Comparative Analysis of Customer Churn Prediction

Dr. Anithaa Kingly K
CSE department
Bharath Institute of Higher Education and Research
Tambaram, Chennai
anithaa.cse@bharathuniv.ac.in

Rama Narayana B
CSE department
Bharath Institute of Higher Education and Research
Tambaram, Chennai
ramanarayana246@gmail.com

Sasidhar Reddy B
CSE department
Bharath Institute of Higher Education and Research
Tambaram, Chennai
bvsasidhar@gmail.com

Raju R
CSE department
Bharath Institute of Higher Education and Research
Tambaram, Chennai
raju.rachakonda2002@gmail.com

Jaya krishna A
CSE department
Bharath Institute of Higher Education and Research
Tambaram, Chennai
jayakrishnaanne33@gmail.com

***Abstract*—The wide variety of service providers are being elevated very swiftly in each business enterprise. For service providers, a rapidly expanding market in each location is leading to a larger subscriber base. Customer acquisition costs are rising as a result of increased competition, new and innovative business models, and better-suited goods. Serviceproviders have learned the value of keeping existing clientshappy in such a short period of time. Because of this, it is critical that service providers reduce churn, the occurrencewhen consumers of a business stop buying from or engaging with the business. Not just in banking and telecommunications, but also in other industries that are heavily reliant on customer engagement, this study examines the most well-known machine learning methods for churn prediction.**

Keywords—Customer churn, Logistic Regression, Xgboost Classifier, K-Nearest Neighbour, Decision Tree Classifier, Random Forest Classifier, GradientBoost Classifier, Ridge Classifier, BaggingRidge Classifier, ExtraTree Classifier.

I. INTRODUCTION

If you're a service provider, "churning" is often defined as the percentage of consumers that end their contracts due to competition. Churners are people who have stopped doing business with a firm because they were unhappy with the service they received. An examination of the likelihood of a customer discontinuing use of a service or product is what is meant by a customer churn analysis. Preventive measures are necessary to avoid a customer's items or services being left behind in the case of a circumstance like this.

The marketplace is very dynamic and distinctly aggressive in recent times. It is because of the supply of a big wide variety of service providers. Customers are a company's most valuable asset since they represent its primary source of revenue. Businesses are now cognizant of the fact that they must pay close attention not just to get new clients but also to keep their existing ones satisfied. A churner is a person who moves around a lot and has a variety of reasons for doing so. Customer churn is minimised when the organisation can accurately forecast the customer's mindset and establish linkages between client attrition and things thatare under their control. Predicting churn rates is a binary classification job that separates churners from non-churners. For any enterprise, vanquishing commercial enterprise from new customers means going through the sales pipeline, using their sales and advertising belongings in the cycle. Customer retention, then again, is generally extra budget- effective, due to the fact they have already won the confidence and loyalty of current customers. So, predicting customer churn rate at the earlier stages is really important for an organization.

In data science, machine learning is a method for creating analytical models automatically. As a result of machine learning, computers are able to discover patterns that would otherwise go unnoticed. Unsupervised, semi-supervised, and supervised machine learning approaches all exist. The goal of supervised learning is to discover patterns in labelleddatasets by using machine learning techniques. Hidden patterns may be discovered in unlabeled data using unsupervised machine learning. For training, semi- supervised learning uses a mix of labelled and unlabeled data, often a small number of labelled records and an enormous number of unlabeled facts. Unsupervised learning and supervised learning are at opposite ends of the spectrum,whereas semi-supervised learning is in the middle.

In order to avoid the previously described inconvenience, businesses must be able to accurately forecast the purchasing behaviour of their customers. There are twoways to control customer churn: First and foremost, (1) Reactive (2) Taking charge. As soon as the consumer asks for a cancellation, a proactive agency presents the customer with appealing ideas in order to maintain their business. The proactive approach anticipates that some customers may depart, therefore it provides a strategy for them to follow. The churners and non-churners are classified in a binary job.In order to deal with this problem, we used the following machine learning techniques: [1]. Using Logistic Elements for Regression Decision Trees, Support Vector Machine,Random Forest Classifier, Extra Tree Classifier and Boosting Algorithm in addition to XGBoost are all examplesof [2]. When it comes to machine learning, it's important that data be linear.

