	≻	Feature encoding		Machine leaning	The outcomes of	
Ahmet Turkmen		dataset consist	⊳	Feature extrac-		algorithm	their work that the	
Cenk Anil Bahcevan		of 35 features		tion			Gradient boosting	
Youssef Alkhanafseh				Pseudonymiza-			algorithm performs	
Esra Karabiyik				tion			better than other	
			\triangleright	Missing-feature			algorithms.	
			≻	Normalization				
Year 2019		SyriaTel tele-		Over sampling		RF	XGBoost gives a	
		Synale tele-		over sampling		INI	VODOOSE BINGS 9	

Abdelrahim Kasem	Τ	com company	۶	Under sampling		Gradient boosted	better perfor-
Ahmad		dataset				Machine Tree	mance of 93%.
Assef Jafar						(GBM)	
Kadan Aljoumaa						DT	
						Extreme Gradient	
						boosting	
						(XGBoost)	
Year 2010	≻	Dataset from	≻	Feature extrac-		DT	Accuracy of
Afaq Alam Khan		spenta co.		tion		LR	89.08% in churn
M.M Sepehri				Modeling		Neural Network	prediction rate by
Sanjay Jamwal				C		Algorithm (NN)	using feed-forward
						5 ()	Neural Networks.
Year 2015	>	Public domain	-			ANN	Support Machine
K.I. Diamantaras		dataset				SVM	classifier Poly with
G. Sarigiannidis K.Ch.						DT	AdaBoost gain ac-
Chatzisavvas						Naïve Bayes	curacy of 97% and
T. Vafeiadis						LR	F-measure over
							84%.
Year 2001	>	Dataset from		Deviation analy-		Decision Tree	Data mining gives a
Kiansing NG	ĺ	transactional		sis	Í	induction	good consequence
Huan Liu		database resid-	A	Feature selection			for customer re-
		ing in Oracle.	1				tention.
Year 2015	>	Warehouse of a		Feature Extrac-	Δ	Neural Network	NN gain the best
M.ridwan ismail	Í	telecommuni-	ĺ	tion	ſ	Analysis	accuracy of
M.Nordin a rehman		cation company	Δ	Normalization	Δ	Multiple Regres-	91.28%.
M.Khalid Awang		cation company	ĺ	Normalization	ſ	sion Analysis	51.20%.
Mokhairi Makhtar					D	-	
						Logistic Regres-	
Voor 2011	~	Dank datasat		Data raduatian	~	sion Analysis	The outcome of
Year 2011	>	Bank dataset		Data reduction		DT	their work that
K. Chitra						CART(Classificati	
B. Subashini						on and regres-	CART (Classifica-
						sion Tree)	tion and regression
							Tree) gives the
							overall greater
	Ļ						classification rate.
Year 2005	≻	Warehouse of a		Data validation		Random Forest	They conclude
Dirk Van den Poel		large European		Data sampling		(RF)	their Model pro-
					1		

Bart Lariviere		financial Ser-				Regression Forest	vided a better es-	
		vice Company					timate and valida-	
							tion sample as a	
							match to the re-	
							gression model.	
Year 2016	>	Telecom com-	≻	Data cleaning		Logistic Regres-	DT and LR is best	
Siddhi K. Khandge		pany dataset		Feature extrac-		sion (LR)	to design customer	
Preeti k.Dalvi				tion		DT	retention and it	
Aditya Bankar							will also easy to	
Ashish Deomore							design or maintain	
Prof V.A Kanade							the customer with	
							a high probability	
							to churn.	
Year 2008	≻	UCI database of	۶	Data acquisition	۶	SVM	the outcome of	
XIA Guo-en		University of	۶	Over-sampling	۶	Artificial Neural	their work were	
JIN Wei-dong		California				Network (ANN)	SVM give the bet-	
		Home Tele-			۶	DT	ter accuracy rat,	
		communication				LR	strong generation	
		dataset	\mathbb{Z}			Naïve Bayesian	ability, and good	
)	- (-		Classifier	fitting Precision.	
Year 2005	≻	Subscriber da-	-		>	ANN	SVM, ANN, DT, and	They suggested
Yu Zhao, Bing Li		tabase pro-			►	DT	Naïve Bayesian	that more re-
Xiu Li		vided by Oracle			۶	Naïve Bayesian	Classifier gain were	search be done
Wenhuang Liu		dataset				Classifier	78.1%, 62%,	on how to se-
Shouju Ren					۶	SVM	83.24%, and	lect the appro-
							87.15% respective-	priate kernel
							ly.	parameters
								and input fea-
								tures to get
								accurate re-
								sults.
Year 2015	>	UK mobile Tel-	≻	Imputation of	≻	Logistic Regres-	the outcome of	
Mohammed Has-		ecommunica-		missing values		sion	their work DT was	
souna		tion operator	\blacktriangleright	Discretisation of		Decision Tree	Preferable for the	
Tariq Elyas		data ware-		Numerical va-			customer churn	
Mohammad Saeed		house		riables			and accuracy rate	
1	1		1		1		1	

