

## Potential Customer Prediction

Suyash S. Tambe<sup>1</sup>, Vivek Ambetkar<sup>2</sup> and Prathamesh Ambovkar<sup>3</sup>

<sup>1</sup>Student, Department of Information Technology, Padmabhushan Vasantdada Patil College of Engineering and Visual Arts, Mumbai, INDIA

<sup>2</sup>Student, Department of Information Technology, Padmabhushan Vasantdada Patil College of Engineering and Visual Arts, Mumbai, INDIA

<sup>3</sup>Student, Department of Information Technology, Padmabhushan Vasantdada Patil College of Engineering and Visual Arts, Mumbai, INDIA

<sup>1</sup>Corresponding Author: [suyashtambe444@gmail.com](mailto:suyashtambe444@gmail.com)

Our project "Potential Customer Prediction" helps the organization to find the customers who are potential and customers which are non-potential buyers and so it helps them to enhance their services for better customer experience.

### ABSTRACT

Major industries today are dealing with large amount of data even small shops are no oblivion for the huge pile of customer data. To gain profit in competitive marketing is imperative that the useful information shall be extracted out of this data. The processing of this huge pile of data becomes monotonous task and the different types of software and algorithms are developed to process and acquire result out of this data. This project deals with same kind of problem of dealing with data. Taking customer purchase history as an input our system using "Apriori Algorithm" classifies these customers as potential and non-potential customers. Customers who are potential are more likely to buy more from the store and help enhance business.

**Keywords--** Customer, Prediction, Machine Learning, Apriori Algorithm

## I. INTRODUCTION

Analyzing customer behavior has been the tryst of development of business relations since long time. Bring a new customer to the business is cost-intensive because it involves various marketing strategies without having prior knowledge about the customer. But when the business approach to its existing customers, they can easily apply specific marketing strategies based on previous customer information.

Data accumulation is highly increased in recent years which leads to a great interest in the field of machine learning (ML). The computer has proven its usefulness in predicting and pattern findings by learning the data.

Grocery superstore business has been accumulating huge data from customers. Due to the digital revolution, most of the grocery superstore is using computer software to manage sales and customer data. Exploring and analyzing this data can provide new insights which can lead to profit acceleration

According to The Pareto Principle twenty percent of customers create eighty percent of the profit. These twenty percent customers are considered as the potential customer. The purpose of this research is to explore how machine learning technique can be used to identify potential customers in a business-to-customer (B2C) sales context, using customer sales data. There are two goals in this research. (i) Implementing machine learning models in a real-life scenario to find the limitations and possibilities. (ii) Application of machine learning in business data to predicts potential sales

## II. OBJECTIVES

The core objectives of project are:

1. Enhance business customer Relations
2. Increase knowhow about customer.
3. Find loyalty based customers.
4. Better management of business according to customer requirements.

Realizing above objectives will help any organization to enhance its profits and make better use of their data to promote know-how about their customers by which organization can form policies which benefit their business customer relations.

## III. MOTIVATION

While reading a book "Power of habit - by Charles Duhig" I came across an interesting information that how in late 90's people used to analyze the business data very rigorously to come up with certain pattern that their customer uses while making a purchase and capitalize on that habit of customer. Later many algorithms were used for doing this job. The sole purpose of this operation was to capitalize on psychological aspect of customer to make them buy more. We in our project drew motivation

from this idea of using customer data to know which of them are more likely to buy more and what changes can we make to business model so we can capitalize more on our customers data. the simple idea behind our project is to classify the customers in potential and non potential category so the organization can focus on potential customers to enhance the business profits.

#### IV. WORKING

Machine learning has many successful footprints in various domains. We used the machine learning approach to analyze a customers purchase behavior. A supervised machine learning technique is used in our study to classify potential customer. At first, we collect the data from sales software and pre-processed it. The quantity of item sold by the unit is considered as features. We used to engineer the dataset to label the data. Using this labeled data, we had trained our supervised model for classification. After assessing the purchase pattern of a customer the system generates a potentiality score between 0 and 1, where 1 stands for higher potentiality and 0 of no potentiality. Classification related computational measures are learned from the methods. This system involves in a large amount of data pre-processing. Complete knowledge of the data required to determine the present features. To increase the accuracy, features were engineered.

#### V. METHODS

In this section, we have represented detailed explanations of data profile along with data pre-processing and labeling. We have also discussed the methods and tools used to apply machine learning algorithms in our problem domain to gain the insight knowledge and achieve the goal.

##### Apriori Algorithm

Apriori is one of the algorithm that is used for frequent pattern mining. Frequent pattern mining methods calculate the probabilities of occurrence of each patterns and filter them with a threshold probability value. For instance, Let's suppose that we have entered a e-commerce web site and start to shopping. After we took the products to our basket, we might seen such notifications or sidebar information that says "People who bought these products also bought these". Most probably, It is the result of one of the Frequent Pattern Mining algorithm. These algorithms scans the previous transactions of customers and finds that buying one or more product how affects the buying other products. Basically, It generates probabilities of products that we might purchase according to our basket and list them through a threshold value. It finds the products mostly bought at the same time. It creates a pattern from these products and this pattern called frequent pattern. Frequent

Pattern Mining methods are used in order to find the probabilities of occurrence of these patterns.

