

B2B SALES AND PRIDITION FOR RESTAURANTS

Prof. Mahajan Anita S.¹, More Sneha R.², Shinde Amruta B³., Katvate Shubhangi A.⁴

^{1,2,3}Student, Computer Engineering, G. H .Raisoni Institute of Engineering and Technology Maharashtra, India

⁴Assistant Professor, Dept of Computer Engineering G. H .Raisoni Institute of Engineering and Technology Maharashtra, India

Abstract - For efficient and economical operation, restaurant owners need to estimate the number of future customers. In this paper, we propose an approach to predict how many future visitors will go to a restaurant using supervised learning. The big data includes restaurant information, historical visits and the overall details of restaurants .

The results show the effectiveness of our approach. Popular machine learning algorithms like Decision Tree and Random Forest were applied over a dataset records. The results have proved that the Decision Tree Classifier is more effective than Random Forest

Key Words: *Machine learning ; Linear regression, Random forest algorithm*

1. INTRODUCTION

Operating a profitable restaurant isn't easy. In addition to managing day-to-day operations, restaurant owners and managers must also constantly evaluate ways to increase efficiency, grow sales and reduce costs all while managing risk and compliance. A restaurant sales inventory prediction enables you to know the count of dish, what inventory to order for the next month, and how many customers should be visit next day or in two weeks. Restaurant management is a procedure to monitor the raw material and supplies you have for schedule in real time, and it also helps you make more food and supply orders for proper management. Effective inventory management is essential to our day-to-day operations and long term Goals.

Based on historical data of business, predictive models of machine learning can estimate future moves of customers . the number of future visitors can help restaurant owners make the best operations to maximise the particular dish count. In addition, restaurant owners can also schedule a suitable number of dishes count to dataset. Previous work such as focused only on visitors' revisit or restaurants review. Hence, it is necessary to develop a new tool that can predict the total number of future visitors or extract raw material from vendors. Our approach is to gather big data that include restaurant information and the historical visits for restaurants. With these features, our approach generates predictions by performing regression using a Linear regression, Random Forest. Restaurants can easily collect historical data without any complex computing . Several Machine Learning algorithms are applied over the dataset for retrieving results. The algorithms used for classifying the

overall class in Random Forest Algorithm and Decision Tree. Based on the outcomes generated by the algorithms applied the results are visualized for better understanding and decision making. We here check the algorithm that performs better on our dataset. It compares the machine learning algorithms by comparing the class of the restaurant in the dataset.

2. RELATED WORK

The historical data which was considered to be scrap when the documents were being preserved on the hard copies or register books has now tend to become the most important aspect for the businesses to take the important decisions in the stock/inventory management. So the historical data generates the pattern of sales on weekly or daily basis and can be fed to the machine learning algorithms to learn from the patterns and predict the further sales of specific cuisine. Although various techniques are proposed for many areas to predicting the number of future visitors in service industry, little is known for restaurant owners to forecast the number of future visitors using big data focused only on visitors' revisit. Some inputs in the system were automatically generated and easily readable by machine. To clean the data and cut out inessential information prior to modeling, the team first filtered out all entries created then the system was formally launched for the company¹. Variables with a high percentage of null values were then excluded to ensure a sufficient sample size. The remaining variables were further screened based on potential importance determined by conversations between the team and key company stakeholders.

3. Proposed Methodology

Based on big data, many machine learning algorithms for regression can find correlations between factors. Support Linear Regression is a widely used method for regression. It can also work for classification.