AbouTrab				Feature Selection			of DT was 70% and
Ali Tarhini				of the most in-			LR was 68%.
				formative va-			
				riables			
				New Variable			
			<i>.</i>	derivation			
				Transformation			
			ĺ	from one set off			
				discrete values to			
				another			
				another			
Year 2016	۶	Electronic bank	۶	Data Cleaning	≻	Decision Tree	They used DT and
Abbas Keramati		customer data		Feature Selection			the accuracy of DT
Hajar Ghaneei		from bank's da-		Sampling			was 99.70%.
Seyed Mohammad		tabase					
Mirmohammadi							
Year 2007	≻	Dataset Pro-		Handling of miss-	A	daBoost	These algorithms
Shao Jinbol		vided by Ano-		ing values		real-AdaBoost,	are proven im-
Li Xiu		nymous bank in	4	Sampling		Gental AdaBoost	proved for the
Liu Wenhuang		China				Modest Ada-	predicting accuracy
						Boost	than the other
							algorithm.
Year 2009	≻	Real bank cus-	۶	Sampling		Improved ba-	IBRF produced
Xiu Li		tomer dataset				lanced Random	better accuracy
Yaya Xie						Forest	than the other RF
Weiyun Ying							alogorithm (ba-
E.W.T. Ngai							lanced and
							weighted random
							forest).
Year 2016	≻	Dataset of pre-		Data Modeling		Neural Network	Accuracy of
Ionuț Brândușoiu		paid mobile te-				SVM	99.10% for Baye-
Gavril Toderean		lecomm-				Bayesian Net-	sian Network,
Horia Beleiu		unication in-				work	99.55% for NN,
		dustry					99.70% for SVM.
Year 2016	≻	Dataset PAKDD		Handling Missing		Naïve Bayes	Random Tree with
Suban Ravichandran		2006 available		Values		Radial Basis	three staged clas-
Chandrasekaran		in "dat"		Chi-square		Function Net-	sifiers Chi2 gives
Ramasamy	\triangleright	Dataset of 3G		·		work	the high accuracy
						-	

		Network by an	۶	Feature Selection		Random Tree	rate of 87.67%.
		Asian Telco				J48 Algorithm	
		Operator					
Year 2006	≻	Customer data-	\checkmark	Sample extrac-		Decision Tree	The outcome of
LI-SHANG YANG		set from data		tion		(DT)	their model was DT
CHAOCHANG CHIU		warehouse				Logistic Regres-	Performed better
						sion	than the Logistic
							Regression.
Year 2014	۶	Customer data-	≻	Sampling		Decision Tree	performance by
Stanislav Stakhovych		set of a major				with Cost Sensi-	the area under the
Michael Ewing		Australian on-				tive	ROC curve are
Ali Tamaddoni Ja-		line Fast mov-				DT	0.83, 0.91, 0.85,
hromi		ing consumer				Simple Decision	and 0.92 for Cost-
		Goods				Tree	Sensitive DT, LR,
						Boosting	Simple DT and
						Logistic Regres-	Boosting respec-
						sion	tively.
Year 2008	۶	Dataset from	۶	GA algorithm for		GA- based Novel	outcome was
Hongmei Shao		Japanese credit		feature selection		Bayesian Classifi-	compared with NB,
Gaofeng Zheng		debit and credit				er	TAN, and APRI
Fengxian An		company			4		classifier and indi-
							cate GA-based
							identified custom-
							er churn is better
							for a large number
							of customers
							shows higher clas-
							sifying Precision.
Year 2015	۶	Dataset from	-			Echo State Net-	accuracy rate for
Ali Rodan		Jordanian cellu-				work (ESN)	the KDD dataset of
Hossam Faris		lar telecommu-				SVM	ESN with SVM-
		nication com-					readout was 93.7%
		pany.					and 87.9% and for
		KDD dataset					the dataset from
							Jordan cellular
							telecommunication
							company was

