OPEN ACCESS INTERNATIONAL JOURNAL OF SCIENCE & ENGINEERING

SMART MEASUREMENT OF MARKETING ACTIVITIES USING MACHINE LEARNING ALGORITHM

Rohit Babulal Yadav[1] , Deepak Laxman Singh[2] Shubham Chaurasia [3],Dr. Vinayak D. Shinde[4]

Department of Computer Engineering, Shree L.R. Tiwari college of Engineering, Thane.[1]

rohitbyadav786@gmail.com

Department of Computer Engineering, Shree L.R. Tiwari college of Engineering, Thane.[2]

sdeepak247@gmail.com

Department of Computer Engineering, Shree L.R. Tiwari college of Engineering, Thane.[3]

chaurasiashubham48@gmail.com

Associate Professor, Department of Computer Engineering, Shree L.R. Tiwari college of Engineering, Thane.[4]

vdshinde@gmail.com

Abstract: *Advertisement is a costly affair, but very rare serious attempts were there to judge how far it is effective. Measurement is the key to optimizing any process, and marketing campaigns are no exception. When you establish and measure key performance indicators (KPIs) for your marketing campaigns, you can clearly see what works and what does not. You can then direct your marketing resources toward the most effective campaigns to achieve marketing success. Till recently, the effectiveness of advertisement has been measured in relation to sales. Thus now design and develop a system that helps to measure the effectiveness of the marketing activities. To do that use techniques and algorithms like PCA (Principle component Analysis), DA (Dominant Analysis) and machine learning Algorithms like RF(Random Forest Regression), to analyse the data related to marketing activities and build a prediction model that will predict the effectiveness of the activity.*

Keywords-*Marketing Activity, web application, Machine learning.*

-- ∴∴∴ --

I INTRODUCTION

For most businesses, especially the larger ones, marketing is a fluid and constantly changing process that never stops. Many companies run several different types of marketing campaigns at once. The activities include Melas, Bazaars, tagging, paper inserts, in-cinema branding, apartment activities, exchange melas, Finance schemes etc. There is no single metric or measurement technique that can be deployed horizontally across geographies which can provide the right measure of effectiveness for these activities. Having a good measure of whether these activities have yielded desired results will help companies in the rationalization of resources and more importantly, it will help them in planning and implementing need-based activities.

It is understood that marketing is one of the biggest and most important expenditures that any company will make. After all, if you are not doing the right things to add to your customer base, then your company is not going to grow in both size and revenue. That is why effective marketing campaigns are critical to the continued success of any business [1].

The problem is that most companies do not have an unlimited marketing budget. This results in the need for each business to examine the effectiveness of each marketing campaign to see if that type of campaign was successful or should be avoided in the future.

Without having the proper measure of the effectiveness of activities organized by the companies, the company's resources like human power, money, time etc are wasted on

the activities that do not give effective sales or inquiries.

With the help of sufficient and correct data of the activities organized by the company and efficient algorithm, design a model that provides the right measure of effectiveness for these activities.

II LITERATURE REVIEW

Utku kose, Ahmet demir help us to understand trendiest marketing approach. Content marketing is one of the trendiest marketing approaches employed by companies. Because of its connections with especially social media, it is always important to obtain effective content marketing processes in the context of a dynamic, flexible communication environment. So, there is a remarkable research interest in making everything better for content marketing [2].

Jonathon Shlens help us to understand how PCA works is explained. Performing PCA is quite simple in practice.Organize a data set as an $m \times n$ matrix, where m is the number of measurement types and n is the number of trials. Subtract off the mean for each measurement type or row xi. Calculate the SVD or the eigenvectors of the covariance. So it will help to understand what is PCA and how and when it is used [3] [4].

Machine learning, all the more explicitly the field of prescient displaying is basically worried about limiting the blunder of a model or making the most precise forecasts conceivable, to the detriment of logic. In this paper linear and multiple linear regression are compare, Linear Regression endeavours to display the connection between at least two factors by fitting a straight condition to watched information. One variable is viewed as a logical variable, and the other is viewed as a needy variable. Multiple linear regression (MLR), also known simply as multiple regression, is a statistical technique that uses several explanatory variables to predict the outcome of a response variable. The goal of multiple linear regression (MLR) is to model the linear relationship between the explanatory (independent) variables and response (dependent) variable [5].

