

Predicting The Success of Bank Marketing using Classification Techniques

S.Vaishnavi

Assistant Professor

*RMKCET College of Engineering & Technology Chennai,
India*

Gattu Divya Lakshmi

Student

*RMKCET College of Engineering & Technology Chennai,
India*

Dandu Vinisha Reddy

Student

*RMKCET College of Engineering & Technology Chennai,
India*

Majeti Satya Naga Sulochana

Student

*RMKCET College of Engineering & Technology Chennai,
India*

Abstract

Marketing plays prominent role in any business field. It is very important in any field to promote its product to gain profit. This marketing or campaigning can be done in two ways one is direct interaction and other is through telecommunication that means communicating or interacting with the people through messages, emails or any other social media approach. For a person who is campaigning a product it is very important to know the reach of the product so that he could do any modifications that are needed, in his way of approach for promoting a product. In this project we are considering banking system, we are considering the data and going to find the persons who have subscribed for a scheme through different data mining techniques. By finding the number of persons subscribed to the scheme we can know up to what extent that scheme has reached to the people. So if the bank system was not happy with the subscriptions they change the way of promoting. In our project we are comparing the predictions using different data mining techniques like Decision tree(DT), Naïve Bayes, Neural Networks(NN), KNN and support vector machine(SVM). Then we find the techniques that gives us the highest accuracy of prediction.

Keywords: SVM, DT, NN, KNN

I. INTRODUCTION

The main aim of any business system is to gain good profit. A well planned strategy plays a key role in profit gaining. Generally business systems have many schemes that give them good profit. So whenever a scheme is implemented it is very important to know its success rate. Marketing plays very important role in the success of any business scheme. Whenever, any product or a scheme is introduced it is very important for the producers of that product to take that scheme or product into the people. So that if people are interested, they can buy it or subscribe to the scheme. In today's technology there are many schemes such as Amazon, Bank subscriptions etc., for all these channels it is very important to know its reach to people.

In our project we are going to consider banking system. We are going to predict the success of the scheme of the bank telemarketing. We are going to use different classification techniques for this purpose. In our project we are going to use different classification techniques like Decision tree (DT), Neural Networks (NN), Random forest, KNN and support vector machine (SVM). These are different supervised and unsupervised machine learning techniques. All these are very efficient classification techniques. We are using these classification techniques and going to predict the number of persons who have subscribed for the scheme. So that we can know up to which extent the product has reached the people.

We are going to consider this problem as a machine learning classification technique.

If we are having a group of data, classification is technique of categorizing the data into groups based on different properties. Some of the examples of this classification are pattern recognition, differentiating spam and non-spam mails etc., the property of classification varies for different types of data. This classification techniques are of two types. Supervised and no supervised classification techniques.

In supervised technique first we are going to train the machine with prior data and then we are going to introduce new set of data. Based on the training data the machine will classify the new testing data. In this supervised classification techniques we will train the data with a well labeled and structured data and we know the class or label of the output data in prior.

The Unsupervised classification technique is completely different from the supervised classification technique. In this technique we can also take unorganized and unlabeled data as input. This technique is also called as self-organizing technique. When the data is not even or labeled then we prefer this type of classification technique.

II. LITERATURE SURVEY

In industry the major role is marketing promotion. In order to inculcate the changes it is thus necessary to understand the human decision making process. By finding the number of persons subscribed to the scheme we can know up to what extent that scheme has reached to the people [1]. As the world is technologically advancing and going more individualistic there is a need to understand the human decision making process. A scheme is implemented it is very important to know its success rate [2]. In today's technology there are many schemes such as Amazon, Bank subscriptions etc..., for all these channels it is very important to know its reach to people [3]. DT, KNN, Random forest, SVM are supervised and unsupervised machine learning techniques. All these are very efficient classification techniques [4].

Going deeper on human communications we understand that the human body stimulates to the five senses. They are: Sight – Seeing, Ear –Hearing, Skin-Touch, Tongue-Taste, and Nose-Smell. The decisions are taken by people based on these five senses. Further on as the technology develops these four mediums could incorporate the other senses like touch, taste, smell. But as of present the only two mediums which are used for communication could be utilized.