Year 2008 > Dataset of VIP > Sampling > SVM SVI Zhao Jing customers do- - - - - -	spectively. M with higher curacy rate
Zhao Jing customers do- Customers do- Customers do-	-
	curacy rate
Dang Xing-hua mestic branch sion (LR) 0.5	5974, 0.5148 for
of CCB as the Decision Tree C4.	.5, 0.5890 for LR,
core data (C4.5) 0.5	5549 for Naïve
> ANN Bay	yesian Classifier,
Naïve Bayesian and	d 0.5479 for
classifier AN	IN.
Year 2006 > Dataset from > Random Sam- > Decision Tree The	e output of their
Shin-Yuan Hung wireless tele- pling > Neural Network wo	ork was Both DT
David C. Yen com company in (NN) and	d NN give accu-
Hsiu-Yu Wang Taiwan rat	te churn predic-
tio	n.
Year 2000 > Dataset Pro- - > Logistic regres- NN	l gives better
Michael C. Mozer vided by wire- sion not	nlinear structure
Richard Wolniewicz less carrier > DT in t	the sophisti-
David B. Grimes > Neural Network cat	ted representa-
Eric Johnson > Boosting tio	on than the LR,
Howard Kaushansky DT,	, and Boosting.
Year 2005 > Dataset from > Data cleaning > Decision Tree DT	based metho- In Future this
Hee Seok SongKorean online> Discretizationdol	logy can be used methodology
Hyea Kyeong Kim Shopping mall. to	detect the can be ex-
Jae Kyeong Kim cha	anges of cus- tended to dis-
Tae Seong Kim tor	mer behavior cover changes
fro	om sales and of customer
cus	stomer profile behavior for
dat	ta at different three or more
tim	nes. dataset.
Year 2017 > Dataset pro- > Data cleaning > Logistic Regres- LR	is a more ad-
Kristof Coussementvided by a large>Data reductionsionvar	nced and as-
Stefan Lessmann European mo- > Sampling ser	mbled algorithm.
Geert Verstraeten bile telecom- > Missing value It o	correctly pre-
munication handling par	res data general-
provider > outlier Iy I	less cumber-
sor	me than other

							algorithms.	
Year 2006	۶	Dataset pro-	۶	Over sampling	۶	Bagging	Boosting and Bag-	
Aurelie lemmens		vided by Tera-			۶	boosting	ging Provided bet-	
Christophe croux		data Center at					ter classifiers than	
		Duke University					a binary logistic	
							model. Predicted	
							churn performance	
							gain 26% for top-	
							decile lift and 16%	
							of Gini coefficient .	
Year 2008	۶	Dataset from	≻	Over Sampling	۶	Multilayer Per-	Synthetic Minority	
Dudyala Anil Kumar		Latin American	≻	Under Sampling		ceptron (MLP)	Oversampling as	
V. Ravi		Bank.	≻	Synthetic Minori-	۶	Decision Tree	well for under and	
				ty Oversampling		(J48)	oversampling. Syn-	
					۶	Radial Basis	thetic Minority	
						Function (RBF)	oversampling pro-	
						Network	duced excellent	
					>	Random Forest	results with	
						(RF)	91.90% overall	
					>	Support Vector	accuracy.	
						Machine		
					\succ	Logistic Regres-		
						sion (LR)		
Year 2006	۶	Dataset from	۶	Randomly Un-	۶	Support Vector	accuracy for a real	Deriving a solu-
Kristof Coussement		Belgian News-		dersampling		Machine	test set of SVM(tion to select
Dirk Van den Poel		paper publish-					SVMacc and	correct kernel
		ing company					SVMauc) are	and parameter
							84.90% and 85.14	value according
							respectively, LR is	the problem is
							84.60% and are	an interesting
							87.21%.	topic for the
								future re-
								search.
Year 2011	۶	Dataset from	-			Neural Network	The output of their	They suggested
Anuj Sharma		UCI Repository					Prediction was NN	we should pro-
Dr. Prabin Kumar		database at the					can Predict the	posed Pre-
Panigrahi		university of					customer churn	processing and

