Customer – Churn Prediction Using Machine Learning

LR+NB

Dr. Varsha Agarwal
Department of Management
Atlas Skilltech University
Mumbai, Maharashtra
always.varsha5685@gmail.com

Shwetkranti Taware
Department of Computer Engineering
Indira College of Engineering and Management
Pune, India
shweta.taware@indiraicem.ac.in

Suman Avdhesh Yadav
Department of CSE
Amity University
Greater Noida, Uttar Pradesh
sayadav@gn.amity.edu

Dr. Durgaprasad Gangodkar
Department of CSE
Graphic Era Deemed to be University and Graphic Era Hill University
Dehradun, Uttarakhand
dr.gangodkar@geu.ac.in

Dr. ALN Rao
Department of CSE
Lloyd Institute of Engineering and Technology
Greater Noida, India
dean.engineering@liet.in

V K Srivastav
Division of Research & Innovation
Uttaranchal University
Uttarakhand, India
dean.uttaranchaluniversity.ac.in

2022 2nd International Conference on Technological Advancements in Computational Sciences (ICTACS) | 978-1-6654-7657-7/22/$31.00 ©2022 IEEE | DOI: 10.1109/ICTACS56270.2022.9988187

***Abstract*— The gradual but consistent decrease in the number of customers retained over time is referred to as "customer churn," and it is a word that is frequently used in the business and financial sectors. If a company can identify the customers who are most likely to leave, they are more likely to take preventative efforts to keep those customers as clients. It is to the bank's advantage to have knowledge about which customers are theoretically and practically most likely to switch banks in the relatively close future. This article explains how to use machine learning algorithms to identify banking customers who may be considering switching financial institutions. This article demonstrates how machine learning models such as Logistic Regression (LR) and Naive Bayes' (NB) can effectively forecast which customers are most likely to leave the bank in the future by using data such as age, location, gender, credit card information, balance, etc. The article also uses data such as age, location, gender, credit card information, balance, etc. In addition, this article demonstrates the probabilistic predictions that may be generated using machine learning models such as Logistic Regression (LR) and Naive Bayes (NB). The findings of this research ultimately point to the conclusion that NB is superior to LR.**

Keywords*— Banking, Customer Churn Prediction, Machine Learning, LR, NB

I. Introduction

One of the most important facets of expansion for subscription-based products is keeping existing customers happy. With so many vendors offering similar services, competition is fierce in the software as a service (SaaS) industry. One negative interaction is all it takes to lose a customer forever. And if many dissatisfied clients suddenly abandoned the firm, it would suffer heavy financial and public relations losses.

This article discusses the use of predictive modelling by SaaS businesses to deal with customer churn, based on interviews with professionals from HubSpot and ScienceSoft. As you read on, you will discover effective strategies for addressing this issue. Here, we'll discuss the best ways to gather information about customers' interactions with a brand, the types of customer behaviour most strongly associated with customer turnover, and the criteria for selecting the most effective machine learning models.

Customer Churn an Overview:

Churn, often referred to as "customer churn," occurs when a company's formerly loyal customers cease buying its products. The churn rate, also known as the customer attrition rate, is the rate at which a company loses clients over the course of a given period of time. One technique to calculate the churn rate is to divide the total number of customers lost over a certain time period by the total number of new customers acquired during that same time period, then multiply the result by 100%[19]. This provides the churn rate for customers. If you gained 150 new customers in the preceding month but also lost 3, your customer churn rate is 2%[20].[33]

The churn rate is a crucial number to assess the health of a company, especially one whose consumers are paying subscribers, according to Alex Bekker, who oversees data analytics at ScienceSoft. To demonstrate this concept using a passage from the article as an example: "Customers choose a product or service to get again for a predetermined period of time, which could be as short as one month, in subscription-based enterprises. Because of this, the buyer is compelled to always search for more alluring discounts. Customers always have the choice to end their association with the business if they change their minds before the end of their existing contract with the company. Natural attrition is inevitable, but the precise quantity will differ from firm to company. But if your company's turnover rate is greater than that, something is undoubtedly wrong.

One of the numerous ways that companies can fail is through poor onboarding, when consumers aren't given clear instructions on how to use the product and what it can achieve, and through poor communication, where customers don't receive feedback or have their questions answered right away. When clients aren't provided clear instructions on how to utilise a product and what it can achieve, poor onboarding results. Long-term consumers who do not receive as many incentives as new customers could feel that they are not valued as highly.

978-1-6654-7657-7/22/$31.00 ©2022 IEEE

The most important component in establishing the value that a customer places on the goods and services provided by a firm is the general perception that a client has of the brand.

Any one of your clients will stop being loyal to you if they have a bad experience with your business. PricewaterhouseCoopers (PwC) found that 17% of respondents in the United States would stop using a brand after just one unpleasant experience, compared to 59% of respondents who would stop using it after several negative interactions.

Even loyal customers can lose trust after a negative experience. Source:PwC Survey 2017/18.