II. LITERATURE SURVEY

Churn prediction in the telecom business and related studies suggested by notable scholars are summarized below. [1,5,7,8,11-14,16]

The real-world Telecom dataset was used by Guo-en Xia et al. [6] to demonstrate their Churn Prediction Model. SVM and Artificial Neural Network were compared to Decision Tree, Nave Bayes, SVM, and weighted selected ensemble churn prediction algorithms in this study (ANN). Base Classifiers perform less well than the suggested technique, according to the findings of experiments.

Decision trees, random forests, GBM tree methods, and XGBoost were used by Abdulrahim et al. [3] to forecast customer attrition. XGBoost outperformed the competition in terms of AUC accuracy in our tests. However, the feature selection process may be enhanced by utilizing optimization methods.

These researchers reviewed all of the machine learning models that were taken into account, as well as gave a thorough study of current feature selection methods. They discovered that decision trees outperformed the rest of the models in the prediction models. Predictive accuracy is enhanced by using optimization approaches in feature selection. Thereafter, author's proposed future study paths following a comparison of current methodologies.

Chuanqi Wang et al. [4] described Churn Prediction Model as a classification issue that is cost sensitive. The term "cost sensitive" was created to describe how the model classified consumers into churners and non-churners and how it determined the maximum profit the organization could receive. According to the authors' research, they were able to demonstrate classification accuracy as well as a high rate of misclassification.

This model was developed by Adnan Idris and Asifullah Khan [5] using an Orange Dataset and a Cell2Cell Dataset. To begin, we used a feature selection strategy that prioritized relevancy above duplication. The Ensemble-basedClassifiers were combined with the basic Classifiers in orderto acquire majority vote casting, so that it could forecast future events more precisely. Three methods were used: random forest, rotation forest, and ok-NN. AUC (Area Under the Curve), sensitivity, specificity, and Q-facts were employed as performance indicators in this work.

By utilizing a Boosting algorithm, the Churn Prediction Model presented by Ning Lu and colleagues [7] was created. The Boosting set of criteria was used to separate consumers into clusters according on the weight they were allocated. To estimate the likelihood of churn, the Logistic Regression was used to each Base Classifier. Results showed that Boosting is a reliable predictor of Churn Prediction.

Loyalty to a pre-paid mobile phone service provider has been shown to reduce customer turnover, according to A.T. Jahromi [11]. Segmenting functions and using many techniques, such as selection bushes and neural networks, were used in this study. A mixed approach was shown to be more effective than a single algorithm. The KNN-LR is a combination of logistic regression and the KNN algorithm. In computing, the KNN algorithm [12] is used. A study of KNN-LR and the radial foundation feature (RBF) network indicated that KNN-LR verified the good performance of logistic regression and the radial foundation feature network. An allotted approach for predicting customer attrition has been developed by Y. Zhang [13]. CRM service quality is enhanced because to the framework.

'Customer churn' and 'churn management' are two termsused in the cellular mobile phone service industry to describe the movement of customers from one firm to the next. An example of churn is when a customer leaves a company in search of another one, which is also known as attrition. A similar word, "churn control," is used in the cell community services business to describe the procedures employed to ensure that a company has the greatest number of essential consumers [11]. Effective customer management, in essence, requires the capacity to predict when a consumer will switch service providers. It also assumes a high level of customer profitability, as well as a unique strategy and method to keep customers from leaving the company.

III EXISTING SYSTEM

In this research, many machine learning algorithms are available for prediction of customer churn. Some of the machine learning algorithm like Logistic Regression, K Nearest Neighbor (KNN), Decision Tree classifier, Random Forest classifier, Extra Tree classifier, Ridge classifier, Bagging Ridge classifier, Gradient Boosting algorithms etc. Researchers used service failure rate, length of customer association, and customer complaints to evaluate the level of dissatisfaction across the operator’s database. Accordingly, considering the limitations in the available data in the banks or telecom database, in this research, length of customer association and customer complaints were used to evaluate level of customer’s dissatisfaction. By using these dataset authors extract and select various features and completed their research.

DRAWBACK: The problem of over-fitting of the non-churners and under-fitting of our churners affects our model in a large way.