		California.				with an accuracy of	we also im-
						92.35%.	plemented
							deep methods.
Year 2008		Dataset Pro-	۶	Sampling	Improved bal-	Improved bal-	
Ellis Johnson		vided by Chines			ance random	ance(RF) gives	
Weiyun Ying		Bank			forest	highest accuracy of	
Yaya Xie						93.4% and DT, ANN	
Xiu Li						and CWC-SVM	
						gives 62%, 78.12%,	
						and 87.15% re-	
						spectively.	
Year 2014	٨	Dataset from	۶	Undersampling	Bayesian Net-	classification per-	
Thomas Verbrakena		European tele-		(removing non	work	formance was	
Wouter Verbekea		communication		churn from the	Naïve Byes Clas-	measured area	
Bart Baesensa,		operator		dataset)	sifier	under the receiver	
	۶	Three datasets	۶	Feature selection	Augmented	operating charac-	
		from Center for		(to limit the	Naive Bayes clas-	teristic curve (AUC)	
		Customer Rela-		available	sifiers	and the maximum	
		tionship Man-		attributes and	General Bayesian	profit (MP) crite-	
		agement at		Markov Blanket	network classifi-	rion and both give	
		Duke University		based algorithm	ers	a different ranking	
	۶	Synthetic data-		for feature selec-	Search-and-score	of the classification	
		set		tion is used)	algorithms	algorithm.	
					Constraint based		
					algorithms		
					Hybrid methods		
Year 2019	۶	Telecom churn	≻	Random over-	Extreme Learning	Testing accuracy of	
Koh Guan Li		dataset from		sampling	Machine	49% and Training	
P. M. Booma		kaggle	۶	Feature Scaling	Particle swarm	accuracy of 50%	
				(Particle swarm	Optimization	and also Scale fea-	
				Optimization)	(PSO) algorithm	ture Training accu-	
						racy 84% and Test	
						accuracy of	
						81.16%.	

Unsupervised

Musadig et.al (2020) in the field of banking for the retention of their customers used unsupervised Machine Learning Methodology

(K-mean clustering algorithm) and they also utilized the real-time dataset of a bank. They had utilized some Pre-processing techniques data reduction, data cleaning, data transformation and normalization. Conclude of their work is decided to avoid this model due to high computational complexity and by using the different clustering algorithm which generates similar outcomes. They suggested methodology used in this paper dilated, so that it applied to the other behaviors of bank customers, such as a behavior based on deposits, loans, investments, etc. in addition to the ones extracted from the transactions history. On the other hand, **Priyanka et.al (2017)** used Artificial Neural Network for the analysis of banking data. They used two different datasets German credit dataset from UCI and the dataset of a bank customers and implemented threshold. The outcome of their work was ANN gives accuracy for dataset 1 and dataset 2 are 72% and 98% respectively. The proposed Model works efficiently for both datasets.