Machine learning algorithms:

Linear regression:

Is a linear approach for the relationship between a scalar response (or dependent variable) and one or more explanatory (or independent variables). It determine the linear relationship between a dependent variable and one or more independent variables. In simple linear regression a single independent variable is used to determine the value of

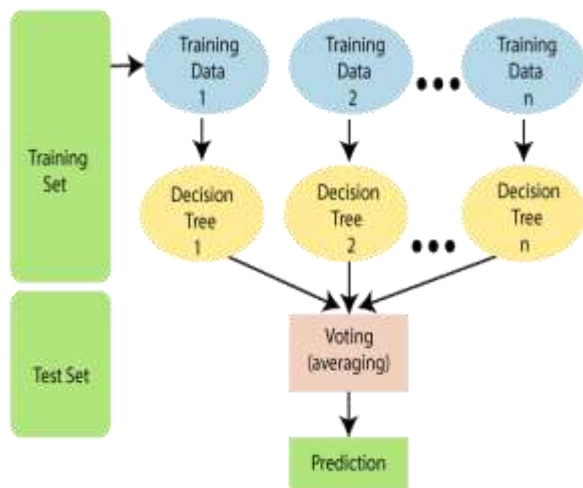
a dependent variable. Regression models are used to determine a continuous value. Determine the prices pizza of a given the features of pizza like size, price, raw material etc is one of the common examples of Regression.

Random Forest algorithm:

Random Forest algorithm is also known as supervised classification algorithm. from its name of random forest algorithm it create a forest by many way and make it random. There is a relationship between the number of trees in the forest and the results from the trees. If there are larger the number of trees then it is possible to get more accurate result. One thing to observe is that creating the forest is not the like as making the decision with information gain .

Random forests are a way of calculating multiple decision trees and tests on that parts of the same training data , with the goal of decreasing the variance. Random Forests (RF) is a decision-tree-based method also for both regression and classification. As decision trees are non-linear by nature, RF can work for either linear or nonlinear data without any prior knowledge about linearity. Relying on a number of decision trees, it uses the mean prediction to generate the final prediction.

The main difference between Random Forest algorithm and the decision tree algorithm is that in Random forest algorithm it divides tree in different root nodes randomly.



Why random forest is efficient ?

Random forest can be used for regression analysis. There is a grouping of different regression trees and it is used for nonlinear multiple regression.

It firstly splits the tree in different sub trees using the features then it calculates based on maximum number of votes receives from the more different tree. It recognize the count of maximum number of votes and then that tree is considered for the final output of that particular random

forest. We count the approximation and the count of customer can that visit next day or not to restaurant.

2.1 Different between linear regression and Random forest algorithm

Linear Regression

- Linear learning algorithm is based on supervised learning.
- This method is used for forecasting and finding out cause and effect relationship between variables.
- It determine the linear relationship between a dependent variable and one or more independent variables.

Random Forest

- Random forest algorithm is a supervised classification method.
- Random forest algorithm can use for classification and regression problem.
- Decision tree concept is a rule based system used for high accuracy result.

Abbreviations and Acronyms

B2B — Business-to-Business

RFA-Random Forest Algorithm

EOD -End of Day

LR- Linear Regression

P&L - Profit and loss

SFDC - Salesforce.com

Equations

Linear regression is a linear approach to modeling the relationship between a scalar response (or dependent variable) and one or more explanatory (or independent variables).

The Linear Regression Equation

The form of equation is $Y = a + bX$,

Y - dependent variable,

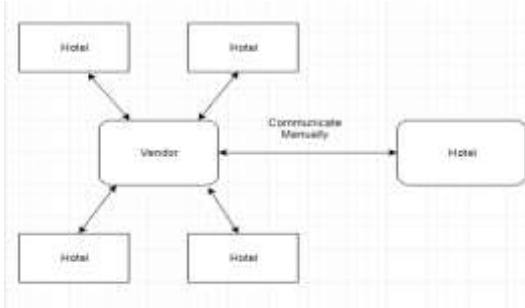
X - independent variable

b - slope of the line and

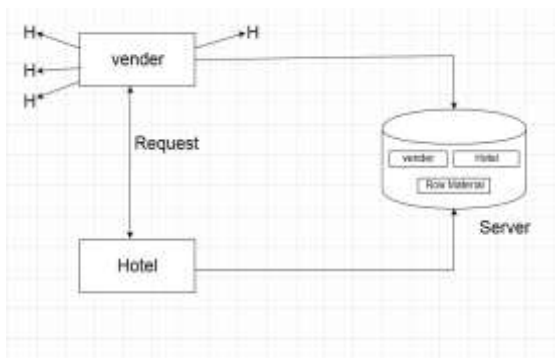
a - y-intercept.