III PROPOSED SYSTEM

Logistic Regression

For the prediction of continuous values, Linear Regression [1] is a supervised Machine Learning technique. linear regression [1] assumes a linear connection between the dependent and independent variables. In layman's words, it looks for the line or plane that best reflects two or more variables.

Building a linear equation describing the connection between the dependent and independent variables is what Linear Regression is about. We have a dataset with the variables x and Y as independent and dependent variables. As a consequence, we will use Linear Regression to generate the following equation [1].

$$\mathbf{y = mx + c}$$

'm' denotes the slope and 'c' denotes the intercept in this equation, which is a simple straight line. For binary classification tasks, logistic regression is a strong supervised ML technique [1]. When the objective is a certain category, as a linear regression technique, logistic regression can be used to solve classification problems, which makes it a useful tool. Using logistic regression, we may model a binary output variable using the logistic function described below. Because the range of logistic regression is confined to 0 and 1, it stands apart from other types of regression. It isnot necessary for input and output variables to be linearly related in logistic regression. Nonlinear log transformations are to blame for this.

$$\boldsymbol{Logistic\ function} = \frac{\mathbf{1}}{\mathbf{1} + \boldsymbol{e}^{-\boldsymbol{x}}}$$

XGBoost Classifier

Due to its emphasis on speed and efficiency, the decision-tree-based ensemble learner, Extreme Gradient Boosting (XGBoost) was developed [2]. XGBoost enhances prediction accuracy by iteratively constructing decision trees, each one taking into account the mistakes and flaws of the preceding one and learning from them [2]. The tree structure learns from the preceding tree's results and residuals in each iteration or sequence. Differences between actual and predicted values are known as residuals [2]. These weak learners are combined into one strong classifier using XGBoost in order to produce more accurate predictions about what is going to happen in the future.Cache optimization using XGBoost has both advantages and disadvantages. While XGBoost can predict outputs with excellent accuracy, it requires a lot of training time [2]. For the first time, the XGBoost classifier is able to handle datasets with missing values, unlike most other tree learning methods [2]. Because of this, it is regarded an excellentmodel in terms of its high prediction accuracy.

$$\boldsymbol{obj}(\mathbf{0}) = \sum_{i}^{n} \boldsymbol{l}\,(\boldsymbol{y}_i - \boldsymbol{y}_i^{\wedge}) \sum^{j} \boldsymbol{\Omega}\,(\boldsymbol{f}_j)$$

It's possible to estimate the probability that a data point will belong to one group or another depending on which group its closest neighbours belong to using the k-nearest neighbours (KNN) algorithm [3]. Supervised Machine Learning algorithms such as the k-nearest neighbour method may be utilized to handle classification and regression issues. As a result, it is mostly utilised to solve categorization difficulties. Lazy learning and non-parametricalgorithms are the hallmarks of KNN [3]. As soon as you provide the training data, it is referred to as a "lazy learning algorithm" or "lazy learner." In place of doing any computations, it just saves the data throughout the training period. Until a query is run, it doesn't develop a model for the dataset When it comes to data mining, this makes KNN
[3] excellent.

$$\boldsymbol{d} = \sqrt{((\boldsymbol{x1} - \boldsymbol{x2})^2 + ((\boldsymbol{y2} - \boldsymbol{y1})^2)}$$

Decision Tree Classifier

There is a non-parametric supervised learning approach called Decision Trees (DTs) [4] that is utilised in Classification Regression. Data characteristics may be used to infer basic decision rules that can be used to forecast the value of a target variable. As a piecewise constant approximation, a tree may be considered. It is widely accepted that decision tree classifiers are one of the most often used approaches for representing data classification. This issue has been addressed by academics in a variety of domains, including machine learning, pattern recognition,and statistical analysis. It has been suggested that Decision Tree classifiers [4] may be used in a variety of ways in a variety of industries. The decision trees discussed in this work are explained in great depth. Also included is an evaluation of the paper's details, such as the algorithms/approaches employed as well as the datasets and results produced. To further demonstrate the authors' points and determine the most accurate classifiers, we described every option we looked at. It is as a consequence of this that various datasets and their results are debated and examined.