Author/ Year	Dataset	Pre-processing	Methods	Result	Future Work.
Year 2020	Real time	data Reduc-	K-means	They decided to	it applied to
Musadig Aliyev	dataset of a	tion	clustering	avoid this model	the other be-
Elvin Ahmadov-	bank	data cleaning	algorithms	due to high com-	haviors of bank
Sayel Abualigah		data Trans-		putational com-	customers,
Habil Gadirli		formation		plexity and by	such as a be-
Arzu Mammadova		Normaliza-		using the different	havior based
Emin Alasgarov		tion		clustering algo-	on deposits,
				rithm which gene-	loans, invest-
				rates similar out-	ments, etc. in
				comes.	addition to the
					ones extracted
					from the
					transactions
					history.
Year 2017	> German	> Threshold	Artificial	ANN gives accura-	
Priyanka S. Patil	credit data-		Neural Net-	cy for dataset 1	
Nagaraj V. Dhar-	set from		work	and dataset 2 are	
wadkar	UCI.			72% and 98% re-	
	Dataset of a			spectively. The	
	bank cus-			proposed Model	
	tomers			works efficiently	
				for both datasets.	

Hybrid

Jina et.al (2021) offers a deep hybrid learning model for customer repurchase behavior for smart phone company. They used customer dataset set which gain by the two survey of a company in December 2017 and in December 2019. The outcome of their model for the customer purchase behavior gains the accuracy of 90.71%. On the other hand, **Saad et.al (2013)** offers a model for the churn prediction by using Regression Analysis, DT, ANN, and K-mean clustering and they used a dataset from the customer DNA website.

They applied re-sampling and data acquisition on the given dataset. The follow-up of their work is that DT is the most accurate algorithm to identified customer churn and the accuracy achieved by DT is 70%.

Jorge et.al (2004) used Neural Network, DT, Neuron Fuzzy, and (GA) rule evolver for the evaluation of wireless churn. They used dataset from major wireless carrier operating in Brazil and implemented oversampling technique. The output of their result was NN with 15 hidden units that accomplish the best classification. On the other hand, **Duyen et.al (2017)** used AdaBoost, KNN, Extra Trees, NN, and XGBoost for the customer churn in ISP. They used a dataset of ISP companies in Vietnam and implemented oversampling. The outcome of their work was XGBoost gives the better result with the precision and recall of 45.71% and 42.06% respectively. Similarly, **S.Ummugulthum et.al (2014)** proposed customer relationship management by using Naïve Bayes, SVM, KNN, and J48. They used a dataset of KDD Cup 2009 and implemented feature selection by Information gain values. Information gain measures the amount of information for split over a feature. The output of their work was accuracy rate of Naïve Bayes, SVM, KNN and J48 are 93.78%, 98.84%, 98.25, and 98.78 respectively. SVM gives the highest accuracy.

Author/ Year	Dataset	Pre-processing	Methods	Result	Future Work.
Year 2021	> Customer	-	Deep hybrid	Proposed model	
Angel P. del Pobil	dataset of a		learning	gain the accuracy	
HongGeun Ji	smart		model	of 90%.	
Soyoung Oh	phone				
Syjung Hwang	company				
Jina Kim					
Eunil Park					
Year 2013	> Dataset	Re-sampling	≻ K-mean	DT is the most	
Aatif Kamal	from Cus-	Data acquisi-	Clustering	accurate algo-	
Ammar Saleem	tomer DNA	tion	Decision Tree	rithm to identified	
Rehman	Website		Artificial	customer churn	
Ali mustafa Qamar			Neural-	and the accuracy	
Saad Ahmed Qure-			Networks	achieved by DT is	
shi			Regression	70%.	
			Analysis		
Year 2004	Dataset	Oversam-	Decision	The output of	
Jorge B. Ferreira	from Major	pling	Tree	their result was	
Marley Vellasco	wireless		> Neuroz	NN with 15 hid-	
Marco Aurélio Pa-	carrier op-		Fuzzy	den units that	
checo	erating in		≻ (GA) Rule	accomplish the	
Carlos Hall Barbosa	Brazil		Evolver	best classification.	
			Neural Net-		
			work		
Year 2017	➢ Dataset of	Over Sam-	> XGBoost	XGBoost gives the	
Duyen DO	ISP compa-	pling	Extra Trees	better result with	
Phuc HUYNH	nies in		≻ KNN	the recall and Pre-	
Phuong VO	Vietnam.		> AdaBoost	cision of 42.06%	
Tu VU			Neural Net-	and 45.71% re-	

					work	spectively.
Year 2014	۶	KDD	Cup	Feature	Naïve Bayes	Accuracy rate of
S.Ummugulthum		2009		Selection	Support	Naïve Bayes, SVM,
Natchiar					Vector Ma-	KNN and J48 are
Dr.S.Baulkani					chine	93.78%, 98.84%,
					KNN	98.25, and 98.78
					J48	respectively.

