***Predictive Marketing Campaign***

*St. Francis Institute of Technology*

|  |  |  |  |
| --- | --- | --- | --- |
| Mr. Sumeet Sharma | Ms. Aaditi Parab | Mr. Avikumar Shah | Mr. Dhruvin Shah |
| *Information Technology* | *Information Technology* | *Information Technology* | *Information Technology* |
| BE IT-B | BE IT-B | BE IT-B | BE IT-B |

*Abstract* — Marketing Campaigns is a series of coordinated activities which are used to communicate with the market to reinforce their positioning and for customer acquisition. Predictive analysis offers a glimpse into how to build or optimize campaigns that derive results. Predictive model is built in R to perform the analysis that is displayed to the user through a web application developed in R Shiny. It determines that whether there will be any deposit generated from the customer based on the campaigns driven through phone calls. The prediction analysis is performed and the accuracy is determined on the model generated by Neural Networks.

Keywords—prediction; marketing campaign; rshiny; predictive modelling; web application, neural networks, machine learning, deep learning, r, rstudio, confusion matrix

# INTRODUCTION

Marketing Campaigns is a series of coordinated activities which are used to communicate with the market to reinforce their positioning and for customer acquisition. These are like advertisements which are used to attract customers on the website and convert them into sales. It becomes very important to determine if there is any loss of sales due to the poor marketing campaigns released. Predictive analysis offers a glimpse into how to build campaigns that derive results. Hence, predictive marketing campaign determines that how should we improve our marketing campaign, who should be our target audience, what should a company offer in the marketing, and when to make the offer of marketing campaign available to customer, i.e. a proper time when there are maximum chances of converting leads to sales. Predictive models make marketing campaigns more effective and profitable by enabling marketers to include only customers that are likely to accept a particular offer or respond to a certain message. Also, we will be using R language an Open source tool to perform the data analysis (here: Predictive). The R tool is easy to use and deploy.

Benefits of predictive marketing campaign

* + - Facilitates result- oriented sales
    - Better conversions
    - Saves time
    - Cost effective
    - Optimizing marketing campaign

After generating marketing campaigns for your business, there can be three cases to evaluate it:

* Best Case:You plan and execute your campaigns to hit specific goals. You don’t always hit them, but you test and improve different elements.
* Neutral Case: The campaigns aren’t the most creative or splashiest, but you’ve hit many of your marketing goals. You don’t test but your response rate is fine. Maybe you don’t know your ROI, but you know generally which campaign works fine.
* Worst Case: You’re low on leads or falling short of your goals and you launch a campaign to fix your problem

Hence, to increase the sales and improve on the business by using marketing campaigns it is necessary to predict the campaigns created for marketing and finding out the amount of sales it is generating. To determine this objective we make use of Predictive Analysis.

# REVIEW OF LITERATURE

## What is R and RStudio?

R is an open-source language and environment for statistical computing and graphical techniques. It is called an environment to distinguish that it is a planned and coherent system rather than a collection of tools.

## One of the most popular IDE for R is RStudio and open-source IDE provided by RStudio Inc [2] (www.rstudio.com). RStudio provides a number of visual tools to assist programming in R (environment viewer, plotting devices, file management, version control, history, etc). The majority of this functionality is provided by packages written by the RStudio team. [2]

## Need for Predictive analytics.

## In the time of traditional advertising, a sales team would resort to calling every lead they could acquire and chasing them around, directionless, in order to attain conversion. If the enquiry fell into their lap, good. Otherwise, they had to prepare a list of potential leads and indulge in frantic cold calling, much to the ire of people.

## How does predictive analytics work?

In a marketing scenario, predicting the efficiency and reachability of ad campaigns based on the past history of data and analysis helps the company to put up or invest in ad campaigns in a more better and useful manner so that it drives better results for sales to the company.

## Building predictive models using algorithms.

A predictive model is made up of a number of predictors, or variables, that are likely to influence future behavior or results. The process is refined and validated as more data becomes available. The model may employ a simple linear equation or a complex artificial intelligence algorithm, mapped out by sophisticated software.