Random Forest classifier

A supervised learning method, the Random Forest classifier [5] is an ensemble learner. Using a bagging strategy, the Random Forest classifier employs many classifiers, with the basic learners being decision trees. As a result, some data is utilised more than once in the training process for a decision tree. There are several decision trees where not all characteristics are put to good use. Nodes are divided in the Random Forest classifier [5] using a random selection of characteristics rather than the optimal split using all features, as in other classifiers. Minimizes the correlation between trees and reduces generalization error. Because various subsets of the training data are utilised for each tree in Random Forest, the classifier becomes more stable and resilient in the face of noise and overtraining as a result of the increased tree diversity. The Random Forest relies on themajority vote of each of its decision trees to produce its ultimate output. The class that garners the most votes isselected as the end product.

PE * = P(x, y) (mg(X, Y)) < 0

K-Nearest Neighbour

$i=1$

**Gr
adi
ent
Bo
ost
ing
cla
ssif
ier**

Powerful machine-learning algorithms, such as gradient boosting machines, have been widely used in a broad variety of practical applications. Like being trained with regard to various loss functions, they are very adjustable to the application's demands. With an emphasis on machine learning, this article provides a pedagogical introduction to the concept of gradient-boosting approaches. All steps of the gradient boosting model design are described and illustrated using examples and diagrams. Complexity management considerations are highlighted. Gradient boosting applications are given and thoroughly analyzed in this article, which includes three examples.

$$L = -\sum_{i=1}^{n} y_i \log(p) + (1-p)\log(1-p)$$

Ridge classifier

Regression Classification for Kernels is an extension of ridge regression to the kernel version that we provide (KRRC). Many researchers employ Kernel analysis to discover the nonlinear structure of their data manifold. KRRC implicitly transforms the observed data into a potentially much higher dimensional feature space using the kernel trick and ridge regression classification in feature space Class-specific subspace distances may be used to establish the new test sample's classification if the new test sample is described as a linear combination of gallery components that are unique to each of these classes. Research at the University of California has proven that the proposed method is accurate and efficient when applied to multiple benchmark datasets.

Y = XB + e

Bagging Ridge classifier

Classifiers based on bagging ridge regression [8] are shown here. In addition to being quicker to calculate, bagging ridge regression may be used for ensemble techniques on small and medium-sized datasets since it has a closed-form solution. To train bagging ridge regression classifiers, we employ random vector functional link networks. Each base classifier uses a separate sample of training data to build multiple kernel ridge regression classifiers. In contrast to utilizing all N training samples for bagging matrix inversion,the partitioning of the training data into separate subsets results in a decrease in computing cost. A well-known multi-class UCI data set is used to assess the suggested approach. The suggested ensemble technique outperforms the single bagging ridge regression classifier and its bagging variant in an experiment.

$$\min_{Q} ||Y - XQ||^2$$

the predicted class label of x:

$$Y^T (K + \lambda I)^{-1} k$$

where $k = (k_1,....,k_N)^T$, $k_n = x_n \cdot x$ and $n = 1,...,N$

ExtraTreesClassifier [9] is an ensemble learning approach that relies on decision trees as its foundation. Unlike Random Forest, ExtraTreesClassifier [9] randomizes specific choices and subsets of data to reduce overfitting and overlearning from data. Extra Trees and a Random Forest have two major differences: As a result, no optimum splits are used and data is not bootstrapped (i.e., samples without replacement). Briefly summarized: Extra Trees bootstrap = false implies that data is sampled without replacing nodes bydefault, meaning that nodes are divided based on random splits among a random subset of the feature sets selected for each of the nodes. Instead of using bootstrapping to generate randomness, Extra Trees uses random splits to generate randomness [9].

$$s = \sum_{i=1}^{c} -p_i \log_2(p_i)$$

Extra trees classifier

IV EXPERIMENT RESULT

In this section, the proposed work has been explained by considering data collection and preparation of dataset.

Data collection and preparation

Categorical and continuous variables are converted into a form that is appropriate for further analysis during the data preparation process Several data preparation strategies are investigated in order to increase the logit model's prediction potential. Using typical input data, including cross-sectional data from a large European telecom's provider, an enhanced logit model is put to the test in this churn prediction modelling scenario. The following conclusions may be reached as a result of the research. Analysis of churn prediction performance may increase by as much as 14.5 percent in the area under the receiving operational characteristics curve and by as much as 34% in the top decile lift, depending on the data preparation approach utilised. Additionally, the improved logistic regression can hold its own against more sophisticated single and ensemble data mining techniques, as well. To wrap things up, there are a few management implications and ideas for further study, including evidence that the findings may be applied to other corporate situation s.