In this paper, Decision Trees and their extension Random Forests are robust and easy-to-interpret machine learning algorithms for Classification and Regression tasks. Decision Trees and Decision Tree Learning together comprise a simple and fast way of learning a function that maps data x to outputs y, where x can be a mix of categorical and numeric variables and y can be categorical for classification, or numeric for regression. Methods such as SVMs, Logistic Regression and Deep Neural Nets pretty much do the same thing. However despite their power against larger and more complex datasets, they are extremely hard to interpret and neural nets can take many iterations and hyper parameter adjustments before a good result is had. As well, one of the biggest advantages of using Decision Trees and Random Forests is the ease in which is what features or variables contribute to the classification or regression and their relative importance based on their location depth wise in the tree [6].

This article compares the two approaches linear model on the one hand and two versions of random forests on the other hand and finds both striking similarities and differences, some of which can be explained whereas others remain a challenge. The investigation improves understanding of the nature of variable importance in random forests.

III PROPOSED SYSTEM

The system here developed will require a huge amount of data from the company that will use the data. This data will be stored in our system and data pre-processing will be done on the stored data . Since this is real time data there might be some missing data and so on so pre-processing is required. The data can be collected from the surveys taken by the companies or may be generated by analysing the customers activities. The proposed system will analyses the previous marketing data of the company then it will predict the most effective marketing activity that brings better profit to the company than the other activities. So the System is made using mainly two algorithms first is PCA and the second one is Random Forest Regression. Initially, create a dataset that contains a lot of parameters. So to reduce the number of parameters from the data set use dimensionality reduction algorithm , PCA [3] , it returns us the most effective parameters from the dataset. This most effective parameter called as principal components , extract limited principal component as per our requirement Then after reducing the parameters, build a model using Random Forest Regression for each marketing activity. This model predicts the Profit of the marketing activity. The System will take some input parameters like Area, Firm, Investment, Frequency of the marketing activity by the manager or marketing head of the company.
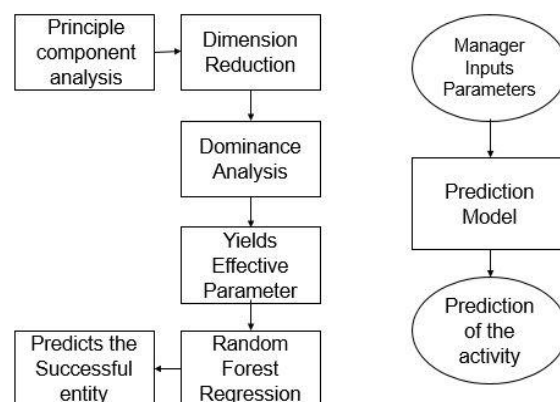

Figure 1: Behavioural Block Diagram Of Prediction System for Marketing Activity

Then it will pass these parameters on all the model for each activity . Then at the output predicted profit of all the activities will be displayed. So the Activity with the highest Profit will be suggested to the manager or marketing head of the company. These results are shown with progress bar so as to visualize easily the activity with their profit with corresponding parameters. So prediction model runs in very less time to predict the profit every time for the new parameters and predicts the results within second.

IV METHODOLOGY
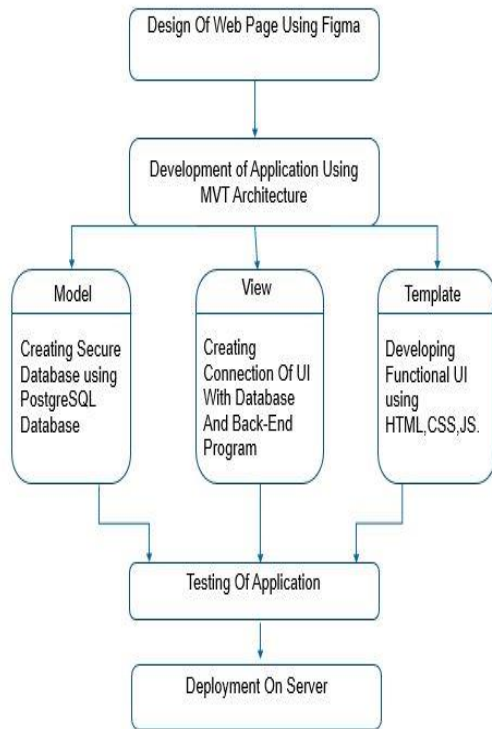
Figure 2: Architecture Of Prediction System for Marketing Activity