There are several classification models, such as the classical Logistic Regression (LR), decision trees (DTs) and the more recent neural networks (NNs) and support vector machines (SVMs). [1]

## Which algorithm is appropriate to perform predicitve analytics?

There various algorithms that can be used to perform predictive analytics. We can perform regression as well as classification techniques to drive results. Linear or Multiple Regression is used to derive linear relationship between dependent variable and independent variable(s).

Whereas, we use classification to derive result with respect to state of variables.

Regression performs better for extremely small sample size, and also when theory or experience indicates an underlying relationship between dependent and predictor variables. And classification performs small as well as large datasets as compared to regression. Accuracy of the model is high when we choose select classification technique as compared to regression.

## Understanding Neural Networks

Neural networks are machine learning framework that attempts to mimic the learning pattern of natural biological neural networks.

A neural network is made up of interconnected information processing units that do not work in linear manner. A neural network draws its strength from parallel processing of information, which allows it to deal with non-linearity.

## How does a neural network works to perform analysis?

A neural network is a model characterized by an activation function, which is used by interconnected information processing units to transfer input from output.

A neural network works in the following manner:

* The first layer of neural network receives the raw input, processes it and passes the information to the hidden layers.
* The hidden layer passes the information to the last layer which then, produces the output

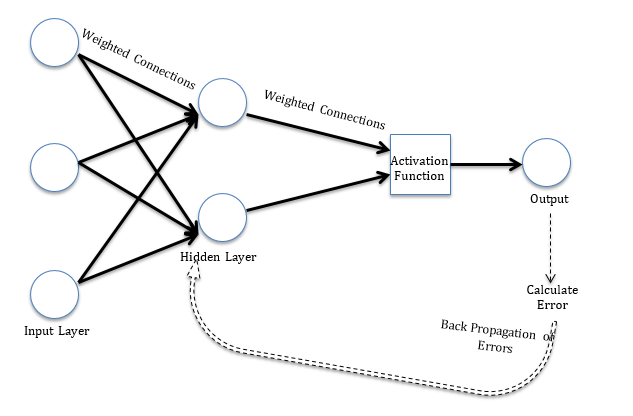


Figure 1: Working model of a neural network

## Advantage of a neural network

* A neural network is adaptive in nature
* It learns from the information provided, i.e. it trains itself from the data, which has a known outcome and optimizes its weight for a better predictions in situations with unknown outcome.

## Understanding decision trees

## A Decision Tree is a robust and transparent Machine Learning model. The tree starts with a single node and then branches out, with a decision being made at every branch point.

## It can be used to predict whether a particular variable would have mattered in the customer’s decision to subscribe or not to the bank’s term deposit.

## How does decision tree regression work?

Decision tree builds regression or classification models in the form of a tree structure. It brakes down a dataset into smaller and smaller subsets while at the same time an associated decision tree is incrementally developed. The final result is a tree with **decision nodes** and **leaf nodes**.

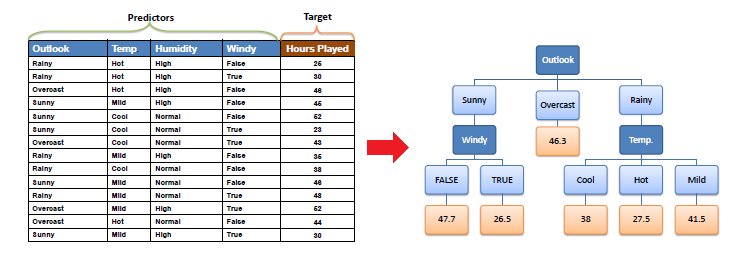


Figure 2: Example of decision tree

# PROPOSED WORK

## Knowing about the dataset.

The dataset we worked on for performing regression analysis is taken from IBM Watson Analytics. It includes total of 11 attributes and 1495 records.

The target / dependent variable here is Sales which depends on the different media of marketing campaigns such as Newspaper, TV and Online. Now let us perform and understand regression analysis on this dataset.

## Working with Linear Regression and its implementation in R.

Regression algorithm is used for prediction analysis and driving better results. It is used to forecast continuous data such as predicting trend for a stock movement given its prices.

Steps in performing analysis using regression algorithm:

• Selecting the Predictors

• Assessing the accuracy of the prediction

• Assessing stability of the model for prediction

• Comparing two different prediction models

1. **Using multiple regressions for prediction analysis.**

Multiple Regression is a technique which is performed considering one attribute or variable and multiple. By multiple regression, we mean models with just one dependent and two or more independent (exploratory) variables. The variable whose value is to be predicted is known as the dependent variable and the ones whose known values are used for prediction are known independent (exploratory) variables.