Traditionally the marketing promotion has been divided into five or more categories depending upon their way of promotion. Promotion is defined as one of five pieces in the promotional mix or promotional plan. They are personal selling, advertising, sales promotion, direct marketing, and publicity. Thus the traditional view of promotion mix categories are based on the traditional approaches to the marketing egg.

- Direct marketing: talks about the approach of the marketer to communicate with the customers.
- Advertising: which mostly deals with the mass media and publicity of the product.
- Sales promotion: It deals with the customer attraction and offerings.
- Personal Selling: which deals with the customization of the services or product according to the needs of customer.
- Publicity: It is gaining public visibility or awareness for a product.

The world has been changed a lot since the definitions of these traditional ways to promote were innovated defined and utilized. Today we can see a sea change in the way in which practically the promotions, campaigns are planned and done. The definitions which were set earlier about the approaches to the marketing and understanding the marketing management are changing day by day, hour by hour, minute by minute and second by second. Today what is relevant is irrelevant tomorrow and un-useful the next hour. So fast is the understand-ability of the business is changing that some new ways and models have become the ways to conduct the business. The businesses have become so proactive, innovative and competitive that the traditional means are insufficient to understand them in a better way. The buying behavior analysis of customer are playing a key role in the in fast changing market approach. Decision tree [5] generally classifies the attributes and predicts whether it is suitable for given statement or not and returns the result associated to the leaf nodes. A Naive Bayes classifier may be a probabilistic machine learning model that's used for classification task. The naïve Bayes classifier follows the Bayes theorem [6]

The traditional ways on which customer decisions were made are fast changing.

III. THEORETICAL ANALYSIS

A. Existing Methods:

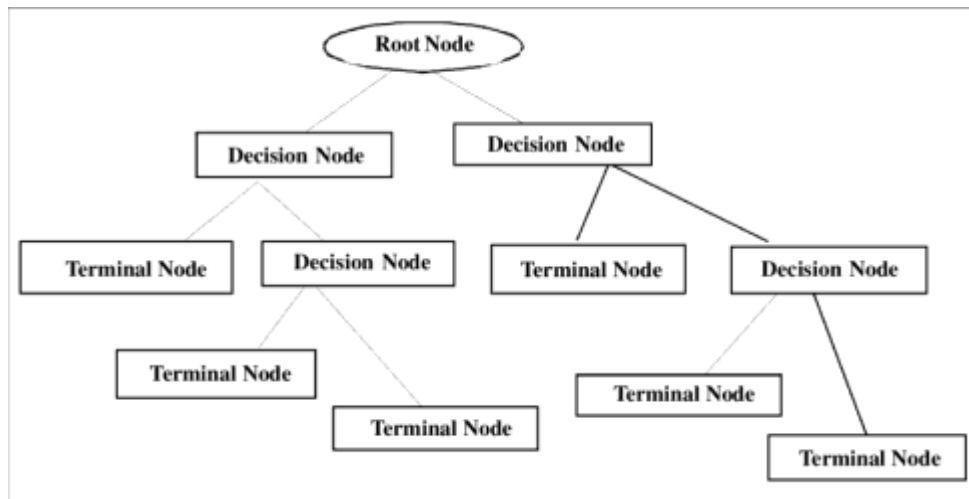
In literature survey we have seen about how the marketing works and how the marketing is very essential to gain profit. We have seen the traditional way of marketing ways and also need to change the ways of promoting and also seen the modern approaches of promoting. All these are part of marketing and then the reach of the product is very important in our project we are going to use different classification techniques to find the reach of the product that means we are going to know how many people have subscribed for a given scheme in banking system.

1) Dataset Description:

The main aim of our project is to find how many persons have subscribed to a scheme of banking system. So based on some details we are going to find the subscription rate. We are going to consider different kind of factors like age, job, marital status, income etc..., in our data totally we are going to consider totally 16 attributes with 4521 instances.

a) Data pre-processing:

Data pre-processing is a process of converting an unstructured or an unlabelled raw data into useful structured data or labeled data. There are different steps involved in data pre-processing they are Dealing with missing data, categorical data, dividing data into training and testing data, future predictions.