LIMITATIONS:

In previous research we have found some limitations such as lack of unsupervised learning methods being has been for customer retention. We have also struggled to find the review on deep learning methods and hybrid models based of feature selection methods.

In future we suggest that there is a need to develop hybrid model based on feature selection methods to handle the large customer's datasets. Feature selection methods can help in finding the important features without compromising on the quality of result. The combination of feature selection methods and deep learning can produce better results.

Conclusion

Customer retention has become more important topic for the business now days. To gain more customers is very important for the success of business. Customer data is imbalanced and huge that is difficult to handle manually. For the prediction of customer retention and handling of imbalanced data many paper has been reviewed in this paper. The limitations shows lack of supervised and hybrid models used for customer retention. There is a need to develop hybrid model based on feature selection methods. The combination of feature selection methods and deep learning can produce better result.

References

- [1] Kumar, M. V. N., Babu, M. G. R., & Kishore, M. A. R. An Enhanced Methodology to predict Customer Churn and Retention model using Machine Learning.
- [2] Al-Mashraie, M., Chung, S. H., & Jeon, H. W. (2020). Customer switching behavior analysis in the telecommunication industry via push-pull-mooring framework: A machine learning approach. *Computers & Industrial Engineering*, *144*, 106476.
- [3] Dalmia, H., Nikil, C. V., & Kumar, S. (2020). Churning of Bank Customers Using Supervised Learning. In *Innovations in Electronics and Communication Engineering* (pp. 681-691). Springer, Singapore.
- [4] Labhsetwar, S. R. PREDICTIVE ANALYSIS OF CUSTOMER CHURN IN TELECOM INDUSTRY USING SUPERVISED LEARNING.
- [5] Sabbeh, S. F. (2018). Machine-learning techniques for customer retention: A comparative study. *International Journal of ad*vanced computer Science and applications, 9(2).
- [6] Dharwadkar, N. V., & Patil, P. S. (2018). Customer retention and credit risk analysis using ANN, SVM and DNN. International Journal of Society Systems Science, 10(4), 316-332.
- [7] Singh, P., & Agrawal, V. (2019). A collaborative model for customer retention on user service experience. In *Advances in Computer Communication and Computational Sciences* (pp. 55-64). Springer, Singapore.
- [8] Sivasankar, E., & Vijaya, J. (2019). A study of feature selection techniques for predicting customer retention in telecommunication sector. *International Journal of Business Information Systems*, *31*(1), 1-26.