Impact of customer churn on businesses

Researchers believe that churn is bad. But exactly how does it affect the long-term success of businesses?

Michael Redbord, general manager of HubSpot's Service Hub, cautions not to undervalue the effects of even a tiny bit of churn. "In a subscription-based business, even a tiny percentage of monthly or quarterly churn will add up over time. A yearly churn rate of approximately 12% is achieved with a monthly churn rate of just 1%. Due to the fact that it is far more expensive to acquire new customers than to keep existing ones, businesses with high customer churn rates can quickly find themselves in financial trouble as they are forced to devote more and more resources to this task.

Numerous questionnaires with a focus on client acquisition and retention costs are available online. According to this report by conversion rate optimization company Invesp, it might cost up to five times as much to get a new customer as it does to keep a current one.

There is a link between increased churn rates and lower revenue. They also have a more subdued impact on a company's capacity for growth, according to Michael. Consumers of today don't hesitate to discuss their contacts with suppliers on peer-to-peer networks, social media, and review websites. 49% of customers reported sharing a favourable interaction with a company on social media, according to HubSpot Research. In a world when consumer trust in businesses is dwindling, word-of-mouth has more influence on the purchasing choice than ever before. The same HubSpot Research survey claims that 55% of consumers no longer have the same level of faith in merchants, 65% don't believe retailer press releases, 69% don't believe commercials, and 71% don't believe sponsored social media postings[34].

Scientists generally agree that turnover has a detrimental effect. What effect does it have, though, on a company's performance over the long haul?

General Manager of HubSpot's Service Hub Michael Redbord issued a warning that any churn, no matter how small it may seem, can have a substantial effect on a business. In a company that uses a subscription model, even a small percentage of monthly or quarterly client churn can have a big impact over the course of a full year. A churn rate of about 12% annually translates to a turnover rate of about 1% every month. Companies with high churn rates may quickly find themselves in a financial hole as they are compelled to commit an increasing amount of resources to the acquisition of new customers because it is substantially more expensive to acquire a new client than it is to keep an existing one.
One of the many sorts of surveys that are readily available online focuses on the expenses associated with acquiring and keeping customers. A new customer can cost up to five times as much to recruit as it would to keep an existing one, according to research done by Invesp, a business that specialises in the optimization of conversion rates.

High churn rates and low profits, as well as high costs for customer acquisition, are related. They "also have a more subtle impact on the possibility for growth of a company," according to Michael. In today's environment, customers are not shy about sharing their experiences with businesses on peer-to-peer networking sites, social media platforms, and review websites. HubSpot found that 49% of customers had tweeted about a great experience they had with a company. In a time when consumers are losing faith in the businesses they use, the power of word-of-mouth marketing is stronger than ever. The same HubSpot Research survey found that about 55% of consumers no longer have faith in stores, 65% do not accept press releases from retailers, 69% do not believe advertisements, and 71% do not believe sponsored social media posts.

Challenges:

There is a lot of transition happening in the banking and financial industries right now. Customers came and went at a somewhat slow rate annually. However, with the development of fintech and the increase in competition, customer churn has become an increasingly important factor in the struggle for control of the financial services market. A customer's financial obligations span a wide range of product and service categories.

A flowchart for the processing of customer churn data is shown in figure 1.

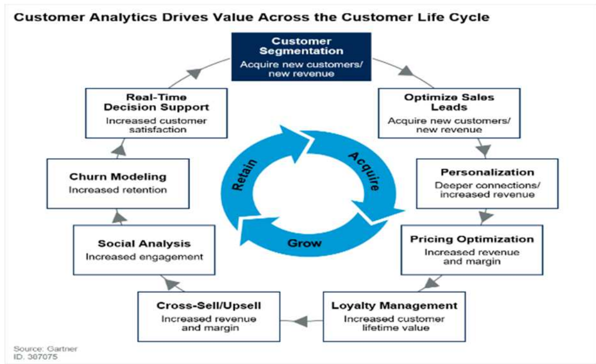


Figure.1: Step – by – Step Process of Churn Prediction Processing
Copyright Source: Heavy.AI

II. LITERATURE WORK

The facts presented in the Trust Survey by HubSpot Research article on Churn is likely to be discussed in the following discussion. As a direct consequence of this, customers' levels of tolerance and trust in corporate organizations have

significantly decreased. In today's hypercompetitive market, the majority of acquisition methods used by firms are based on incorrect mathematical assumptions.
You have the option of investing money into improving the quality of customer service, which is one of the most effective alternatives.

An authority in the field has discovered that businesses with high churn rates not only fail to keep contact with their former customers, but also hinder their efforts to bring in new customers by spreading negative word of mouth about the quality of the goods and services that they offer. This is because the former customers spread negative word of mouth about the products and services that the business provides.