The development Of Prediction System for Marketing Activity is divided into stages. Collect data by conducting review of few papers and getting some basic understanding of the technology and algorithms in use. Design web pages using Figma to serve as a guide. Developing The Application Based on the web page designs using MVT architecture framework i.e. Django. The MVT pattern of Django divides the system into three parts: the first model, where the database or all of the data is generated. second component is the view, where the Functionality of the application or the back-end of the application is written, and the last component is the template, which consist of nothing but the UI of the Application [7]. After the implementation, the Application enters the testing phase, which is divided into two parts: black box testing and unit testing. Debug all bugs before

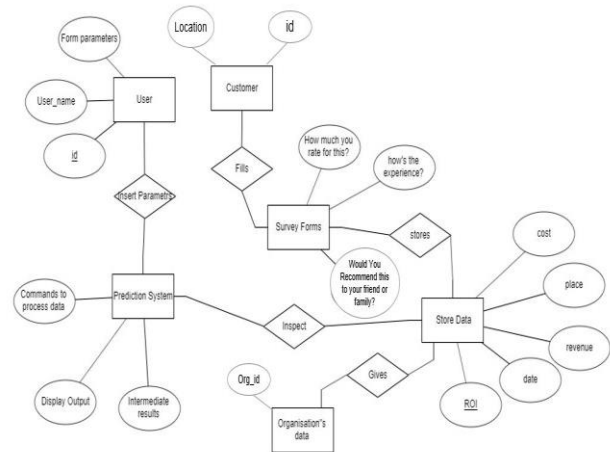deploying the prediction system for marketing activity to the server (Refer figure:2).

Figure 3: ER Diagram Of Prediction System for Marketing Activity

The ER Diagram illustrated in figure 3 gives the visual representation of the model used in this prediction system. This describes how the survey forms are connected to the stored data and those data will be used in the future for the prediction about the best marketing activity [8].

The control flow diagram shows how the user will flow through the system, and how the user's data will flow. The diagram below shows how the user input will be converted to the output, and based on what the user wants to do. The diagram shows the decisions that the system will perform to get the desired output. The intermediate phases are shown in the figure 4.
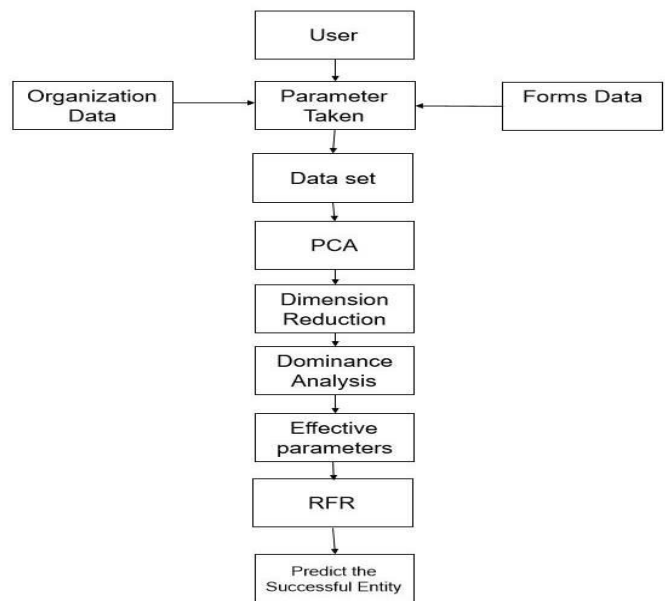

Figure 4: Control Flow Diagram Of Prediction System for Marketing Activity

Prediction is done by using (RFR) Random Forest Regression. RFR is a computational method that optimizes a problem by iteratively trying to improve a candidate solution with regard to a given measure of quality. It solves a problem by having a population of candidate solutions, here dubbed particles, and moving these particles around in the search-space according to simple mathematical formulae over the particle's position and velocity. Each particle's movement is influenced by its local best known position, but is also guided toward the best known positions in the search space, which are updated as better positions are found by other particles. This is expected to move the swarm toward the best solutions [6].

Random forest prediction pseudocode:

To perform prediction using the trained random forest algorithm uses the below pseudocode.

- Takes the **test features** and use the rules of each randomly created decision tree to predict the outcome and stores the predicted outcome (target)

- Calculate the **votes** for each predicted target.

