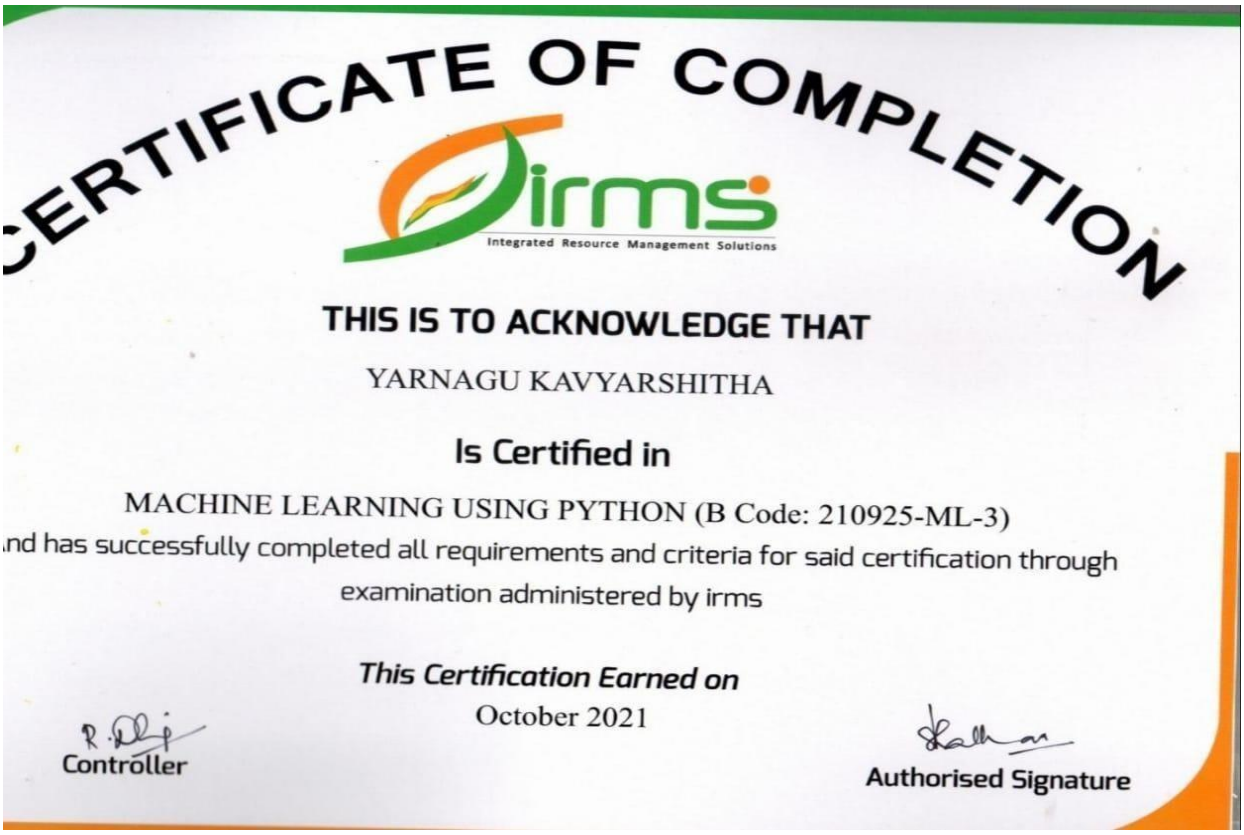
TRAINING CERTIFICATE



ABSTRACT

Now a -days there are a lot of service providers are available in every business. There is no shortage of customers in any options. Mainly, in the banking sector when want to keep their money safely they have a lot of options. As a result, customer churn and loyalty of customers have become a major problem for most banks. In this paper, a method that predicts customer churn in banking using Machine learning with ANN. This research promotes the exploration of the likelihood of churning by customer loyalty.