- [9] Chen, Z. Y., Fan, Z. P., & Sun, M. (2012). A hierarchical multiple kernel support vector machine for customer churn prediction using longitudinal behavioral data. *European Journal of operational research*, *223*(2), 461-472.
- [10] AL-Shatnwai, A. M., & Faris, M. Predicting Customer Retention using XGBoost and Balancing Methods.
- [11] Turkmen, A., Bahcevan, C. A., Alkhanafseh, Y., & Karabiyik, E. (2020). User behaviour analysis and churn prediction in ISP. *New Trends and Issues Proceedings on Advances in Pure and Applied Sciences*, (12), 57-67.
- [12] Ahmad, A. K., Jafar, A., & Aljoumaa, K. (2019). Customer churn prediction in telecom using machine learning in big data platform. *Journal of Big Data*, 6(1), 1-24.
- [13] Khan, A. A., Jamwal, S., & Sepehri, M. M. (2010). Applying data mining to customer churn prediction in an internet service provider. *International Journal of Computer Applications*, *9*(7), 8-14.
- [14] 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.
- [15] Ng, K., & Liu, H. (2000). Customer retention via data mining. Artificial Intelligence Review, 14(6), 569-590.
- [16] Ismail, M. R., Awang, M. K., Rahman, M. N. A., & Makhtar, M. (2015). A multi-layer perceptron approach for customer churn prediction. *International Journal of Multimedia and Ubiquitous Engineering*, 10(7), 213-222.
- [17] Chitra, K., & Subashini, B. (2011). Customer retention in banking sector using predictive data mining technique. In ICIT 2011 The 5th International Conference on Information Technology.
- [18] Larivière, B., & Van den Poel, D. (2005). Predicting customer retention and profitability by using random forests and regression forests techniques. *Expert systems with applications*, *29*(2), 472-484.
- [19] Dalvi, P. K., Khandge, S. K., Deomore, A., Bankar, A., & Kanade, V. A. (2016, March). Analysis of customer churn prediction in telecom industry using decision trees and logistic regression. In 2016 Symposium on Colossal Data Analysis and Networking (CDAN) (pp. 1-4). IEEE.
- [20] Xia, G. E., & Jin, W. D. (2008). Model of customer churn prediction on support vector machine. *Systems Engineering-Theory* & *Practice*, *28*(1), 71-77.
- [21] Zhao, Y., Li, B., Li, X., Liu, W., & Ren, S. (2005, July). Customer churn prediction using improved one-class support vector machine. In *International Conference on Advanced Data Mining and Applications* (pp. 300-306). Springer, Berlin, Heidelberg.
- [22] Hassouna, M., Tarhini, A., Elyas, T., & AbouTrab, M. S. (2016). Customer churn in mobile markets a comparison of techniques. arXiv preprint arXiv:1607.07792.
- [23] Keramati, A., Ghaneei, H., & Mirmohammadi, S. M. (2016). Developing a prediction model for customer churn from electronic banking services using data mining. *Financial Innovation*, *2*(1), 1-13.
- [24] Jinbo, S., Xiu, L., & Wenhuang, L. (2007, June). The Application ofAdaBoost in Customer Churn Prediction. In 2007 International Conference on Service Systems and Service Management (pp. 1-6). IEEE.
- [25] Xie, Y., Li, X., Ngai, E. W. T., & Ying, W. (2009). Customer churn prediction using improved balanced random forests. *Expert* Systems with Applications, 36(3), 5445-5449.
- [26] Brânduşoiu, I., Toderean, G., & Beleiu, H. (2016, June). Methods for churn prediction in the pre-paid mobile telecommunications industry. In 2016 International conference on communications (COMM) (pp. 97-100). IEEE.
- [27] Ravichandran, S., & Ramasamy, C. (2016). Customer Retention of MCDR using Three-Stage Classifier based DM Approaches. International Journal of Innovative Research in Science, Engineering and Technology (IJIRSET), 5, 11190-11196.