CallMiner, a company that provides solutions for conversational analytics, carried out a study with a sample size of one thousand people to uncover the elements that influence customers' decisions regarding which businesses they will patronise. According to a recent survey, the annual losses sustained by U.S. corporations as a result of a drop in their client base are projected to be somewhere in the region of $136 billion. Because of this, businesses were given the opportunity to address the problems that had led customers to stop purchasing their products.

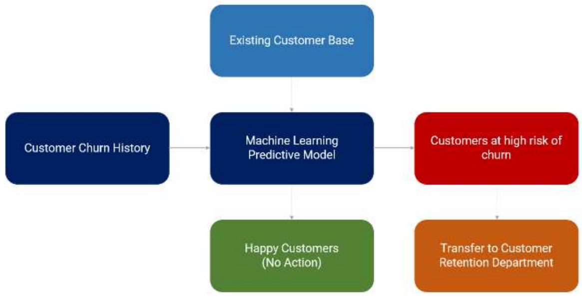


Figure 2.1 Churn Rate Prediction Using ML

Customers are more likely to stick with a business if it monitors how its products are used, solicits customer input, and promptly resolves any issues that are brought up by customers.

It is the goal of machine learning to create self-improving computer programmes that can sift through data on their own in search of recurrent patterns and make progress in the absence of explicit instructions. These are examples of the kind of digital actions that can be examined to foretell whether or not a consumer would stop using a given service.

Classification. Assigning a name to every possible cluster of information serves the overarching goal of categorization (here, customers). For classification problems, data scientists frequently use historical data annotated with labels (churner/non-churner) that need to be forecasted in order to train an algorithm. To do this, the data is utilised for algorithm training.

Regression. Regression analysis can also help you forecast how many customers will cancel your service. Regression analysis is a statistical technique used to identify relationships between a dependent variable and the independent factors that influence it. These supplementary benefits also take the form of an ever-expanding menu of alternatives. If that's too lot to take in, just remember that a regression model produces a numeric value, whereas a classification model produces a categorical label. If that sheds light on the situation, then wonderful! Regression analysis allows one to create informed estimates on the number of data elements that influence a target variable. Using regression, organizations can calculate the likeliness of customer attrition and the predicted duration of a customer's tenure as a customer.

The general SVM approach[21][22][23] creates a "black box" model that doesn't demonstrate how the information gained during training is translated into words that people can understand; M.A.H. Farquad [1] devised a hybrid strategy to circumvent this difficulty. He took this action to avoid more hassles. The hybrid method consists of the following three components: As a first step, we use SVM-RFE to prune the feature set (SVM-recursive feature elimination)[24]. Second, a support vector machine (SVM) model is constructed using a dataset with less features. Naive Bayes Trees[26] are produced at the conclusion of rule generation by combining decision trees with naive Bayesian classifiers[25]. It improves categorization thanks to this (NBTree).

Here, we make use of the Bank Credit Card Customer Dataset that was compiled for the 2004 Business Intelligence Cup. Recurring customers make up a large percentage of the total (93.24%), whereas inactive accounts make up a much smaller percentage (6.76%). The study showed that the model performed poorly when presented with a huge dataset.

By analysing a customer relationship management (CRM) dataset from American telecom providers, Chih-Fong Tsai [2] was able to foresee which consumers would churn. The dataset was sourced from customer relationship management systems. Using a combination of back-propagation artificial neural networks and self-organizing maps, we developed two hybrid models for churn prediction (SOM)[27].

"ANN + ANN" and "SOM + SOM" refer to hybrid models that combine SOMs and ANNs, respectively. In particular, one of the two hybrid models can cut down on the amount of data that needs to be processed by omitting unnecessary examples from the training set.

The second method of prediction relies on the findings from the first.

To evaluate the performance of these models, we employ three distinct testing sets: one general testing set, two fuzzy testing sets, and a testing set derived from the output of the initial ANN and SOM hybrid models. The outcomes demonstrate that the hybrid models outperform the single-network baseline model in terms of prediction accuracy. Specifically, it is demonstrated by the fact that the SOM + ANN hybrid model outperforms the ANN + ANN hybrid model.

In order to create churn prediction models that are not only accurate but also simple to understand, Wouter Verbeke [3] recommended employing the Ant-Miner+ and ALBA algorithms[29] on a publicly available churn prediction

dataset. This action was taken to better inform churn prediction algorithms. You can efficiently mine data with Ant-help.
To better estimate a customer's departure time, Ning Lu [4] proposed using boosting algorithms. These algorithms separate clients into two categories, the relative importance of which is determined by the boosting algorithm. This allowed us to determine which of our clients constituted the greatest risk to the security of our establishment.

Benlan He [5] suggested that the SVM model be used to discover out why clients quit. Because the customer data sets he was working with were unbalanced, he used random sampling to strengthen the SVM model. A support vector machine's job is to build a hyper-plane in a space with a lot or an infinite number of dimensions so that it can be used to sort items[30].

Ssu-Han Chen [6] developed a novel approach (CUSUM). Reference value and decision interval are calculated using a finite mixture model and a hierarchical Bayesian model, respectively, to account for the wide variety of consumer types. Using both of these methods, the gamma CUSUM chart records everyone's IAT (IAT).