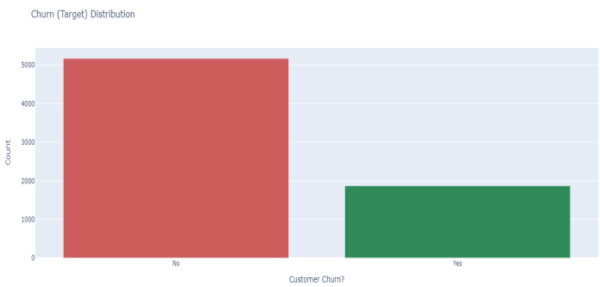

Fig: Churn (Target) Distribution dtype: float64 CUSTOMER % CHURN.

We have 26.5% of our data that is about the Churned customers, and I will try to understand the pattern of these groups I will filter the dataset and set a dataset for Churn and Non-Churn Customers. Also, I will see if monthly

Charges has some difference to Churn and Non-Churn Customers. I have the hypothesis that maybe Churn customers has a highest mean value of no churn customers.

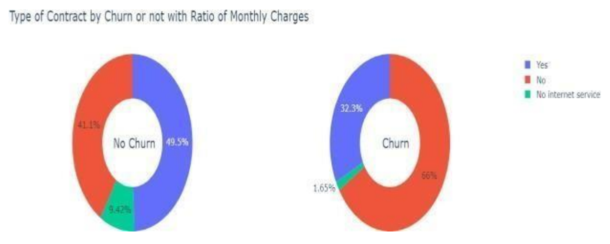

Fig: Type of contact by Churn or not with Ratio of Monthly Charges.

If you don't know how many customers are churning out, you won't know how well your firm is doing, and you won't know how much money you're losing. The more clients yourfirm loses, the greater the churn rate.

$$\textit{Churn Rate} = \frac{\textit{Number of churned customers}}{\textit{Total number of customers}}$$

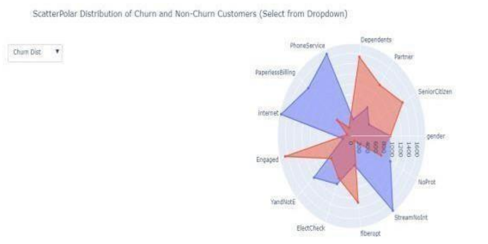

Fig: Scatter Polar Distribution of churn and non-Churn customers.

Customers' churn rate is an important indication of their satisfaction. If your turnover rate is low, it means your consumers are satisfied; if it is high, then means your customers are dissatisfied. Over time, even a low rate of monthly/quarterly churn might add up. 1 percent turnover per month equates to about 12 percent churn per year.

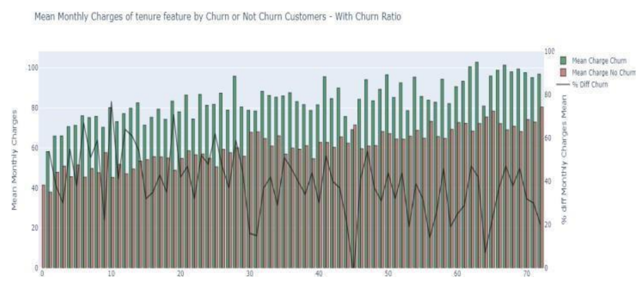

Fig: Mean Monthly Charges of tenure feature by Churn or non-churn Customers-with Churn Ratio.

The percentage of accounts that cancel or chose not to renew their subscriptions is known as churn. A high churn rate can reduce Monthly Recurring Revenue (MRR) and indicate buyer dissatisfaction with product or service.