The Random Forest, SVM, KNN, and Decision Tree Machine Learning algorithms are used in this study. Keras and TensorFlow are ANN concepts that are also used in this study. This study is done on a dataset called churn modeling. The dataset was collected from Kaggle. The results are compared to find an appropriate model with higher accuracy. As a result, the Random Forest algorithm achieved higher accuracy than other algorithms. And accuracy was nearly 87%. The least accuracy was achieved by the Decision tree algorithm and it was 78.3% accuracy.

The number of service providers are being increasing very rapidly in every business. In these days, there is no shortage of options for customers in the banking sector when choosing where to put their money. As a result, customer churn and engagement have become one of the top issues for most banks. In this project, a method to predict the customer churn in a Bank, using machine learning techniques, which is a branch of artificial intelligence is proposed. The research promotes the exploration of the likelihood of churn by analyzing customer behavior. Customer Churn has become a major problem in all industries including the banking industry and banks have always tried to track customer interaction so that they can detect the customers who are likely to leave the bank.

TABLE OF CONTENTS

LIST OF FIGURES

LIST OF TABLES

LIST OF ABBREVIATIONS

ABBREVIATION	EXPANSION
SVM	Support Vector Machine
KNN	K-Nearest Neighbor
RF	Random Forest
DT	Decision Tree

1. INTRODUCTION

Churning means a customer who leaves one company and transfers to another company. It is not only a loss in income but also other negative effects on the operations and also mainly Customer Relation Management is very important for banking when the company considers it as they try to establish long-term relationships with customers and also it will lead to increase their customer base. The service provider's challenges are found in the behavior of the customer and their expectations. In the current generation, people are mostly educated compared to previous generations. So, the current generation of people is expecting more policies and their diverse demand for connectivity and innovation. This advanced knowledge is leading to changes in purchase behavior. This is a big challenge for current service providers to think innovatively to reach their expectations.

Private sectors need to recognize customers Liu and Shih strengthen this argument in their paper by indicating that increasing pressures on companies to develop new and innovative ideas in marketing, to meet customer expectations and increase loyalty and retention. For Customers, it is very easy to transfer their relations from one bank to another bank. Some customers might be keeping their relationship status null that means they will keep their account status inactive. By keeping this account inactive it might be the customer transferring their relationship with another bank. There are different types of customers are in the bank. Farmers are one of the major customers to the banks they will expect fewer monthly chargers as they were financially low. Businessperson, are also one of the major and important customers because a lot of transactions with huge amount is done by them only usually. These customers will expect better service quality. One of the most important categories was Middle-class customers, mostly in every bank these peoples are more than the type of customers. These people will expect fewer monthly charges, better service quality, and new policies.

So, maintaining different types of customers is not that easy. They need to consider customers and their needs to resolve these challenges delivering reliable service on

time and within budget to customers. While maintaining a good working partnership with them is another significant challenge for them. If they failed to resolve these challenges this may cause churning. Recruiting a new customer is more expensive and harder than keeping already customers. Customers holding on the other hand is usually more expensive because they have already gained the confidence and loyalty of present customers. So, the need for a system that can predict customer churn effectively in the early stages is very important for any banking. This paper aims at a framework that can predict the customer churning banking sectors using some machine learning algorithms with ANN.

2. LITERATURE SURVEY

There are a lot of studies done on customer churning and a lot of research has been done on this concept. One of the studies was done by D-R.Liu &Y.Y.sinh [1] on “Integrating ahp and data mining for product recommendation based on customer lifetime value [1]" in the year 2005. In this study they mainly focused on improving the quality of recommendation to fulfill customer's needs is important in most competitive environments. The major advantage in this was although various recommender system has been found, few have labeled the lifetime value of a customer to the firm. They developed a product recommendation methodology that combined group decision making and data mining techniques and the major disadvantage was, it is more fruitful for loyal customers but not done for less loyal customers.

In another study, G.Canning Jr[2] has done a study on “Do a value analysis of your customer base[2]” in the year 1982. In this study, he mainly focused on the company’s customer base can yield valuable data on customer sales volumes, profits, growth trends, and technical strength. The main advantage in this research was, there he used the CVA process that can establish this value in a way that leads to an effective marketing program and the major disadvantage was marketing source of a company is its customers the experience accumulated through serving them in the past is of substantial value in serving them in future.