For predicting when a client will leave, Koen W. De Bock [7] developed a number of rotation-based ensemble classifiers, including Rotation Forest and Rotboost. When multiple classifiers are used to make a prediction, the results from each are combined using a fusion rule to form an ensemble classifier. This is only one possible explanation of what an ensemble classifier is. To combine Rotation Forest and AdaBoost is referred to as "RotBoost."[31] In order to transform the input data into training data for the base classifiers, Rotation Forests perform feature extraction on feature subsets. The information used in this research came from four distinct studies that aimed to forecast how many consumers would quit. Rotation Forests do better than RotBoost in terms of both the area under the curve (AUC) and the top-decile lift, although RotBoost is more accurate.

In addition to RotBoost and Rotation Forest, the data classification abilities of three other feature extraction algorithms—Principal Component Analysis (PCA), Independent Component Analysis (ICA), and Sparse Random Projections—were investigated and compared to those of RotBoost and Rotation Forest (SRP) (SRP).
Both the feature extraction algorithm and the criteria used to evaluate classification performance play a role in determining how effective a rotation-based ensemble classifier is.

Lee et al. [8] set out to solve this problem by developing a churn prediction model that was accessible to a wide audience while still maintaining a high level of precision. The Partial Least Squares (PLS) approach was applied to datasets containing variables that were highly correlated with one another in order to accomplish this.

II. PROPOSED MACHINE LEARNING TECHNIQUES

a) Data Set for Customer Churn:

- The most important information about an individual can be found in their demographic data (e.g., age, education level, location, income)
- "User behavior features" are the parts of a service or product that are adapted to the particular activities of the user. This can refer to both software and hardware modifications (e.g., lifecycle stage, number of times they log in into their accounts, active session length, time of the day when a product is used actively, features or modules used, actions, monetary value)

- Interactions with customers are characterized by feature sets related to customer service (e.g., queries sent, number of interactions, history of customer satisfaction scores)

The location of a customer can be ascertained based on a number of different factors.
As can be seen in Figures 3.1a. and 3.1.b, our dataset is comprised of 79.45% was Retained and 20.55% was churn.

Row Number	Customer Id	Surname	Credit Score	Geography	Gender	Age	Tenure	Balance	NumOf Products	Has CrCard	IsActive Member	EstimatedSalary	Exited
1	1.6E+07	Hargrave	619	France	Female	42	2	0	1	1	1	101349	1
2	1.6E+07	Hill	608	Spain	Female	41	1	83807.9	1	0	1	112543	0
3	1.6E+07	Onio	502	France	Female	42	8	159661	3	1	0	113932	1
4	1.6E+07	Boni	699	France	Female	39	1	0	2	0	0	93826.6	0
5	1.6E+07	Mitchell	850	Spain	Female	43	2	125511	1	1	1	79084.1	0
6	1.6E+07	Chu	645	Spain	Male	44	8	113756	2	1	0	149757	1
7	1.6E+07	Bartlett	822	France	Male	50	7	0	2	1	1	10062.8	0
8	1.6E+07	Obinna	376	Germany	Female	29	4	115047	4	1	0	119347	1
9	1.6E+07	He	501	France	Male	44	4	142051	2	0	1	74940.5	0
10	1.6E+07	H?	684	France	Male	27	2	134604	1	1	1	71725.7	0

Figure 3.1.a Data set

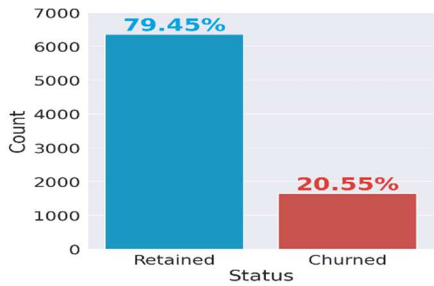


Figure 3.1b: Data set

Pre-processing:

A. 3.1 Classification of Machine Learning Techniques

Classic machine learning models, such as logistic regression, decision trees, NB, SVM and random forest, and others, are often used to predict customer loss.

a) SVM

Support Vector Machine (SVM) is an unsupervised approach of machine learning that trains data points along the hyper planes that divide them. This is in contrast to other supervised machine learning methods. In this essay, I'll show how I projected customer attrition for a made-up telecom company using SVM from Python's great scikit-learn machine learning library.It is feasible to make a prediction based just on the test data by using the support vectors. The SVM model relies heavily on support vectors to do its calculations. In order to make a forecast, take the support vectors and the test data point and multiply them together using the dot product. The table that follows:

- *Eliminate Unwanted Items: Row Number, Customer-id, Surname*
- *ENCODING*

Table: 3.1 Data set

Before encoding	After encoding
France	0
Germany	1
Spain	2
Before encoding	**After encoding**
Female	0
Male	1

The data encoding after pre-processing is shown in the table 3.1, and the balanced and unbalanced data after pre-processing are shown in figure 3.1.C

	CreditScore	Geography	Gender	Age	Tenure	Balance	Num Of Products	Has Credit Card	Is Active Member	Estimated Salary	Churn
0	619	0	0	42	2	0.00	1	1	1	101348.88	1
1	608	2	0	41	1	83807.86	1	0	1	112542.58	0
2	502	0	0	42	8	159660.80	3	1	0	113931.57	1
3	699	0	0	39	1	0.00	2	0	0	93826.63	0
4	850	2	0	43	2	125510.82	1	1	1	79084.10	0

Figure 3.1c: Balanced and Unbalanced Data set

Demonstrates that there is no linear relationship between the three variables of generalization error, model complexity, and the size of the training dataset.