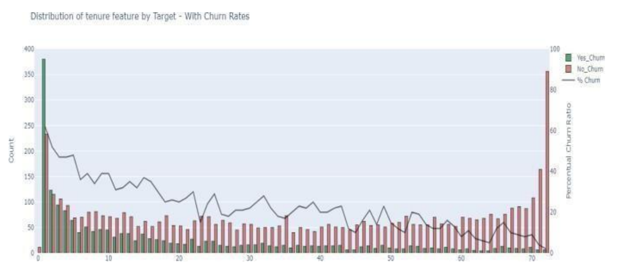

Fig: Distribution of tenure feature by Target - with Churn Rates.

Employee churn, also known as attrition, is the rate at which people leave a company. It's the number of individuals who have left the organization throughout time divided by the average number of workers, no matter how you slice it. Percentage is the most common way to represent it (percent).

Fig: Dispersion of Total Charges explained by Monthly Charges by Target.

Investors can access consensus analyst target prices for free on the internet. The level of dispersion in the individualtarget prices that make it up the consensus influences the predictive link between the consensus target price and future returns in this paper. We find some evidence that returns anticipated by consensus target prices are more favorably correlated with realized future returns when dispersion is moderate.

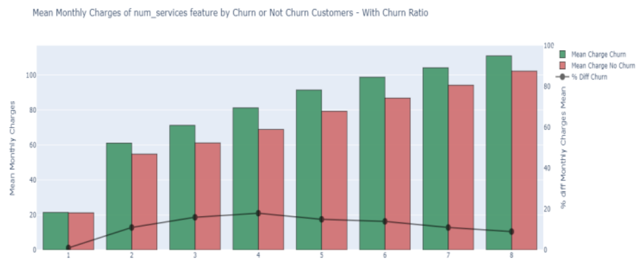

Fig: Mean Monthly Charges of num_services feature by Churn or Not Churn Customers – with Churn Ratio.

Because monthly churn increases with time, the dramatic disparity is evident. While a yearly churn rate of 5% is calculated across the whole year, The following formulas may be used to convert yearly churn to monthly churn and the other way around.

$$\textit{MonthlyChurnRate} = 1 - (1 - \textit{AnnualChurnRate})^{1/12}$$

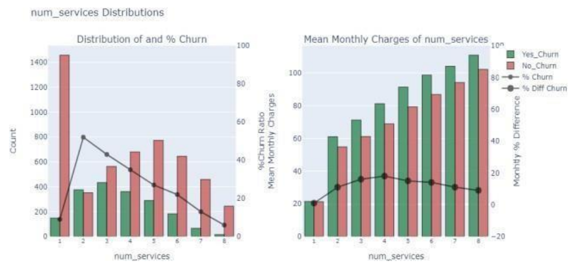

Fig: num_services Distributions.

An increase in customer turnover may have a negative impact on profits and growth. The churn rate is essential in the digital economy. Since many of these businesses compete in a wide range of industries, customers have many options when it comes to changing service providers.

Customer Churn Rate = (Lost Cust ÷ Tot_Cus_at_the_

Start_of_Time_Period) x 100

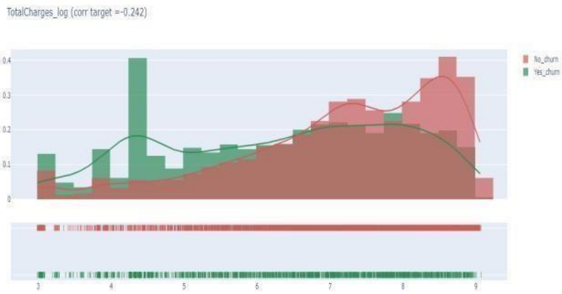

Fig: TotalCharges_log (corr target = -0.242)

We can see that churn subscribers have lower Total Charges levels. I believe that is a signal of changing tenure values; let's see what the tenure feature says.

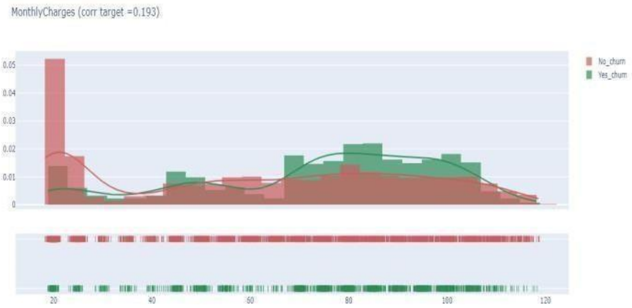

Fig: MonthlyCharges (corr target=0.193).

