Prediction of Best Brand in E-Commerce

Narayanan V[#1],Hariharsuthan K[#2],Veeralakshmi P[#3]

[#1],Student, Department of Information Technology, Prince Shri Venketeshwara Padmavathy Engineering College, Ponmar, Chennai

[#2], Student, Department of Information Technology, Prince Shri Venketeshwara Padmavathy Engineering College, Ponmar, Chennai

[#3]Head Of The Department, Department of Information Technology, Prince Shri Venketeshwara Padmavathy Engineering College, Ponmar, Chennai

Abstract: **Digital marketing is considered the preferred method comparing to traditional marketing. It is useful to both practitioners and academics of social media marketing and purchase intention. The research provides some initial insights into consumer perspectives of social media ads and online purchase behavior. Business, academician, researchers all are share their advertisements, information on internet so that they can be connected with people fast and easily to survey on searchable mobile brand websites. To prevent this problem, web scraping helps collect these unstructured data and store it in a structured form. The aim is to investigate given dataset using machine learning based techniques for brand name forecasting by regression and prediction results in best accuracy. The analysis of dataset by random forest algorithm is to capture several information's like, variable identification, uni-variate analysis, bi-variate and multi-variate analysis, missing value treatments and analyze the data validation, data cleaning/preparing and data visualization will be done on the entire given dataset. Our analysis provides a comprehensive guide to sensitivity analysis of model parameters with regard to performance in prediction of sales ratings with mobile features by finding accuracy calculation. To increase the sales in the Ecommerce based upon the customer requirement and current trend, we also present a fast machine learning algorithm for analyzing purpose. Here, K-Nearest Neighbour & Random Forest algorithm is used for classification & recommending the brand & MLP Regression is used for regression purpose.**

I. INTRODUCTION

Machine learning is to predict the future from past data. Machine learning (ML) is a type of artificial intelligence (AI) that provides computers with the ability to learn without being explicitly programmed. Machine learning focuses on the development of Computer Programs that can change when exposed to new data and the basics of Machine Learning, implementation of a simple machine learning algorithm using php. Process of training and prediction involves use of specialized algorithms. This algorithm has to figure out the clustering of the input data. Finally, Reinforcement learning dynamically interacts with its environment and it receives positive or negative feedback to improve its performance. In machine learning and statistics, classification is a supervised learning approach in which the computer program learns from the data input given to it and then uses this learning to classify new observation. This data set may simply be bi-class (like identifying whether the person is male or female or that

the mail is spam or non-spam) or it may be multi-class too. Some examples of classification problems are: speech recognition, handwriting recognition, bio metric identification, document classification etc.

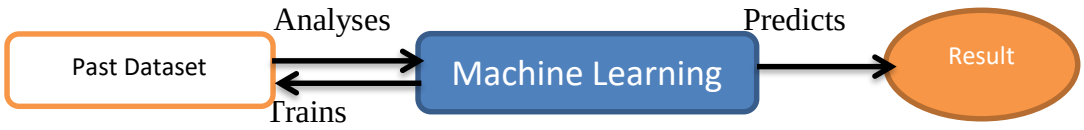

Fig. Process of Machine learning

Techniques of Supervised Machine Learning algorithms include **logistic regression**, **multi-class classification**, **Decision Trees** and **support vector machines etc**. Supervised learning requires that the data used to train the algorithm is already labeled with correct answers. Supervised learning problems can be further grouped into **Classification** problems.

II. RELATED WORKS

[1] Influence of Cognitive Resource Limitation on Consumer Purchasing Decision: An Event-related Potentials Perspective. It is to study the neural influence of conflicts between product appearance and performance on consumer decision when consumer cognitive are limited.

[2] Determining the Effects of Marketing Mix on Customers' Purchase Decision Using the Grey Model GM (O, N) - Case Study of the Western style Coffeehouse Chains in Vietnam. in order to meet coffee drinker demand as well as expand the market share, suppliers of the coffeehouse chain should be taken more consideration on promotions activities and the taste of coffee beverages. Furthermore, products of coffee beverages with good taste and reasonable price can make consumers have a high-level satisfaction and confidence to purchase.

[3] Online Apparel Shopping Behavior. The study revealed that consumers tended to be in need of the new manifestations of online information, such as the comparison information and the e-word-of-mouth information. Consumers comparatively rely on experience or evaluations shared by other consumers' real purchases, instead of the official information provided by the retailers or obtained from the brick-and-mortar stores.

[4] Consumers' intention to purchase remanufactured electronic products: an empirical study in China. Remanufacturing has been characterized as a critical component of the circular economy and an engine of the closed loop supply chain process. However, few researches focus on consumers' intention to purchase remanufactured products.