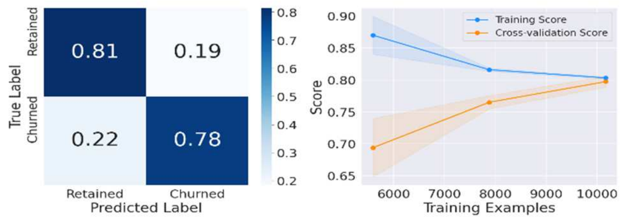
When contrasting the errors made during training with those made after generalization, for example,

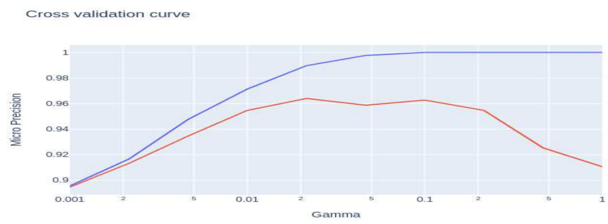
$$et + f\,(c,n,d) \qquad (1)$$

The complexity of a nonlinear function model is represented by F, the number of training data is represented by c n, and the error probability is represented by d.

The likelihood is at least if the mistake is smaller than the value on the right side of the formula (1-d) (1-d). error caused by the complexity of the model for a set amount of available training data" is the second component on the right. Earlier, in one of my comments, I actually used this very line.

Iteration is the only technique capable of finding the smallest possible training set necessary for a given level of complexity and error in the model. This is because of the nonlinear nature of the relationship between the two variables. The VC value, which represents the complexity of the model, is notoriously obscure, which makes the challenge even more challenging. On the other hand, there are some statements that can be made regarding quality that are generally recognized by everyone.





These are undoubtedly promising numbers to consider. Both our test and train curves have unusually high area under the curve (AUC) values, which are concentrated in the upper left corner. Given that we now have solid models, we can concentrate on optimizing a small number of model parameters and hyperparameters to gradually improve our results.

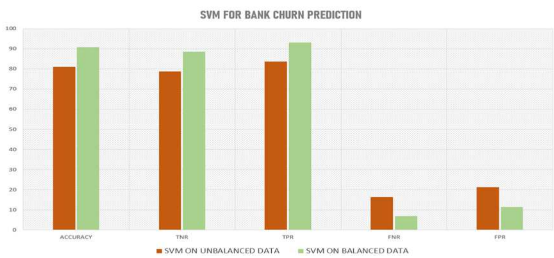


Figure 3.5 Shows the ML Output on Balanced and Unbalanced Data for SVM

Naive Bayes

The Naive Bayes model[32] computes the overall probability of an occurrence by using the probabilities that are linked with each underlying variable. This exemplifies the overall likelihood that the occurrence will take place.

P(a|b): We suffer a loss of clients (A) whenever anything negative (b) takes place (as was anticipated), where b is one of our data points. The value of P is equal to the combination of the two variables (e.g., Tenure, Age, Credit Score). An individual of the age of 30 with a FICO score of 350 who has been a loyal customer for the past two years is an illustration of a churning customer.

P(a|b) is referred to as "naive" due to the fact that it is computed independently for each variable. As a result, it is based on the incorrect assumption that changing one variable will not have an impact on the others. This is due to the fact that the probability P(a|b) is computed in a manner that is distinct for both a and b.

The likelihood that event B will take place if event A has already taken place is represented by the mathematical notation P(b|a) (for example, if the customer is 18 years old and we have a dataset of old customers who have left the bank, we can create a probability that reflects the probability of customer churning if the customer is 18 years old). This is the case independent of any other circumstances, such as the applicant's credit history or score, the length of time the employee has been working for the organization, or any other variables.

P stands for the possibility that a consumer will withdraw their funds from a bank, and it is used to signify this probability (A). One hundred out of every thousand of our customers has the option of terminating their contract with us.

B will occur in the future (B). How much of a difference is there between the probability that one of the customers is 18 and the probability that all of the customers are 18?