In another study which was done by, P.G.Nieto, E.F.Combarro, J.del coz Dia and E.Montanies [3] on "An SVM based regression, model to study the air quality of local

scales in Oviedo urban area [3]" in 2013. They focuses on a method of daily air pollution modeling by using support vector machine procedure in Oviedo urban areas at a regional scale. In this study, the main advantage is SVR fit captures the prime idea of statistical learning theory to get a good anguring of the dependence among the main pollution in the city of Oviedo and the main disadvantage was, the methodology used in this study can be applied successfully to other cities with similar or different sources of pollutants, but it is always compulsory to take into account the particular of each location.

In another study which was done by, M.K.Kim, M.C.Park, and D.H.Jeong [4] on "The results of customer happiness and switching barriers customer loyalty in Korean mobile telecommunication services [4]" in 2004. They mainly focused on, the Korean mobile telecommunication services industry is entering a new transition period. The main advantage of this study was the effects on customer trust of client satisfaction and the switching barrier and the structural relationship between these factors in the Korean mobile telecommunication on the service industry. The main disadvantage is, there are other factors influencing customer loyalty, apart from factors suggested in this study, such as the demographic characteristics of customers and their usage pattern of mobile telecommunication services.

In another study, B.He, Y.shi, Q. Wan, and X.zhao [5] has done a study on the Prediction of customer attrition of commercial banks based on the SVM model [5]" in 2014. Mainly they focused on Chinese commercial banks that are facing triple huge pressure, including financial churn rate, interest rate marketization and internet finance. In this study, the Main advantage was, commercial bank client churn prediction based on the SVM model and uses random sampling method to improve the SVM model, in view of the sample imbalance characteristics of customer data sets. This method can effectively enhance the prediction accuracy of the selected model and the main disadvantage of this study is, Due to the imbalanced characteristics of the actual commercial bank customer churn dataset. SVM model cannot predict the churners effectively and only general rating criteria cannot measure the predictive power of the model.

In another study that was done by A.Bilal Zoric [6] in 2016, on "Predicting customer

churn in banking industry using neural networks [6]". In this study, he mainly focused

on the usage of the data mining methods, neural networks, in knowledge discovery from databases in the banking industry. The main advantage was the risk of leaving and analyzing whether those customers are worth retaining. Clients who use more bank services are more loyal, so the bank should focus on those clients who use less than three products and offer them products according to their needs and Main disadvantage of this study was the analysis focused on churn prediction based on only one method and a neural network. They could access other important information that could help banks to get a competitive advantage by using other methods such as segmentation, Decision trees, self-organizing maps.

In another study, A.K.Ahmed, A.Jafar, and K.Aljoumaa [7] on "Client churn prediction in telecom company using machine learning algorithms in big data platform [7]" in 2019. In this study, they mainly focused on customer churn, it is a major problem and one of the most important concerns for large companies. Due to the effect on the revenue of the companies, especially in the telecom companies. These companies are seeking to develop means to predict potential customers to churn. In this study, the main advantage was, working on a large dataset created by changing big underdone data provided by Syriatel Telecom company. Best results were obtained by applying the XGBOOST algorithm and the main disadvantage of this study was, the method was restricted to customer churn in the telecom industry. This method focuses only on the 2 telecom companies in Syria which are Syria Tel and MTN telecom companies.

In another study which was done by Y.Jiang and C.Li [8] on" MRMR based feature selection for classification of cotton foreign matter using hyperspectral imaging [8]" in 2015. They mainly focused on different cotton and decreases the monetary value of cotton. Hyperspectral imaging foreign matter. The main advantage of this study was, it explored a new 2 stage approach for hyperspectral imaging optimal wavelength selection. A total of 12 wavelengths in the visible/NR range were selected for cotton FM classification and the main Disadvantage was, It is not focused on the implementation of automatic FM image segmentation and classification method using the selected wavelength.

In another study was done by J. Raj and V. Ananthi[9] on "Recurrent neural networks