Dealing with the missing data: There are many ways to deal with the missing data they are deleting the data with missing values, Imputation (mean/median/knn/mode), using different packages in R. These are some of the ways to deal with the missing data.

- Deletion: In this methods the rows or the columns are deleted in order to get good accuracy.
- Imputation: In this method mean/mode/median/knn of the corresponding values are imputed in the place of missing values.
- Packages: R supports many packages that deals with the missing data. Some of the packages are MICE, miss Forest, Hmisc etc.,

B. PREDICTION TECHNIQUES

1) Decision Tree:

Decision tree is the most used machine learning classification technique generally used for prediction. Decision tree consists of root node, internal node and terminal node. Each node consists of branches, which indicates the outcome of the test. Each node is an attribute of the given data and the leaf node has the values of class labels. Decision tree can be constructed by a technique called as recursive partitioning. In recursive partitioning we will split the given dataset into small subsets considering the attribute values. Then the subsets are again divided into subsets in recursive method. When the subset of a node has the same value of the target variable then the recursive method is halted. Decision tree is used to classify the high dimensional datasets. Construction of a decision tree does not require any domain knowledge and setting of parameters. Decision tree gives the result with high accuracy. Decision tree classifies the instances from root to bottom of the tree based on attribute values. Decision tree [5] generally classifies the attributes and predicts whether it is suitable for given statement or not and returns the result associated to the leaf nodes.

Decision tree represents the disjunction of conjunctions of the attribute constraints. Decision tree is used to classify both categorical and continuous variables. Decision trees requires low computation for classifying a dataset. Decision trees identifies which attribute is more important in prediction of classification. In case of small training values decision trees leads to small prone errors. Decision trees are expensive to train computationally.

2) K- Nearest Neighbors (KNN):

K-nearest neighbors (KNN) algorithm is a supervised machine learning algorithm that is used for both classification and regression predictive problems. KNN uses all the data for training while classification. K-nearest neighbors (KNN) algorithm predict the values of new data points by use of 'feature similarity'. Then the new data point will be assigned a value based on how closely it matches the points in the training set. In the first step of the KNN algorithm, we must load the train and the test data. Next, choose the value of K i.e. the nearest data points. K can be any integer. The distance between test data and each row of training data is calculated using the Euclidean distance method. Based on the distance value, they are sorted in ascending order. Then it will choose the top K rows from the sorted array. In the final step, it will assign a class to the test point based on the most frequent class of these rows. It is computationally a bit expensive algorithm and has high memory storage.

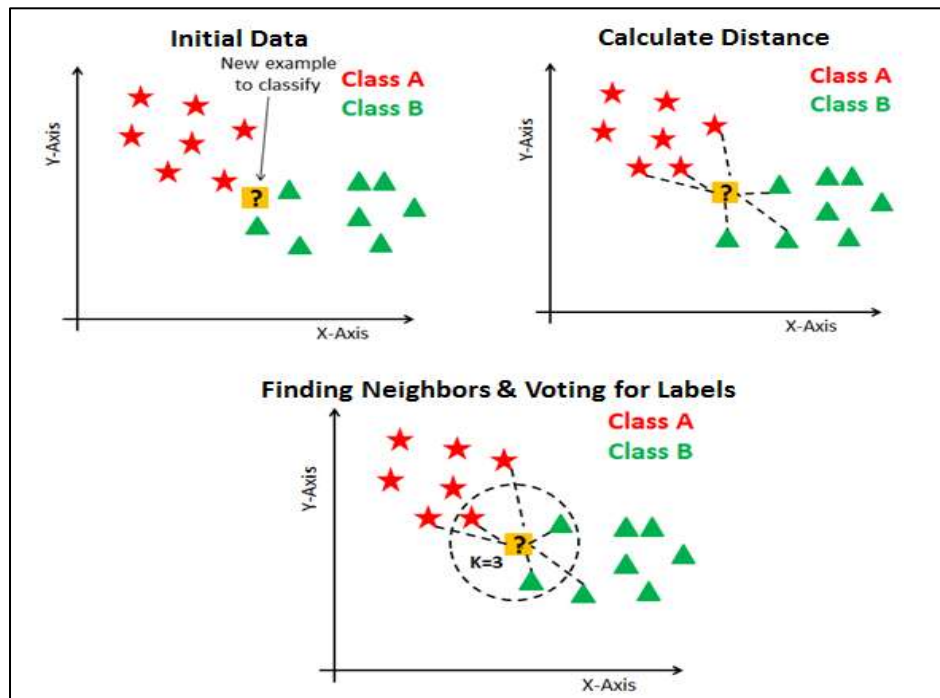


Fig. 4: K-Nearest neighbors

3) Support Vector Machine (SVM):