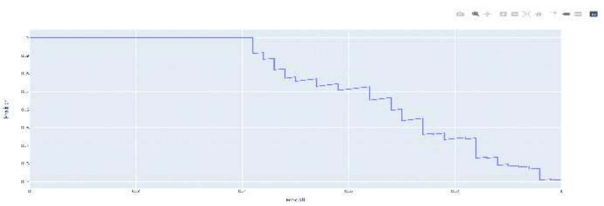


Figure 3.6 Shows the ROC of NB

The Naive Bayes technique receives support from both the ROC curve and the Confusion Matrix. The percentages that fall within the True Positive and True Negative quadrants account for a combined total of ninety percent of the total test results. That indicates that our model correctly labelled or categorized 90% of the data 90% of the time. The area score for the ROC was 0.89, which was 0.12 points higher than the score for the logistic regression (0.71). Based on these findings, it appears that the Logistic Regression model is surpassed by around 20% by the Naive Bayes algorithm when it comes to anticipating customer attrition.

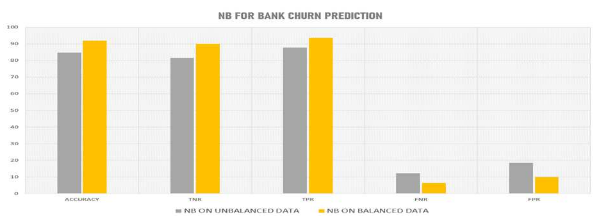


Figure 3.7 Shows the Data set of Number of Subplots

IV. RESULT AND DISCUSSIONS

Table 4.1 and Figure 4.1 shows, the results of classifying items by only employing unigrams as the criterion for classification. Random selection has a far lower possibility of providing the correct outcome in comparison to the machine learning method, which has an accuracy rate of 90%. In spite of having access to fewer test data, the Naive Bayes algorithm performs better than the SVM algorithm. Even though the utilization of LGBM Classifier might result in an increase in accuracy in future.

Table 4.1 ML Output on Balanced and Unbalanced Data

Mod el	Data	Accur acy	TNR	TPR	FNR	FPR
SVM	Unbalance d Data	81.05	78.68 07	83.64 78	16.35 22	21.31 93
	Balanced Data	90.8	88.49 29	93.02 55	69744 6	11.50 71
NB	Unbalance d Bta	84.75	81.58 7	87.80 73	12.19 27	18.41 5
	Balanced Data	91.95	90.08 18	93.73 78	6.262 23	9.918 2

As discussed in Section 3, results from Machine Learning algorithms are considerably superior to human-made baselines. It can be said that the relative performances of Naive Bayes are extremely high comparable to one another. With an accuracy of 84.75% and 91.95%. the NB algorithm is recommended and noticeably more effective than other algorithms.

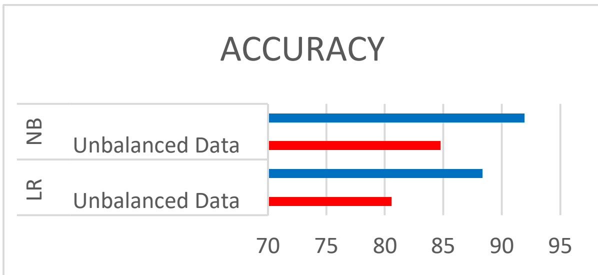


Figure 4.2 shows that NB has better Accuracy.

The results of this study show that customer churn prediction using ML algorithms has been implemented successfully and tested with positive accurate results.

V. Conclusion

In this study, a deeper investigation into the methods that are used to forecast the likelihood of contract churn among existing customers is undertaken. All of these models are flawed, and they are not very good at estimating the number of customers who will stop using a service. Customers can be kept from switching to a competitor if they are presented with a trustworthy model of what the future holds for the company. By combining LGBM – Classifiers with various boosting strategies, you can improve both the accuracy and performance of your work, which will bring you closer to achieving your objective. It is possible that in the not-too-distant future, these data will be used to compute churn, which would be of great assistance. It is necessary to continually improve prediction models, and a combination of the strategies described above is advised as the best course of action.

References

[1] M.A.H. Farquad, Vadlamani Ravi, S. Bapi Raju “Churn prediction using comprehensible support vector machine: An analytical CRM application”, Applied Soft Computing 19 (2014) 31–40.

[2] Chih-Fong Tsai, Yu-Hsin Lu "Customer churn prediction by hybrid neural networks", Expert Systems with Applications 36 (2009) 12547–12553.

[3] Wouter Verbeke, David Martens, Christophe Mues, Bart Baesens "Building comprehensible customer churn prediction models with advanced rule induction techniques", Expert Systems with Applications 38 (2011) 2354–2364.

[4] Ning Lu, Hua Lin, Jie Lu, Guangquan Zhang "A Customer Churn Prediction Model in Telecom Industry Using Boosting", IEEE Transactions on Industrial Informatics, vol. 10, no. 2, may 2014.

[5] Benlan He, Yong Shi, Qian Wan, Xi Zhao "Prediction of customer attrition of commercial banks based on SVM model", Proceedings of 2nd International Conference on Information Technology and Quantitative Management (ITQM), Procedia Computer Science 31 (2014) 423 – 430.