The line of regression of Y on X is given by **Y=b0+b1\*X1+b2\*X2+…+bk\*Xk**, …………..………(1)

where b0, b1, b2,…, bk are unknown constants known as intercept and slope of the equation. This is used to predict the unknown value of variable Y when value of variables X1, X2, X3,…., Xn known.

1. **Linear Regression modelling in R.**

R is a statistical and analytics tool that is used for performing various statistical calculations and displaying result by performing analysis.

To perform the regression modelling, dataset is first divided into two parts with the ratio 80:20, as training and test dataset. We first build our model on training dataset and then we test our model on the training dataset.

R provides various packages that suits best for our analysis. A built-in function in R called as linear model (lm()) function is used for performing regression analysis. Syntax of lm() function in R can be given as:

model <- lm(formula = dependentVariable ~ indpendentVariable, data = traindata).

Output for the summary of the model is given below:

1. **Understanding the output.**

Our output shows that we have some estimated value of the intercepts which are the coefficients in linear regression. To select that which campaign can drive better results we look upon the threshold (t) value. The one with maximum t-value is OnlineExpenditure.

But the accuracy or the efficiency of using the linear modelling is very low or negligible that is the value of Multiple R-Squared when worked on dataset of one thousand records as compared to say 30-40 records which shows much better results. Hence, we cannot rely on this algorithm. Therefore, we worked on to perform classification and working on new dataset with four thousand records using neural networks in R. Now, the dataset is taken that of Bank Marketing

## Fitting a neural network in R

To create a neural network, we simply begin to add layers of perceptrons together creating a multiplayer of perceptron model of neural networks.

Any layers in between are hidden layers because they don’t really see the feature inputs or outputs.

A bias is added to the perceptron which avoids the issue where all inputs could be equal to 0. And an activation function receives input, multiplies with some weight and then passes them into an activation function.

We keep repeating this process until we reached a maximum number of allowed iterations or an acceptable error rate.

## Understanding the dataset

The dataset we are working on to perform neural network analysis is a **Bank Marketing Campaign** dataset. In this dataset there are 17 attributes and 4251 records.

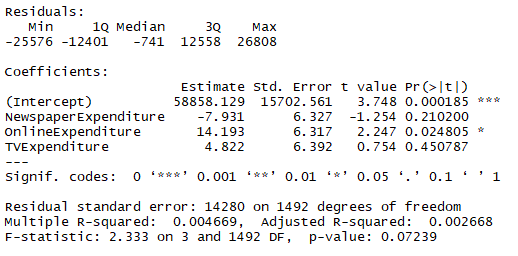
The dataset is based on the campaign driven by bank through phone calls and determining that the customer will perform a deposit into his/her account.

Figure 3: Summary of the linear model

## Implementation of Neural Networks in R

* Data Preprocessing: It is important to normalize the data before training a neural network on it. The neural network may have difficulty converging before the maximum number of iterations allowed if the data set is normalized.
  + **Minmax Normalization:** It transforms the data into a common range, thus removing the scaling effect from all variables. The minmax method retains the original distribution of the variables.
  + **Z-Score Normalization:** Standardization or z-scores is the most commonly used method. It converts all indicators to a common scale with an average of 0 and standard deviation of 1.
* Train and Test Split: The normalized data is split into training and test sets. We divide them in the ratio 80:20. Firstly, we run the neuralnet() function of R on training set and then we test the model built by neuralnet() on test set. Hence, determine the result on the test set.
* Visualizing and understanding the result: Once the training is completed on the train dataset using the neuralnet() function, we plot the model to check how does our neural network is performing. A screenshot of the neural network generated is displayed below:

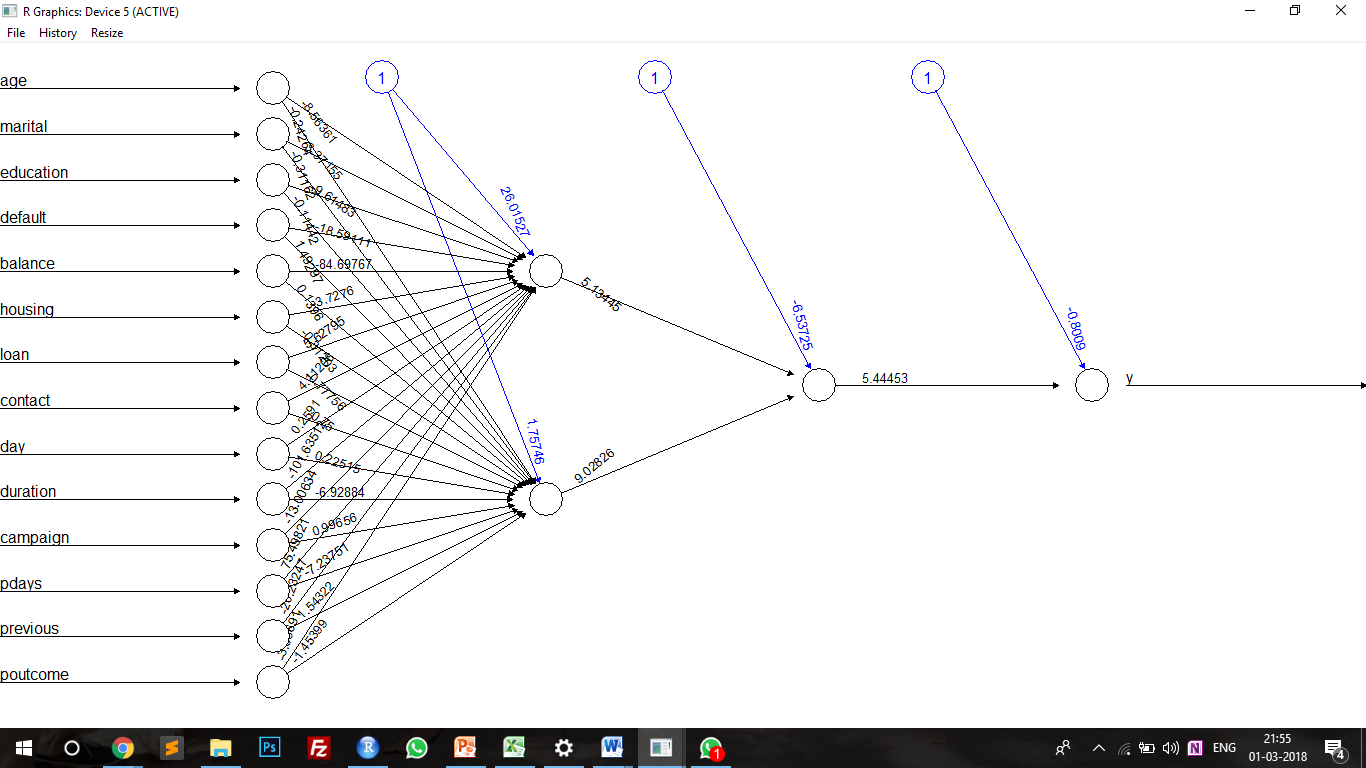
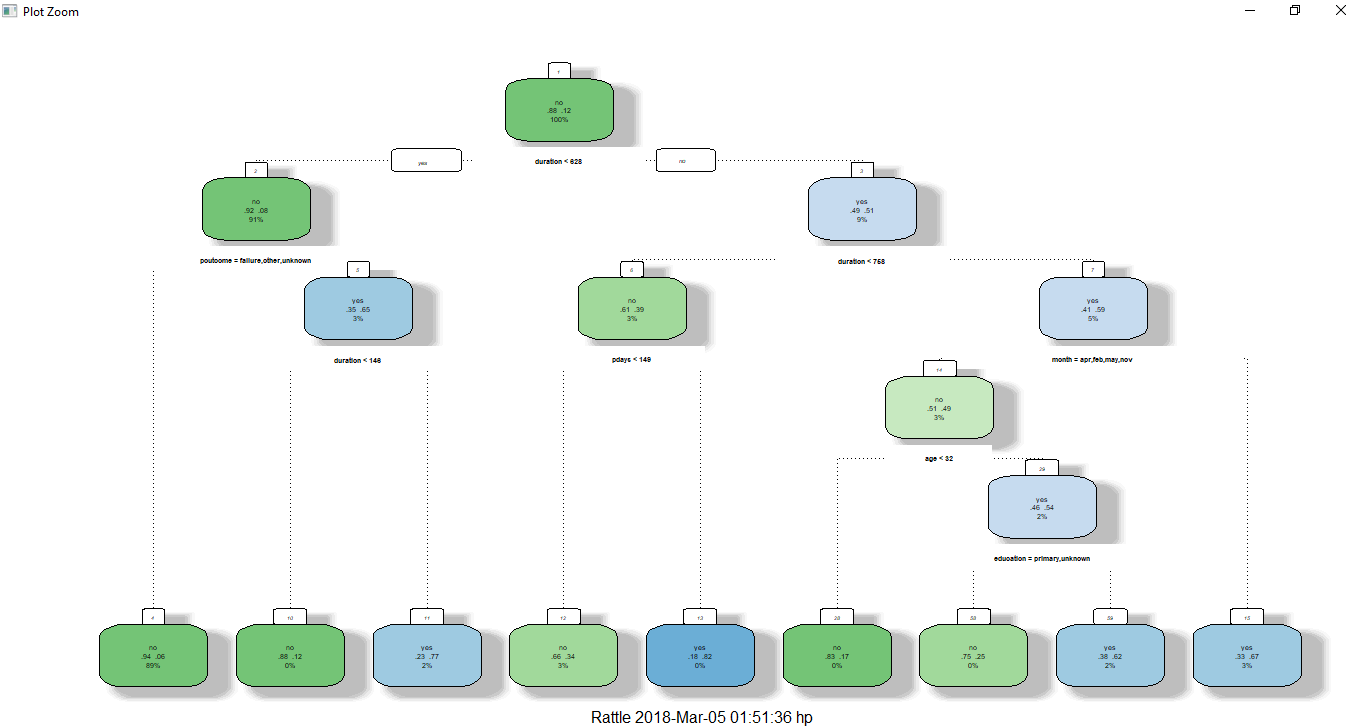