Support vector machines (SVMs) flexible supervised algorithms and have their unique way of implementation as compared to other machine learning algorithms. SVMs have the ability to handle multiple continuous and categorical variables. SVM model is a representation of different classes in a multidimensional hyper plane. The hyper plane is generated in an iterative manner by SVM so that the error can be minimized. Support vectors, Hyper plane, and Margins are the important concepts in SVM. The objective of SVM is to segregate the datasets into classes that are used to find a maximum marginal hyper plane(MMH). First, SVM will generate the hyper planes iteratively that segregates the classes in the best way. Later, it will choose the hyper plane that separates the classes correctly. SVM uses a technique called the kernel. The main motive of the Kernel is to convert a non-separable problems into separable problems by adding dimensions to it. This makes SVM more powerful, flexible and accurate.

It is capable of solving both classification and regression methods. Svm uses a non-linear kernel trick. In machine learning, kernel tricks are class of algorithms for pattern classification. Non-linear SVM means algorithm calculates it does not have to be a straight line.

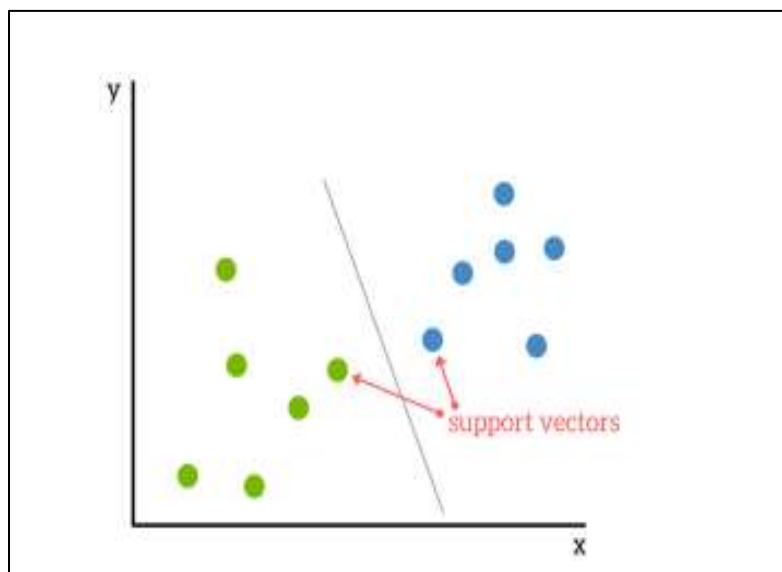


Fig. 5: Support Vector Machine

4) Naive Bayes:

A Naive Bayes classifier may be a probabilistic machine learning model that's used for classification task. The naïve Bayes classifier follows the Bayes theorem[6]

- Bayes Theorem: Bayes Theorem may be a way of finding a probability after we know certain other probabilities.

The Formula used for Bayes Theorem is:

$$P(a|b) = \frac{P(a) P(b|a)}{P(b)}$$

There are different Naïve Bayes:

- Gaussian Naive Bayes: The Gaussian Naive Bayes algorithm used for classification. The predictors take up a continuous value and not discrete.
- Multinomial Naïve Bayes: The multinomial Naïve Bayes algorithm is used for multinomial distributed data i.e data belongs to category of sports, politics, technology etc.
- Bernoulli Naive Bayes: This is a same as Multinomial Naïve Bayes but the Predictors are Boolean Variables. The Parameters that we used to predict the class variable take up only values either yes or no.