- Consider the **high voted** predicted target as the **final prediction** from the random forest algorithm. [6]

V RESULT & DISCUSSION

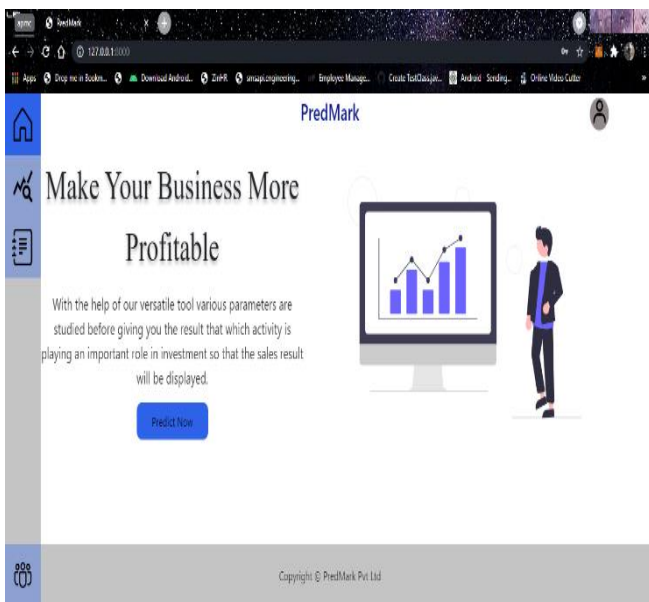Figure 5 shows home page of the prediction system for marketing activity.

Figure 5: Home Page Of Prediction System for Marketing Activity

User or any company can create their account on this web application using **Signup** button given in user section as shown in the below figure.
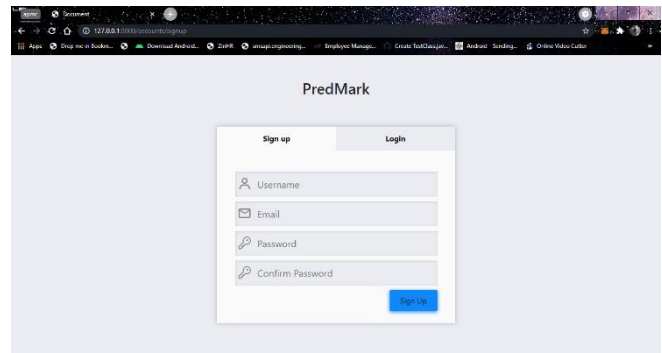

Figure 6: Signup Page Of Prediction System for Marketing Activity

Figure 5.3 shows after account registration user will be able to login to this web application using **Login** button. User login is secured using captcha verification.
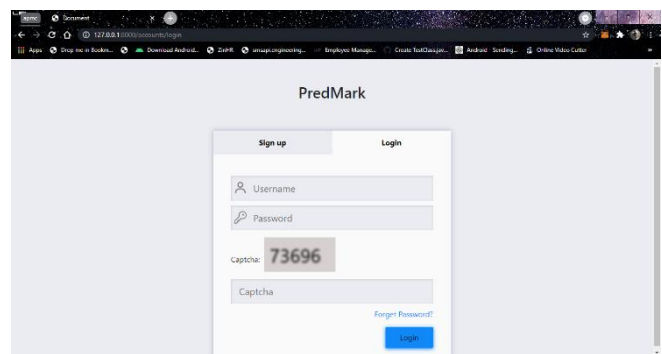

Figure 7: Login Page Of Prediction System for Marketing Activity

Figure 8 shows prediction page where user can predict the best marketing activity for their business. This page is only accessible to only registered user. This Page will ask user to enter few parameters to give the predicted value.
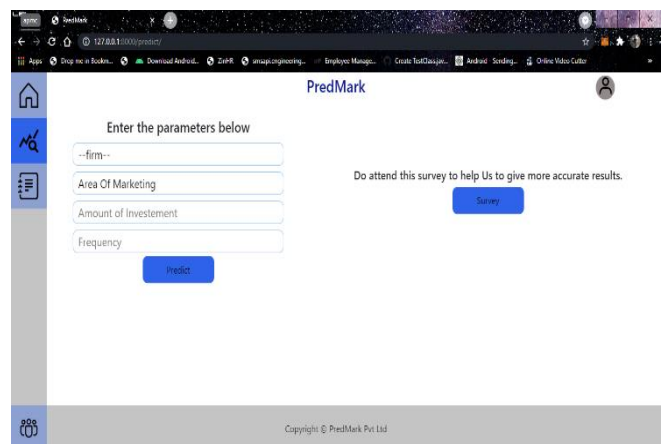