Figure 5: Output of the decision tree

Figure 4: Output of the neural network with hidden layer configuration as (2, 1)

In the above output, the input is passed from the variables to the hidden layers. The hidden layers are arranged in the configuration of 2:1. The above neural network performs classification such that at the end it generates a single output from the inputs given by the hidden layers.

## Interpreting the result after performing testing on the test dataset with respect to the model built.

Table 1: Confusion Matrix generated from the test set

|  |  |  |  |
| --- | --- | --- | --- |
|  |  | Predicted Values | |
|  |  | 0 | 1 |
| Actual values | 0 | 47 | 64 |
| 1 | 26 | 768 |

* The above table shows the analysis in the form of confusion matrix.
* The confusion matrix is such that it is showing total 815 correct predictions and 90 incorrect predictions (values added diagonally).
* Hence, the accuracy is determined by the formula given as: Accuracy of the mode = (correct predictions / total predictions).
* Thus, using neural network, we get an accuracy of 90% whereas an error rate of 10%. Hence, we can say that classification technique for performing predications has a great accuracy and efficiency.

## Working with decision tree

Decision tree is performed in R using the bank marketing dataset. Decision tree is another algorithm performed to check for the accuracy of prediction of data.

In our current dataset, decision tree classifies the dataset into to classes of attribute ‘y’ into *yes* or *no*. These classes determine that will the customer subscribe the bank for the term deposit or no based on the marketing campaigns and other parameters.

## Result of implemented decision tree in R

Output of our decision tree is performed in R. The decision tree function in R itself calculates the information gain and entropy at each level and provides the output for our dataset. The accuracy of our decision tree is 90%. Below is the output image for the decision tree algorithm implemented.

# COMPARATIVE STUDY OF ALGORITHMS

## Study of Algorithms

## We have successfully performed a comparative study of various algorithms such as logistic regression, decision tree, naïve bayes, neural networks, and support vector machines (SVM).

## This comparative study was performed on Orange tool to determine the test score of predictions of each algorithm. Based on the result we have generated a report which contains comparison of these algorithms over the classes ‘yes’ and ‘no’ of bank marketing dataset. A graphical analysis of the precision, recall and accuracy of these algorithms over the training set (75% of dataset) is shown below.