5) *Artificial Neural Network (ANN):*

ANNs are composed of artificial neurons which continue to have the biological concept of neurons. They receive the input, combine the input with their internal state and an optimal threshold using an activation function and produce the output. The inputs are the external data and the outputs are the recognition of a particular task. In ANN, the neurons are organized into multiple layers. Neurons of one layer connect to the neurons of the immediately following layers. The layer which receives the initial data is the input layer and the layer that produces the result is the output layer. In-between the input layer and the output layer there may be zero or more hidden layers. Single layer and unlayered networks can also be used. Between two layers, multiple connection patterns are possible, i.e., they can be fully connected, with every neuron in one layer connecting to every neuron in the next layer. Neural networks are divided into many types based on the number of hidden layers.

- Feed- forward neural network:

It is the simple neural network technique. The data moves in only one direction starting from the input node until it reaches the output node. Sometimes hidden layers will not be present.

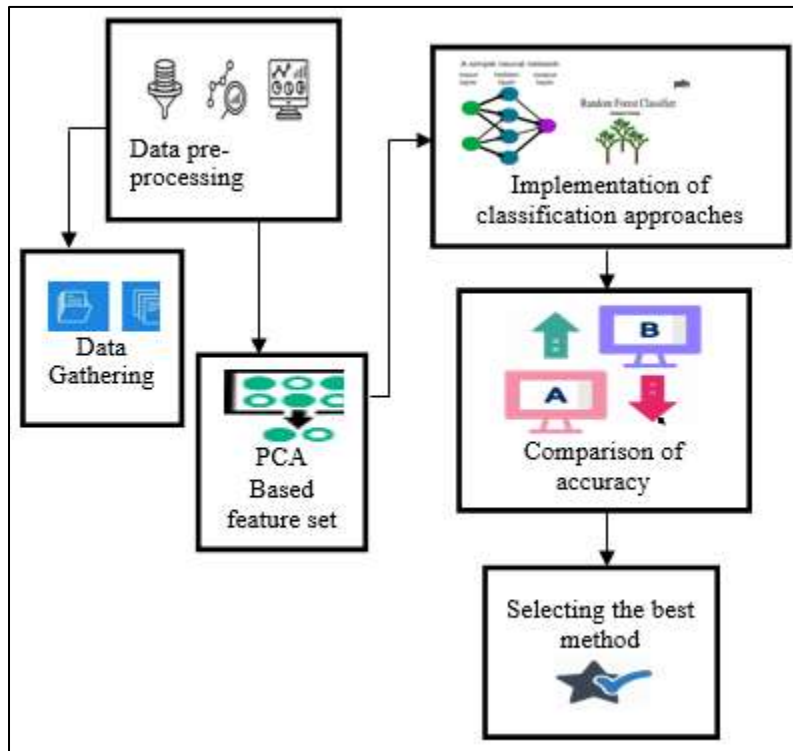
- Radial basis function neural network:

This function considers the distance of any relative points to the center. This function mainly applied to power restoration systems.

- Multilayer perception:

It is used to classify data that not be separated linearly. It has three or more layers. It is one of the types of artificial neural network. It uses nonlinear activation function.

IV. EXPERIMENTAL INVESTIGATIONS PROPOSED ALGORITHM BASIC BLOCK



A. Data Gathering

Data gathering is very important because the quality and quantity of the data you gather will directly affect the level of your prediction model. So we have taken data of different voice recordings of the patient.