[5] How Perceived Factors of Review Contents Influence Consumers' Purchase Decision. All factors of perceived benefit have a positive influence on purchase decision, which proves hypothesis all factors of perceived risk have a negative influence on purchase decision, which proves hypothesis the decision tree model shows that perceived risk has a more all-round influence on purchase decision than perceived benefit, which proves hypothesis.

III. PROBLEM DEFINATION

- **Exploration data analysis of variable identification**
  - Loading the given dataset
  - Analyze the general properties
  - Find duplicate and missing values
  - Checking unique and count values
- **Uni-variate data analysis**
  - Rename, add data and drop the data
  - To specify data type
- **Method of Outlier detection with feature engineering**
  - Pre-processing the given dataset
  - Splitting the test and training dataset

3.1 SYSTEM ARCHITECTURE

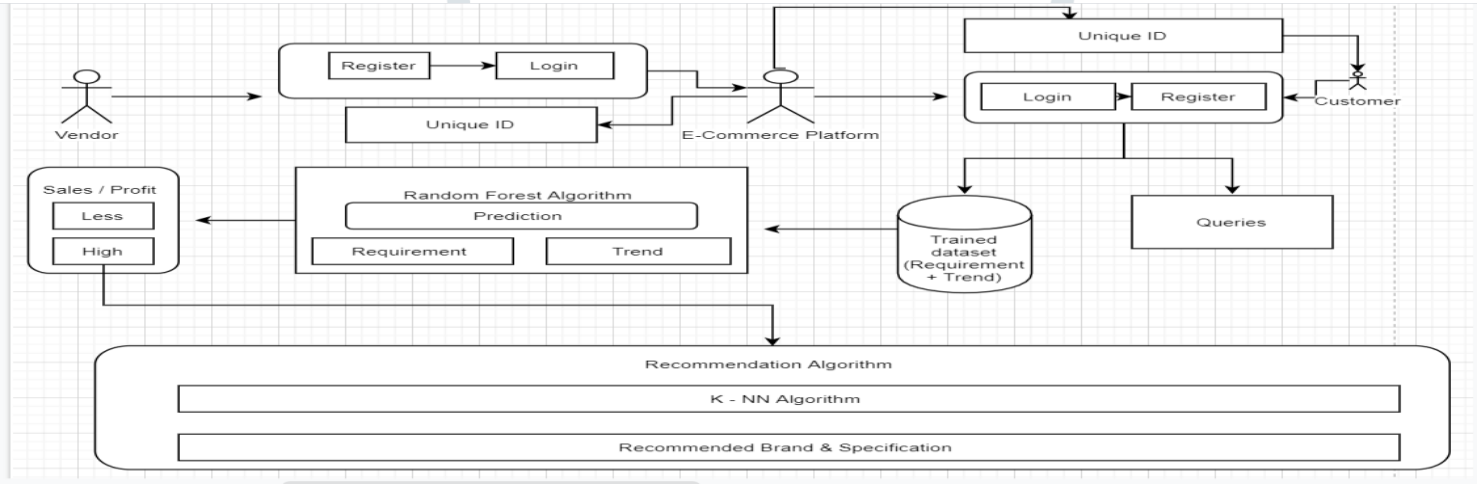

Fig. SYSTEM ARCHTECTURE

IV. PROBLEM DESCRIPTION

4.1 METHODOLOGY: To develop the smart attendance management system, some steps are required to be followed. The steps can be defined in the following ways:

- WEB BASED ANALYSIS FOR MOBILE PREDICTIONS
- IDENTITY MANAGEMENT
- DATA VALIDATION & PREPROCESSING TECHNIQUES
- WORKING MODEL OF REGRESSION METHOD
- COMPARISON OF MACHINE LEARNING ALGORITHM ACCURACY RESULTS
- GUI BASED PREDICTION OF BRAND BY MOBILE FEATURES

4.1.1 WEB BASED ANALYSIS FOR MOBILE PREDICTION

We have to inspect the page to see, under which tag the data and inspect the page extract the Price, Name, and Rating which is nested in the ID tag respectively. After extracting the data, you might want to store it in a format. it will store the extracted data in a CSV (Comma Separated Value) format. Web scraping is an automated method used to extract large amounts of data from

websites. The data on the websites are unstructured. Web scraping helps collect these unstructured data and store it in a structured form.

4.1.2 IDENTITY MANAGEMENT

The actors are Admin (E-Commerce), Vendor and Customer. Both the vendor and customer should enter their details for registration. As soon as the registration is completed, their data is stored in DB. The DB is managed by admin, the admin will provide unique identity to both the customer as well as to the vendor. Once the user logged into the system, there will be session management for 20 minutes. If time exceeds, the user will be logged out to the home page. This will be supported with basic universal English language.