[6] Ssu-Han Chen, "The gamma CUSUM chart method for online customer churn prediction", Electronic Commerce Research and Applications, 17 (2016) 99–111.

[7] Koen W. De Bock, Dirk Van den Poel, "An empirical evaluation of rotation-based ensemble classifiers for customer churn prediction", Expert Systems with Applications 38 (2011) 12293–12301.

[8] H. Lee, Y. Lee, H. Cho, K. Im, Y.S. Kim, "Mining churning behaviors and developing retention strategies based on a partial least squares (PLS) model", Decision Support System 52 (2011) 207–216.

[9] Shrivastava, A.; Prakash Arul Jose, J.; Dilip Borole, Y.; Saravanakumar, R.; Sharifpur, M.; Harasi, H.; Abdul Razak, R.K.; Afzal, A., A study on the effects of forced air-cooling enhancements on a 150 W solar photovoltaic thermal collector for green cities, Sustainable Energy Technologies and Assessments, 2022, DOI: 10.1016/j.seta.2021.101782

[10] Anurag Shrivastava; Ali Rizwan; Neelam Sanjeev Kumar; R. Saravanakumar; Inderjit Singh Dhanoa; Pankaj Bhambri; Bhupesh Kumar Singh; Samarendra Nath Sur, VLSI Implementation of Green Computing Control Unit on Zynq FPGA for Green Communication,Wireless Communications and Mobile Computing2021-11-30, DOI: 10.1155/2021/4655400

[11] D. Haripriya; Keshav Kumar; Anurag Shrivastava; Hamza Mohammed Ridha Al-Khafaji; Vishal Moyal; Sitesh Kumar Singh; Mohammad R Khosravi, Energy-Efficient UART Design on FPGA Using Dynamic Voltage Scaling for Green Communication in Industrial Sector, Wireless Communications and Mobile Computing, 2022-05-05, DOI: 10.1155/2022/4336647

[12] Anurag Shrivastava; D. Haripriya; Yogini Dilip Borole; Archana Nanoty; Charanjeet Singh; Divyansh Chauhan, High performance FPGA based secured hardware model for IoT devices, International Journal of System Assurance Engineering and Management,2022-03,DOI: 10.1007/s13198-021-01605-x

[13] Suresh Kumar; S. Jerald Nirmal Kumar; Subhash Chandra Gupta; Anurag Shrivastava; Keshav Kumar; Rituraj Jain; Punit Gupta, IoT Communication for Grid-Tie Matrix Converter with Power Factor Control Using the Adaptive Fuzzy Sliding (AFS) Method, Scientific Programming,2022-03-29, DOI: 10.1155/2022/5649363

[14] K. Murali Krishna; Amit Jain; Hardeep Singh Kang; Mithra Venkatesan; Anurag Shrivastava; Sitesh Kumar Singh; Muhammad Arif, Development of the Broadband Multilayer Absorption Materials with Genetic Algorithm up to 8 GHz Frequency,Security and Communication Networks,2022-02-17, DOI: 10.1155/2022/4400412

[15] Anurag Shrivastava; Chinmaya Kumar Nayak; R. Dilip; Soumya Ranjan Samal; Sandeep Rout; Shaikh Mohd Ashfaque, Automatic robotic system design and development for vertical hydroponic farming using IoT and big data analysis,Materials Today: Proceedings, 2021-07,DOI: 10.1016/j.matpr.2021.07.294

[16] Anurag Shrivastava; Rajneesh Sharma; Mohit Kumar Saxena; V. Shanmugasundaram; Moti Lal Rinawa; Ankit, Solar energy capacity assessment and performance evaluation of a standalone PV system using PVSYST,Materials Today: Proceedings,2021-07, DOI: 10.1016/j.matpr.2021.07.258

[17] Srivastava, A., Singh, A., Joseph, S.G.Borole, Y.D., Singh, H.K.,WSN-IoT Clustering for Secure Data Transmission in E-Health Sector using Green Computing Strategy,2021 9th International Conference on Cyber and IT Service Management, CITSM 2021, 2021

[18] Shrivastava, A.; Ranga, J.; Narayana, V.N.S.L.; Chiranjivi; Borole, Y.D., Green Energy Powered Charging Infrastructure for Hybrid EVs, 2021 9th International Conference on Cyber and IT Service Management, CITSM 2021,DOI: 10.1109/CITSM52892.2021.9589027.

[19] Soumi Ghosh, Ajay Rana, Vineet Kansal, "A Nonlinear Manifold Detection based Model for Software Defect Prediction", Procedia Computer Science, Volume 132, 2018, Pages 581-594, ISSN 1877-0509, https://doi.org/10.1016/j.procs.2018.05.012.

[20] G. Dubey, A. Rana and N. K. Shukla, "User reviews data analysis using opinion mining on web," 2015 International Conference on Futuristic Trends on Computational Analysis and Knowledge Management (ABLAZE), 2015, pp. 603-612, doi: 10.1109/ABLAZE.2015.7154934.