Figure 8: Prediction Page Of Prediction System for Marketing Activity

When clicked on predict button the application will analyse the data using random forest regression and give the predicted value of profit generated by the activity as show in figure 9.
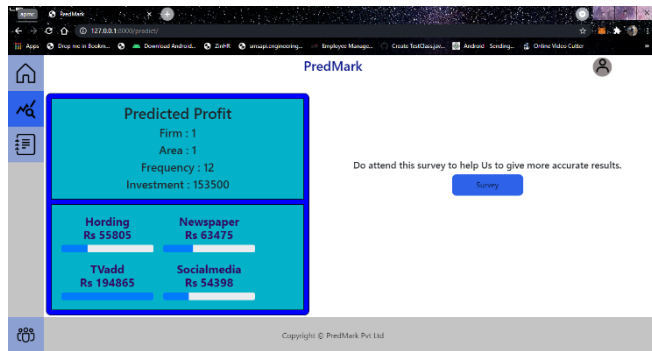
Figure 9: Prediction Page Of Prediction System for Marketing Activity

Apart from prediction, this application will be asking few survey questions shown in following figure 10 to the users or the business owner about their experience, profit etc. Once the survey finished the data will be stored and will be used to give better prediction.
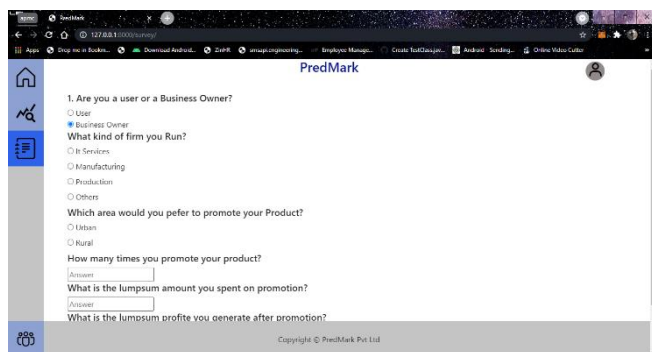

Figure 10:Survey Page Of Prediction System for Marketing Activity

VI CONCLUSION

This paper explains how to find the one marketing activity that will be more beneficial in particular region. Because of this organization can rationalize their resources towards this particular activity. It will help organization in their future planning like giving more budget to that activity which give more profit in particular area. It optimizes the marketing process according to determined success parameters. System improves effectiveness of their current content marketing process and integration of such systems to the objective marketing platforms will enable companies to control and improve their content marketing strategies easily and in a practical way. A proposed system is smart, reliable and efficient. As the traditional systems are more error prone, insecure, and unreliable. In today's world with its huge volume financial data, the traditional system affects towards the organization's budget. Here introducing solution for prediction of profit for each marketing activity and can easily suggest the best marketing activity for the particular region. Finally, you have seen how you can overcome the challenges of organization to invest in right marketing activity.

REFERENCES

[1] T. Goodwin, "Measuring the Effectiveness of Online Marketing," *Market Research Society,* Vols. vol. 41, no. 4, p. 1–6, 1999.

[2] U. Kose, A. Demir, A. Hasan and S. Selcuk, "A software system for intelligent mathematical optimization of content marketing," in *Thr 4th Virtual Multidisciplinary conference*, Usak, Turkey, 2016.

[3] J. Shlens, "A tutorial on principal component analysis.," arXiv preprint arXiv, Mountain View, CA, 2014.

[4] G. K. Uyanık and N. Guler, "A Study on multiple linear regression analysis," in *4th International Conference on New Horizons in Education*, 2013.

[5] Z. M. Y. Z. and L. Z. , "A Flexible Principle Component Analysis Method for Process Monitoring," in *Fourth International Conference on Natural Computation*, 2008.

[6] N. A. C. S. O. and D. K. , "Short term load forecasting using Multiple Linear Regression," in *42nd International Universities Power Engineering Conference*, 2007.

[7] A. Liaw and M. Wiener, "Classification and Regression by randomForest," *Forest,* vol. vol 2/3, pp. 18-22, 2001.

[8] A. Holovaty and J. Kaplan-Moss, The Definitive Guide to Django: Web Development Done Right, NY: Apress, 2009.

[9] S. Shrinivasan and M. H. Dominique, "Marketing and firm value: Metrics, methods, findings, and future directions," *Journal of Marketing Research,* pp. 293-312, 2009.

[10] A. Ortiz, "Web development with python and django," in *SIGCSE'12 :Proceedings of the 43rd ACM technical symposium on Computer Science Education*, 2012.

[11] T. H. Park, "An Analysis of HTML and CSS Syntax Errors in a Web Development Course," in *ACM Transactions on Computing Education*, 2015.

[12] N. Kumari, "SQL Server Query Optimization Techniques - Tips for Writing Efficient and Faster Queries," in *International Journal of Scientific and Research Publications*, 2012.