Existing system

B2B sales and inventory prediction management software is needed to meet the demands of this task.



System Architecture



Hotel Manager add restaurants sales data to dataset.

Algorithm performs operations and gives dish count to Hotel Manager.

Vendor / Retailer add their raw material data to dataset.

Hotel Manager Receives dish count and request to Vendor for raw material.

According to request of raw material from particular hotel vendor provides that much material to Hotel manager

3. CONCLUSIONS

We have proposed the model using machine Learning techniques linear regression and random forest algorithm which makes sure that the products to be purchased and from wherever to be purchased will be in sync with the

The authors can acknowledge any person/authorities in this section. This is not mandatory. predicted number of customers and the historical dishes being preferred as per customers. So it can predict raw ingredient for selected food. it can use for Easy and Efficiently Collection of raw materials so it Reduces the wastage of food. For the purpose of Better planning and optimized performance we are proposed this system.

REFERENCES

- [1] Implicit (Sales Cloud by Salesforce.com). [Online]. Available: <https://www.salesforce.com/blog/2014/08/infographic-7-powerfulpredictors-closed-won-opportunity-gp.html>
- [2] Insight Squared.[Online]. Available: <https://www.insightsquared.com/features/sales-forecasting/>
- [3] J. Yan, C. Zhang, H. Zha, et al, "On Machine Learning towards Predictive Sales Pipeline Analytics." Proceedings of the Twenty-Ninth AAAI Conference on Artificial Intelligence, pp. 1945-1951, 2015. [Online]. Available: <https://www.aaai.org/ocs/index.php/AAAI/AAAI15/paper/download/9444/9488> [Accessed: Mon. 24 Sept. 2018]. \
- [4] M. Bohaneca, M.K. Borstnarb, M. Robnik-Sikonja, "Integration of machine learning insights into organizational learning: A case of B2B sales forecasting." 28th Bled eConference, June 7-10, 2015. [Online]. Available: [https://domino.fov.unimib.si/proceedings.nsf/Proceedings/B12ECF2381AB59EEC1257E5B004B39B7/\\$File/2_Bohanec.pdf](https://domino.fov.unimib.si/proceedings.nsf/Proceedings/B12ECF2381AB59EEC1257E5B004B39B7/$File/2_Bohanec.pdf) [Accessed: Tue. 25 Sept. 2018].
- [5] M. Bohaneca, M.K. Borstnarb, M Robnik-Sikonja, "Explaining machine learning models in sales predictions." Expert Systems with Applications, no. 71, pp. 416-428, 2017. [Online]. Available: <http://lkm.fri.uni-lj.si/rmarko/papers/Bohanec17-ESwA-preprint.pdf> [Accessed: Tue. 25 Sept. 2018].
- [6] Crooks, G. E. 2009. "Logistic approximation to the logistic-normal integral". Technical report, Lawrence Berkeley National Laboratory.
- [7] Simatupang Togar M. and Sridharan Ramaswami, An integrative framework for supply chain collaboration, The International Journal of Logistics Management, 2005, 16(2), 257 – 274.
- [8] Bowersox, D.J., Closs, D.J. and Keller, S.B., How supply chain competency leads to business success, Supply Chain Management Review, 2000, 4(4), 70-8.
- [9] L. Breiman, J. H. Friedman, R. A. Olshen, and C. J. Stone, Classification and Regression.
- [10] A. Jain, R. Dubes, Algorithms for Clustering Data, Prentice-Hall, Englewood Clis, NJ, 1988.