The churn rate, or the rate of attrition or customer churn, is the percentage of customers that leave an organization. Percentage of customers who have cancelled their subscription over a particular length of time is the most common way to portray this figure It's also a gauge of how often employees quit their employment after a specific amount of time. Increasing a company's client base requires growth (as measured by the number of new customers) to beat attrition.

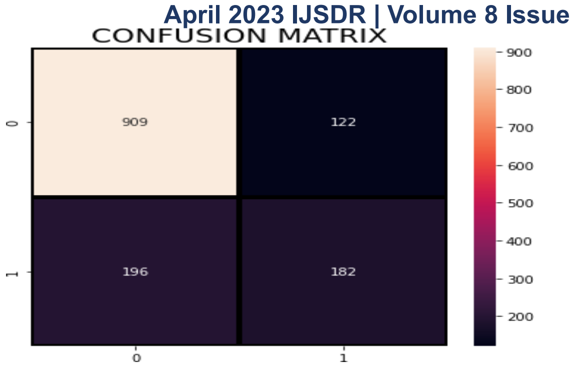

Fig: Confusion Matrix Predictions.

Confusion matrices are widely used to describe the performance of a classification model (or "classifier") on a set of test data for which the true values are known. The vocabulary used to describe the confusion matrix might be difficult to grasp.

$$\boldsymbol{Accuracy} = \frac{\boldsymbol{T(N+P)}}{\boldsymbol{TN+F(P+N)+TP}}$$

In order to better understand the performance of your classification model, classification accuracy is a major issue.

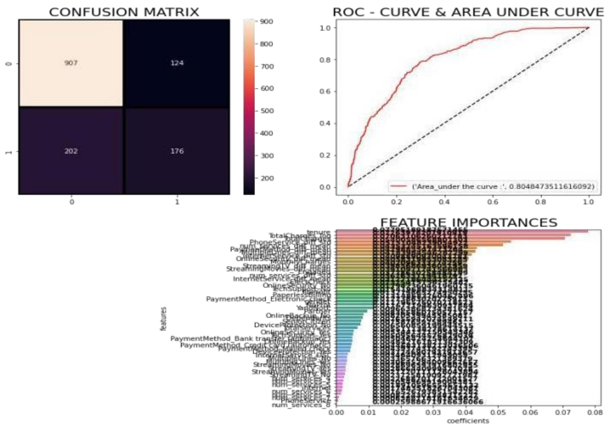

Fig: X has feature names, but RandomForestClassifier was fitted without feature names.

The average impurity decrease estimated from all decision trees in the forest can be used to compute the feature importance using the Random forest methodology. This holds true despite of whether the data is linear or non-linear (linearly inseparable).

$$\boldsymbol{RFfi_i} = \frac{\sum_{j \in all\ trees} \boldsymbol{normfi_{ij}}}{\boldsymbol{T}}$$

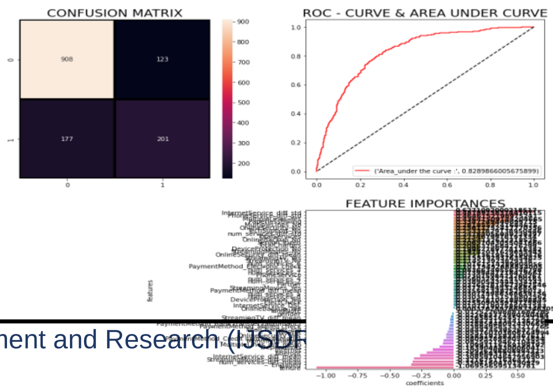

Fig: X has feature names, but Logistic Regression was fitted without feature names.