[21] H. Walia, A. Rana and V. Kansal, "A Naïve Bayes Approach for working on Gurmukhi Word Sense Disambiguation," 2017 6th International Conference on Reliability, Infocom Technologies and Optimization (Trends and Future Directions) (ICRITO), 2017, pp. 432-435, doi: 10.1109/ICRITO.2017.8342465.

[22] A. Gupta, A. Pachauri, P. Pachauri, S. V. Singh, P. Chaturvedi and S. Sharma, "A review on conglomeration of Technologies for Smart Cities," 2021 International Conference on Technological Advancements and Innovations (ICTAI), 2021, pp. 526-530, doi: 10.1109/ICTAI53825.2021.9673458.

[23] V. Kumar, D. Singh, S. V. Singh and T. Anand, "The Role of Converged Network in Disruptive Technology," 2020 International Conference on Intelligent Engineering and Management (ICIEM), 2020, pp. 483-486, doi: 10.1109/ICIEM48762.2020.9160337.

[24] Singh, G., Bhardwaj, G., Singh, S. V., and Kumar, V. 2020. "Technology Acceptance Model to Assess Employee's Perception and Intention of Integration of Artificial Intelligence and Human Resource Management in IT Industry," Technology (29:3), pp. 11485–11490.

[25] Singh, G., Bhardwaj, G., Singh, S. V. & Garg, V. Biometric Identification System: Security and Privacy Concern. in Artificial Intelligence for a Sustainable Industry 4.0 (eds. Awasthi, S., TraviesoGonzález, C. M., Sanyal, G. & Kumar Singh, D.) 245–264 (Springer International Publishing, 2021).

[26] S. A. Yadav, S. Sharma and S. R. Kumar, "A robust approach for offline English character recognition," 2015 International Conference on Futuristic Trends on Computational Analysis and Knowledge Management (ABLAZE), 2015, pp. 121-126, doi: 10.1109/ABLAZE.2015.7154980.

[27] S. R. Kumar, S. A. Yadav, S. Sharma and A. Singh, "Recommendations for effective cyber security execution," 2016 International Conference on Innovation and Challenges in Cyber Security (ICICCS-INBUSH), 2016, pp. 342-346, doi: 10.1109/ICICCS.2016.7542327.

[28] S. A. Yadav and T. Poongodi, "A Review of ML Based Fault Detection Algorithms in WSNs," 2021 2nd International Conference on Intelligent Engineering and Management (ICIEM), 2021, pp. 615-618, doi: 10.1109/ICIEM51511.2021.9445384.

[29] P. Choudhary, S. A. Yadav, A. P. Srivastava, A. Singh and S. Sharma, "A System for Remote Monitoring of Patient Body Parameters," 2021 International Conference on Technological Advancements and Innovations (ICTAI), 2021, pp. 238-243, doi: 10.1109/ICTAI53825.2021.9673325.

[30] Ruchika Gupta, Siddharth Varma, Garima Bhardwaj, A Structural Equation Model to Assess the Factors Influencing Employee's Attitude & Intention to Adopt BYOD (Bring Your Own Device), International Journal of Recent Technology and Engineering (IJRTE), 2019, PP: 6306-6308 | Volume-8 Issue-3 September 2019 | Retrieval Number: C5942098319/2019©BEIESP | DOI: 10.35940/ijrte.C5942.098319.

[31] A. Singh, A. P. Srivastava, G. Bhardwaj, A. Jain, A. K. Singh and Ritika, "Methods to Detect an Event using Artificial Intelligence and Machine Learning," 2022 3rd International Conference on Intelligent Engineering and Management (ICIEM), 2022, pp. 297-301, doi: 10.1109/ICIEM54221.2022.9853147.

[32] Gupta, R., Phanden, R.K., Sharma, S., Srivastava, P., Chaturvedi, P. (2021). Security in Manufacturing Systems in the Age of Industry 4.0: Pitfalls and Possibilities. In: Phanden, R.K., Mathiyazhagan, K., Kumar, R., Paulo Davim, J. (eds) Advances in Industrial and Production Engineering. Lecture Notes in Mechanical Engineering. Springer, Singapore. https://doi.org/10.1007/978-981-33-4320-7_10.

[33] S. A. Siddiqui, A. Ahmad, A. A. Siddiqui and P. Chaturvedi, "Stability Analysis of a Cantilever Structure using ANSYS and MATLAB," 2021 2nd International Conference on Intelligent Engineering and Management (ICIEM), 2021, pp. 7-12, doi: 10.1109/ICIEM51511.2021.9445357.

[34] P. Surya, P. Pachauri, A. Pachauri, P. Chaturvedi, S. A. Yadav and D. Singh, "Virtualization Risks and associated Issues in Cloud Environment," 2021 International Conference on Technological Advancements and Innovations (ICTAI), 2021, pp. 521-525, doi: 10.1109/ICTAI53825.2021.9673424.