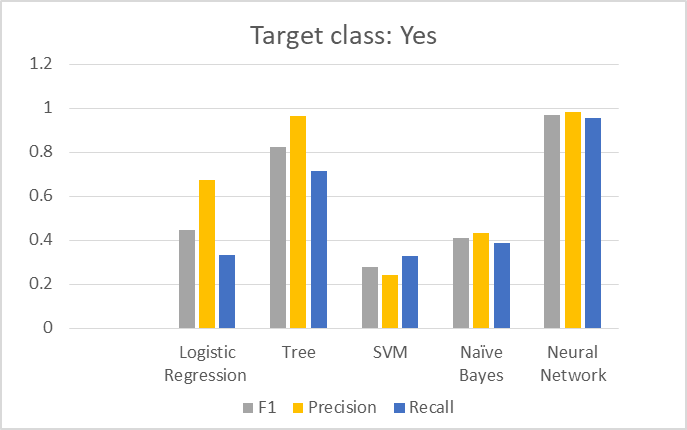


Figure 6: Comparison of algorithms for prediction analysis over target class ‘yes’

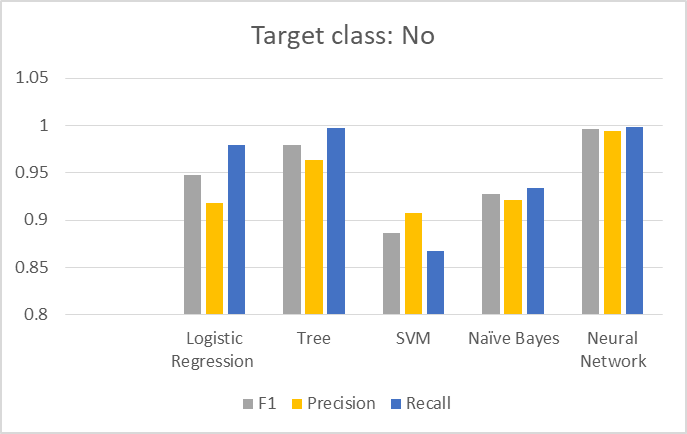


Figure 7: Comparison of algorithms for prediction analysis over target class ‘no’

## Interpreting the result of comparison

From the above graphical analysis about the comparison of various prediction algorithms we can determine that the accuracy of **Neural Networks** is high and good as compared to other algorithms.

We determine the accuracy of these algorithms from the F1 value. The ‘Precision’ attribute determines what was the ratio of the correct predictions over total predictions and ‘Recall’ determines what values had the correct predictions.

Also, from the graphical analysis we can determine that the Neural Network and other algorithms predict that the user will not subscribe for the term deposit. To test these predictions, we will be taking a tuple and selecting random value over each attribute in that tuple. Then, we can determine that whether or not the customer will subscribe for our test tuple.

## Algorithm of selection

Thus, from the above analysis we select **Neural Networks** as our classification algorithm. The results generated from the orange tool were thus helpful in determining and performing predictions over the dataset.

The algorithm will be implemented on the web application developed over RShiny. This will thus ensure that whenever the user takes an input to the system of his / her dataset it will perform the prediction analysis to determine the marketing campaign over the selected target attribute.

# CONCLUSION

We have built a web application for performing predictive analysis on the marketing campaigns set by a company for a product or service. The prediction of the marketing campaigns will help the company to determine where and how to drive the campaigns to get better sales and results. We have performed prediction algorithms to finally conclude which one would provide a better accuracy such as, multiple regression, neural networks and decision tree. Also, we performed a graphical analysis to determine the accuracy with some more algorithms such as support vector machines (SVM), logistic regression, and naïve bayes. Considering all of the above results we have built the web application using neural network algorithm to perform prediction analysis over a bank marketing dataset from UCI repository to determine whether or not a customer would subscribe for term deposit.

# REFERENCES

[1] S. Moro, P. Cortez, P. Rita, “A data-driven approach to predict the success of bank telemarketing”, in Decision Support Systems, Elsevier 62 (2014), pp. 22-31.

[2] Jake Kaupp, “Using R to Collect, Analyze and Visualize Graduate Attribute Data”, Proc. 2016 Canadian Engineering Education Association (CEEA16) Conf, pp. 1-3.