- [28] Yang, L. S., & Chiu, C. (2006, November). Knowledge discovery on customer churn prediction. In Proceedings of the 10th WSEAS International Conference on Applied Mathematics, Dallas, Texas, USA, November 1-3, 2006, pp523 (Vol. 528).
- [29] Jahromi, A. T., Stakhovych, S., & Ewing, M. (2014). Managing B2B customer churn, retention and profitability. *Industrial Marketing Management*, 43(7), 1258-1268.
- [30] Shao, H., Zheng, G., & An, F. (2008, October). Construction of Bayesian classifiers with GA for predicting customer retention.
 In 2008 Fourth International Conference on Natural Computation (Vol. 1, pp. 181-185). IEEE.
- [31] Rodan, A., & Faris, H. (2015, November). Echo state network with SVM-readout for customer churn prediction. In 2015 IEEE Jordan Conference on Applied Electrical Engineering and Computing Technologies (AEECT) (pp. 1-5). IEEE.
- [32] Zhao, J., & Dang, X. H. (2008, October). Bank customer churn prediction based on support vector machine: Taking a commercial bank's VIP customer churn as the example. In 2008 4th International Conference on Wireless Communications, Networking and Mobile Computing (pp. 1-4). IEEE.
- [33] Hung, S. Y., Yen, D. C., & Wang, H. Y. (2006). Applying data mining to telecom churn management. *Expert Systems with Applications*, *31*(3), 515-524.
- [34] Mozer, M. C., Wolniewicz, R., Grimes, D. B., Johnson, E., & Kaushansky, H. (2000). Predicting subscriber dissatisfaction and improving retention in the wireless telecommunications industry. *IEEE Transactions on neural networks*, *11*(3), 690-696.
- [35] Kim, J. K., Song, H. S., Kim, T. S., & Kim, H. K. (2005). Detecting the change of customer behavior based on decision tree analysis. *Expert Systems*, *22*(4), 193-205.
- [36] Coussement, K., Lessmann, S., & Verstraeten, G. (2017). A comparative analysis of data preparation algorithms for customer churn prediction: A case study in the telecommunication industry. *Decision Support Systems*, 95, 27-36.
- [37] Lemmens, A., & Croux, C. (2006). Bagging and boosting classification trees to predict churn. *Journal of Marketing Research*, 43(2), 276-286.
- [38] Anil Kumar, D., & Ravi, V. (2008). Predicting credit card customer churn in banks using data mining. *International Journal of Data Analysis Techniques and Strategies*, 1(1), 4-28.
- [39] Coussement, K., & Van den Poel, D. (2008). Churn prediction in subscription services: An application of support vector machines while comparing two parameter-selection techniques. *Expert systems with applications*, *34*(1), 313-327.
- [40] Coussement, K., & Van den Poel, D. (2008). Churn prediction in subscription services: An application of support vector machines while comparing two parameter-selection techniques. *Expert systems with applications*, *34*(1), 313-327.
- [41] Ying, W., Li, X., Xie, Y., & Johnson, E. (2008, July). Preventing customer churn by using random forests modeling. In 2008 IEEE International Conference on Information Reuse and Integration (pp. 429-434). IEEE.
- [42] Verbraken, T., Verbeke, W., & Baesens, B. (2014). Profit optimizing customer churn prediction with Bayesian network classifiers. *Intelligent Data Analysis*, 18(1), 3-24.
- [43] Li, K. G., & Marikannan, B. P. (2019). Hybrid Particle Swarm Optimization-Extreme Learning Machine Algorithm for Customer Churn Prediction. *Journal of Computational and Theoretical Nanoscience*, *16*(8), 3432-3436.
- [44] Aliyev, M., Ahmadov, E., Gadirli, H., Mammadova, A., & Alasgarov, E. (2020). Segmenting Bank Customers via RFM Model and Unsupervised Machine Learning. *arXiv preprint arXiv:2008.08662*.
- [45] Patil, P. S., & Dharwadkar, N. V. (2017, February). Analysis of banking data using machine learning. In 2017 International Conference on I-SMAC (IoT in Social, Mobile, Analytics and Cloud)(I-SMAC) (pp. 876-881). IEEE.

- [46] Kim, J., Ji, H., Oh, S., Hwang, S., Park, E., & del Pobil, A. P. (2021). A deep hybrid learning model for customer repurchase behavior. *Journal of Retailing and Consumer Services*, *59*, 102381.
- [47] Ferreira, J., Vellasco, M. M., Pacheco, M. A. C., Carlos, R., & Barbosa, H. (2004, April). Data Mining Techniques on the Evaluation of Wireless Churn. In *ESANN* (Vol. 28, pp. 483-488).
- [48] Do, D., Huynh, P., Vo, P., & Vu, T. (2017, December). Customer churn prediction in an internet service provider. In *2017 IEEE International Conference on Big Data (Big Data)* (pp. 3928-3933). IEEE.
- [49] Natchiar, S. U., & Baulkani, S. (2014, November). Customer relationship management classification using data mining techniques. In 2014 International Conference on Science Engineering and Management Research (ICSEMR) (pp. 1-5). IEEE.

CGSJ