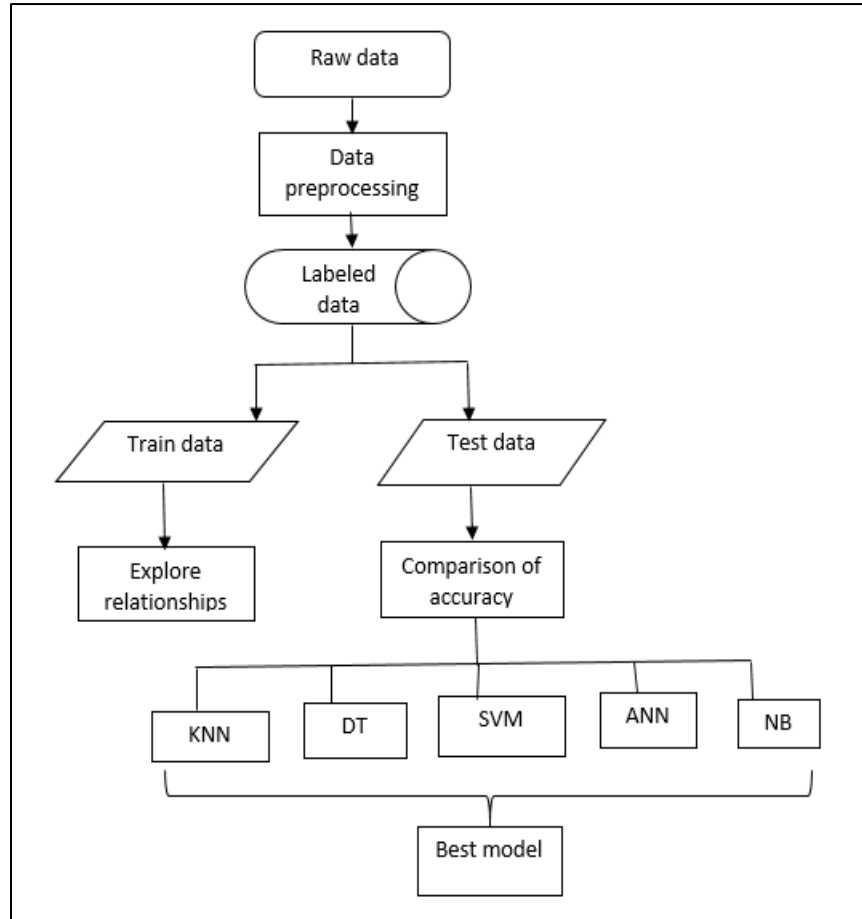
B. Data Pre-Processing

The second step is Data pre-processing. Data pre-processing is a techniques used to convert noisy and raw data in to cleaned data. There are two steps in data pre-processing.

- Selecting of data objects and attributes for the analysis.
- Creating or changing of attributes.
- Selecting the Best Method

The final step is selecting the best method among all algorithms by comparing accuracy, F1score, precision, Recall, etc., values. The algorithm which gets the highest value is the best algorithm among all algorithms.

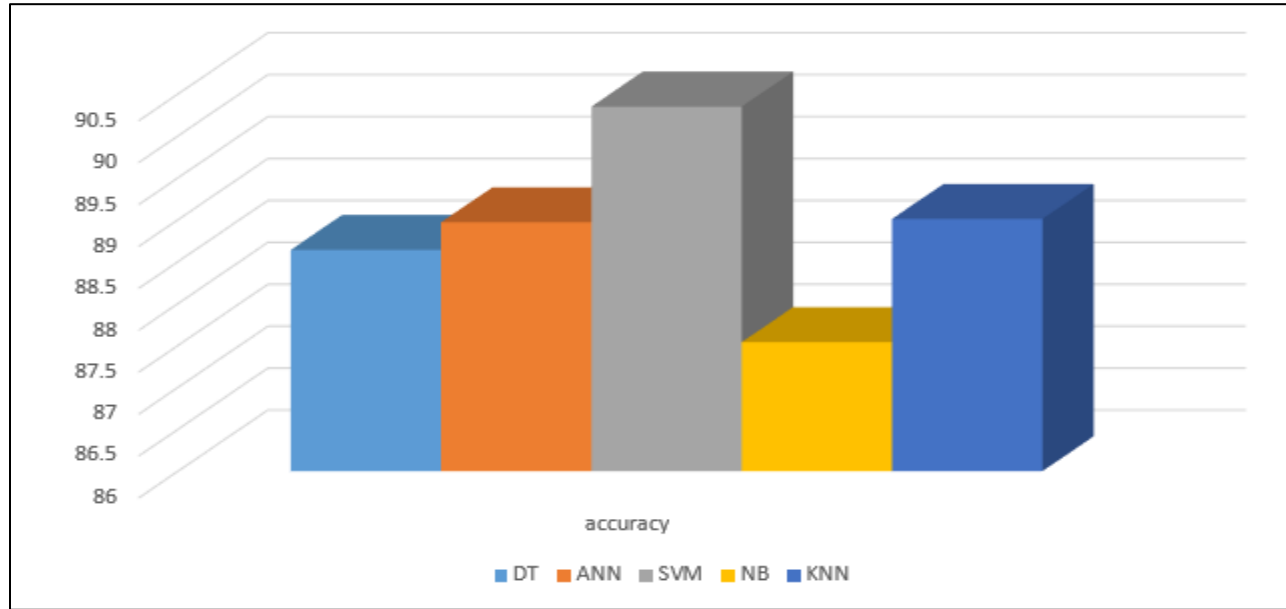
C. Working Procedure:



We have taken the raw data and it is preprocessed to handle missing data.

Now load the data and it is trained and tested. The testing can be done in 5 different methods .Decision tree, K Nearest neighbor, Navie bayes, Support vector machine in all these we are finding the best method with more accuracy.

V. DISCUSSION OF RESULTS



| MODEL | ACCURACY |
|-------|----------|
| DT | 88.64 |
| NB | 87.54 |
| SVM | 90.35 |
| ANN | 88.97 |
| KNN | 89.01 |

We have used different classification algorithms in this project. The accuracy of these different algorithms are Decision Tree(88.64),Naïve Bayes(87.54),Support Vector Machine(90.35),Artificial Neural Networks(88.97),K Nearest Neighbor(89.01).Among all these techniques SVM gives highest accuracy of 90%. So we are using support vector machine for predicting as it has more accuracy.

VI. CONCLUSION

Our project aimed at finding the success rate of companies in promoting their system or products to the people. So that by knowing its success rate they can make any changes needed in their way of promoting. We used different classification technique in this aspect and found Support Vector Machine (SVM) gives the highest accuracy among all the different classification techniques used. So we recommend SVM technique to be used in this kind of predictions regarding classifications problems.

VII. FUTURE SCOPE

This research can be further improvised by implementing various feature selection process on classification techniques. By using this, we can enhance even in different fields not only in the banking sector so that companies can know the reach of their product. We can further improvise this project by using latest machine learning techniques

REFERENCES

- [1] S. Moro, P. Cortez and P. Rita, "A data-driven approach to predict the success of bank telemarketing," ELSEVIER, p. 10, 2004.
- [2] A. Bujari, M. Furini and N. Laina, "On using cashtags to predict companies stock trends," 2017 14th IEEE Annual Consumer Communications & Networking Conference (CCNC), Las Vegas, NV, pp. 25-28, 2017.
- [3] J. A. Frempong and M. Jayabalan, "Predicting customer response to bank direct telemarketing campaign," International Conference on Engineering Technology and Technopreneurship (ICE2T), pp. 1-4, 2017.
- [4] Y. Jiang, "Using Logistic Regression Model to Predict the Success of Bank Telemarketing." International Journal on Data Science and Technology, vol. 4, pp. 35-41, 2018.

- [5] Y. Li, M. Dong and R. Kothari, "Classifiability-based omnivariate decision trees," in IEEE Transactions on Neural Networks, vol. 16, pp. 1547-1560, NOV 2005.
- [6] Y. Ji, S. Yu and Y. Zhang, "A novel Naive Bayes model: Packaged Hidden Naive Bayes," 2011 6th IEEE Joint International Information Technology and Artificial Intelligence Conference, Chongqing, pp. 484-487, 2011.
- [7] G. Guo, H. Wang, D. Bell, Y. Bi and K. Greer, "(2003) KNN Model-Based Approach in Classification," Meersman R., Tari Z., Schmidt D.C. (eds) On The Move to Meaningful Internet Systems 2003: CoopIS, DOA, and ODBASE, vol. 2888, 2003.
- [8] K. kim, C. Lee, S. Jo and S. Cho, "Predicting the success of bank telemarketing using deep convolutional neural network," International Conference of Soft Computing and Pattern Recognition (SoCPaR), pp. 314-317, 2015.
- [9] Z. Liu and L. Bai, "Evaluating the supplier cooperative design ability using a novel support vector machine algorithm," 2008 12th International Conference on Computer Supported Cooperative Work in Design, Xi'an, pp. 986-989, 2008.