The sigmoid function is used extensively in logistic regression. The sigmoid graph seems to be in the shape of a S. It might perplex you and lead you to believe it is a non-linear expression. That, however, is not the case. A linear model is what logistic regression is. As a result, most verifying information to it as a generalized linear model (GLM).

$$Z = W_0 + W_1X_1 + W_2X_2 + W_3X_3 + W_4X_4$$

$$y = {}^{1}/_{1 + e^{-z}}$$

V CONCLUSION

As technology advances, so do the number of services available, making it more difficult for businesses to anticipate which consumers would choose to discontinue their use of those services. The prediction of customer turnover in the telecom industry has been a major study topic in recent years. In order to reduce customer turnover, many machine-learning algorithms have been used. Analysis of Customer Churn Prediction and current base classifiers was the focus of this work. All models were tested on a publicly available dataset of telecom customer turnover. Logistic Regression, Xgboost classifier, and Random Forest Classifier are the best three models based on the experimental findings. Their accuracy, precision, recall, and AUC scores were all higher than those of the competition. Furthermore, it is expected that deep learning methods would improve with time, resulting in better success rates.

REFERENCES

[1]. D. Buo and M. Kjellander, ‘Predicting Churn Rates at a Swedish CRM-System Company, Linkopings Universitet,’ Linkoping, 2014.

[2]. H. Gebert, M. Geib, L. Kolbe and W. Brenner, ‘Knowledge-enabled customer relationship management: integrating customer relationship management and knowledge management concepts,’ Journal of Knowledge Management, vol. 7, no. 5, pp. 107-123, 2003.

[3]. F. Khodakarami and Y. Chan, ‘Exploring the role of customer relationship management (CRM) systems in customer knowledge creation,’ Information & Management, vol. 51, pp. 27-42, 2014.

[4]. N. Gordini and V. Veglio, ‘Customers churn prediction and marketing retention strategies. An application of support vector machines based on the AUC parameter-selection technique in B2B e-commerce industry,’ Industrial Marketing Management, vol. 62, pp. 100-107, 2017.

[5]. I. Ullah, B. Raza, A. K. Malik, M. Imran, S. U. Islam and S. W. Kim, ‘A churn prediction model using random forest: analysis of machine learning techniques for churn prediction and factor identification in telecom sector,’ IEEE Access, pp. 60134-60149, 2019.

[6]. A. T. Jahromi, S. Stakhovych and M. Ewing, ‘Managing B2B customer churn, retention and

profitability,' Industrial Marketing Management, vol. 43, no. 7, pp. 1258-1268, 2014.

[7]. I. Ullah, B. Raza, A. K. Malik, M. Imran, S. U. Islam and S. W. Kim, 'A churn prediction model using random forest: analysis of machine learning techniques for churn prediction and factor identification in telecom sector,' IEEE Access, pp. 60134-60149, 2019.

[8]. K. Casement and D. Van den Poel, 'Churn prediction in subscription services: An application of support vector machines while comparing two parameter-selection techniques,' Expert Systems with Applications, vol. 34, no. 1, pp. 313-327, 2008.

[9]. A. K. Ahmad, A. Jafar and K. Aljoumaa, 'Customer churn prediction in telecom using machine learning in big data platform, Journal of Big Data volume 6,' vol. 6, no. 28, 2019.

[10]. T. Jiang, J. L. Gradus and A. J. Rosellini, 'Supervised Machine Learning: A Brief Primer,'Behaviour Therapy, vol. 51, no. 5, pp. 675-687, 2020.

[11]. V. F. Rodriguez-Galiano, B. Ghimire, J. Rogan, M. Chica-Olmo and J. Rigol-Sanchez, 'Anassessment of the effectiveness of a random forest classifier for land-cover classification, 'ISPRS Journal of Photogrammetry and Remote Sensing, vol. 67, pp. 93-104, 2012.

[12]. V. F. Rodriguez-Galiano, B. Ghimire, J. Rogan, M. Chica-Olmo and J. Rigol-Sanchez, 'Anassessment of the effectiveness of a random forest classifier for land-cover classification,' ISPRS Journal of Photogrammetry and Remote Sensing, vol. 67, pp. 93-104, 2012.

[13]. X. Deng, Q. Liu, S. Mahadevan and Y. Deng, 'An improved method to construct basicprobability assignment based on the confusion matrix for classification problem,' Information Sciences, Vols. 340-341, pp. 250-261, May 2016.

[14]. Nabgha Hashmi, Naveed Anwer Butt, Dr. Muddsar Iqbal, 'Customer Churn Prediction in Telecommunication A Decade Review and Classification,' Telecommunication Churn Prediction, September 2013.