##### Listing 1: Apriori Algorithm

**Step 1:** While the number of customer attribute items in the set is greater than 0:

**Step 2:** Create a list of customer attribute itemsets of length k

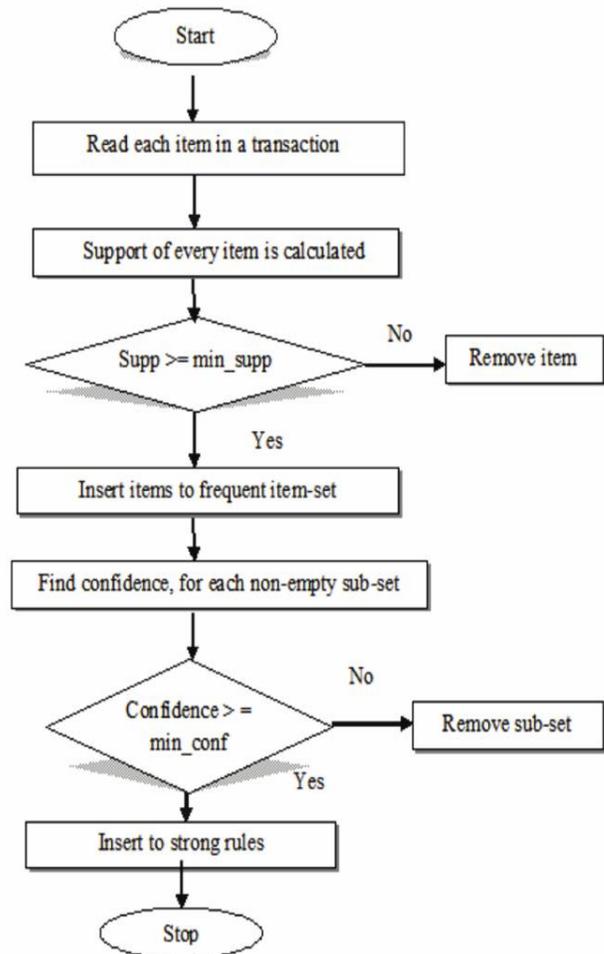
**Step 3:** Scan the dataset to see if each customer attribute itemset is frequent

**Step 4:** Keep frequent attribute itemsets to create itemsets of length k+1

The step 3 (Scan dataset) in Algorithm 1 which is an iterative process is further describes as;

- For each transaction in tran the dataset:
- For each customer attribute itemset, cus:
- Check to see if pat is a subset of tran
- If true, increment the count of cus
- For each customer attribute itemset:
- If the support meets the minimum, keep this item
- Return list of frequent attribute itemsets

##### Flowchart



**Hardware Requirements**

- Desktop or Laptop
- Hard drive 500gb

**Software Requirements**

- Anaconda Navigator
- Intel i3 processor
- Windows 7, 8, 10.

**VI. CONCLUSION**

Better understanding of customer behavior and targeting them with right ads can be achieved.

This system shall provide organizations with benefit to capitalize on cues provided by customers. The potential customer prediction will provide the comprehensive distinction between loyalty based customer and profit based customer. These systems will help to find the potential and non-potential customers. These can target the customer and better analysis of purchasing data. The algorithms can analyse hundreds and thousands of parameters and attributes, millions of data points or log entries, and filter out the four or five important ones for understanding, for example, customer behavior.

**REFERENCES**

- [1] Tan, Y., Xu, X., & Liu, Y. (2016). Improved recurrent neural networks for sessionbased recommendations. In: *Proceedings of the 1st Workshop on Deep Learning for Recommender Systems*, pp. 17–22.
- [2] Suchacka, G. (2016). Analysis of aggregated bot and human traffic on e-commerce sites. In: *Proceedings of IEEE FedCSIS14*, 2, pp. 1123–1130.
- [3] J.E.T. Akinsola. (2017). Supervised machine learning algorithms: Classification and comparison. *International Journal of Computer Trends and Technology*, 48(3), 128-138.
- [4] Mohssen Mohammed, Muhammad Khan, & Eihab Bashier. (2016). *Machine learning: Algorithms and applications*. CRC Press.
- [5] Salehi, F., Abdollahbeigi, B., Langroudi, A. C., & Salehi, F. (2012). The impact of website information convenience on e-commerce success of companies. *Procedia Social and Behavioral Sciences*, 57, 381–387.
- [6] P. Kisioglu & I.Y. Topcu. (2011). Applying bayesian belief network approach to customer churn analysis: A case study on the telecom industry of Turkey. *Expert Systems with Applications*, 38(6), 7151-7157.