4.1.3 DATA VALIDATION AND PREPROCESSING TECHNIQUE

To analyzing the variable identification by data shape, data type and evaluating the missing values, duplicate values. A validation dataset is a sample of data held back from training your model that is used to give an estimate of model skill while tuning models and procedures that you can use to make the best use of validation and test datasets when evaluating your models. Data cleaning / preparing by rename the given dataset and drop the column etc. to analyze the uni-variate, bi-variate and multi-variate process. The primary goal of data cleaning is to detect and remove errors and anomalies to increase the value of data in analytics and decision making. Data Preprocessing is a technique that is used to convert the raw data into a clean data set. To achieving better results from the applied model in Machine Learning method of the data has to be in a proper manner.

4.1.4 WORKING MODEL OF REGRESSION METHOD

To rid of the limitation of linear regression solution, ML scientists applies non-linear functions such as the Sigmoid and inverse tangent functions onto linear functions. Thus, a perceptron is created and took the form h(x) = g(f(x)) where g(z) is a non-linear transformation of a linear function f(x) = w^T x. However, researchers soon realize an import draw back to a single perceptron solution. On the direction that is perpendicular (or orthogonal in multidimensional space) to the weight vector w, the predictions do not change. This flaw would greatly increase the error rate as a whole direction of predictions is error-prone. Ultimately, the solution is simple: use more perceptrons. While each perceptrons has a blind spot, that spot would be covered by other perceptrons. This divide and conquer technique can be seen many other parts of Computer Science, such as quicksort in the sorting problem.

4.1.5 COMPARISON OF MACHINE LEARNING ALGORITHM ACCURACY RESULTS

SVM:

- ❖ Its representation of different classes in a hyper-plane in multidimensional space. The hyper-plane will be generated in an iterative manner. So, that the error can be minimized and to divide the given datasets into classes to find a maximum marginal hyper-plane (MMH).
- ❖ It will generate hyper-planes iteratively that segregates the classes in best way.
- ❖ It will choose the hyper-plane that separates the classes correctly.

KNN:

- Load the given dataset and initialize ‘k’ value by chosen number of neighbors.
- Calculate the distance between trained customer behavior’s and new customer behavior.

- Sort the collection of distances in ascending order of customer behaviors by the distances and pick the first 'k' entries from sorted collection.
- So, get the labels of the selected 'k' entries and return the mode of the 'k' labels.

RFA:

Random forests or random decision forests are an ensemble learning method for classification, regression and other tasks, that operate by constructing a multitude of decision trees at training time and outputting the class that is the mode of the classes (classification) or mean prediction (regression) of the individual trees.

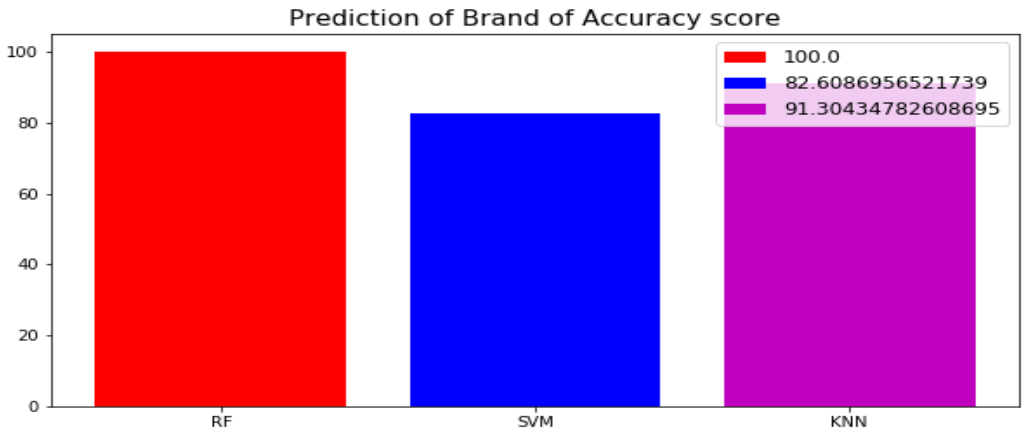

Fig. Comparison of machine learning accuracy results

4.1.6 GUI BASED PREDICTION OF BRAND BY MOBILE FEATURES

Input:

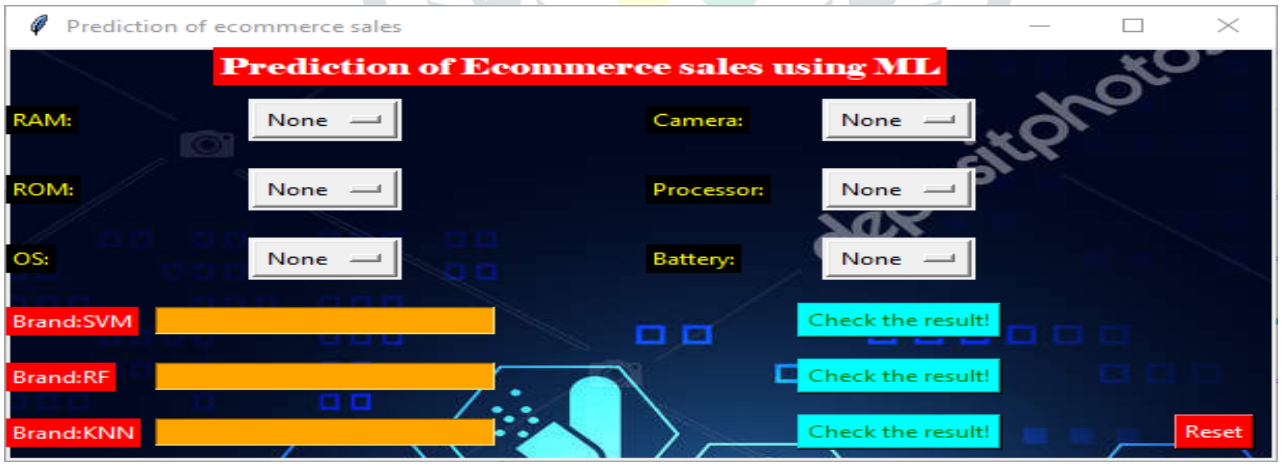

Fig. GUI BASED PREDICTION

V. RESULTS

We can easily predict the sale of the mobile phone using previous year sale as well as also can predict the new product sale before the product launch using machine learning algorithms. This will give a huge profit for both the vendor as well as the E-Commerce platform. The Platform ensures that the customer trust them with product.

VI. CONCLUSION

The analytical process started from data cleaning and processing, missing value, exploratory analysis and finally model building and evaluation.

VII. FUTURE WORK

- To automate this process by showing the result in desktop application.
- To optimize the work to implement in Artificial Intelligence environment.

VIII. REFERENCE PAPER

[1] Herio Susanto, Yudho Giri Sucahyo, Yova Ruldeviyani, Arfive Gandhi "Analysis of Factors that Influence Purchase Intention on Omni-channel Services" 2018 Universitas Indonesia.

[2] Yu-Chien Chai, Ying-Fang Huang, and Hoang-Sa Dang "Determining the Effects of Marketing Mix on Customers' Purchase Decision Using the Grey Model GM(O,N) - Case Study of the Western style Coffeehouse Chains in Vietnam" 2017 International Conference on System Science and Engineering (ICSSE).

[3] Fang Gao "A Study of Online Purchase Intention: Based on the Perspective of Customer Trust" 2019 School of Business, Jianghan University Wuhan, P.R. China.

[4] Weiwei Han , Hua Bai "Influence of Cognitive Resource Limitation on Consumer Purchasing Decision: An Event-related Potentials Perspective" 2018,School of Economics and Management, Beijing University.

[5] M. Octaviano Pratama, Ruci Meiyanti, Handrie Noprisson, Arief Ramadhan, Achmad Nizar Hidayanto "Influencing Factors of Consumer Purchase Intention Based on Social Commerce Paradigm" ICACSIS 2017.

[6] Chen Zhou, Hao Jiang , Jing Wu, Jianguo Zhou, , and Shuwen "Understanding the Patterns behind Purchasing Capability: A Case Study of Smartphone Consumers" Wuhan University, Wuhan, China.

[7] Li-Wen Chuang, Shu-Ping Chiu, "Analysis on consumer repeat purchase behavior of buying green products" 2017 Fuzhou University of International Studies & Trade, China.

[8] Hazliza Haron , Erny Hayati Johar , Zarifah Fadilah Ramli "Online Opinion Leaders and Their Influence On Purchase Intentions" 2016, Department of Business Management, Perak, Malaysia.

[9] Mohamed Bilal Basha "Driving Factors of Purchase Intention Towards Organic Food: A Conceptual Study" 2014, Business School, Al Khawarizmi International University College, Abu Dhabi, UAE.

[10] Yi-Cheng Ku , Yi-Ming Tai "What Happens When Recommendation System Meets Reputation System? The Impact of Recommendation Information on Purchase Intention" 2014, National Pingtung Institute of Commerce, Taiwan.

[11] LIANG Jian-ming , WEI Hai-ying "Impact of Perceived Risk on Purchase Intention in Product-harm Crisis" , Management School, Jinan University.

[12] Hanyang Luo, Jingjing Wang, Ying Bi "Empirical Research on Consumers' Intention to Purchase Online" , College of Management Shenzhen University.

[13] Archana Anand Boppolige , Anjula Gurtoo "Exploring Viral Phenomenon Methodology for Tangible Purchased Products: Case of iPod and iPad" ,IEEE Senior Member.