Prediction of Product Purchase using Machine Learning

LEELAVATHI R
Assistant Professor
Information Technolgy
Bannari Amman Institute of
Technology, Sathyamangalam.

Mail id:leelavathi@bitsathy.ac.in

AISHWARYA G
UG scholar
Information Technology
Bannari Amman Institute
of Technology, Sathyamangalam

aishwarya.it17@bitsathy.ac.in

ABINAYA S
UG Scholar
Information Technology
Bannari Amman Institute of
Technology, Sathyamangalam.

abinaya.it17@bitsathy.ac.in

ELAKIYA E
UG scholar
Information Technology
Bannari Amman Institute
of Technology, Sathyamangalam

elakiya.it17@bitsathy.ac.in

I.ABSTRACT

An online sale is a sale held over the internet. Over the past ten years, the prediction of User purchase has been considered very important. The main purpose of this research is to construct the purchase prediction for users using machine learning. While an attempt made for purchase intention, features such as ID, gender, age and salary are considered. Here, using a Machine learning algorithm it is easy to recommend the products whether they could purchase or not based on their salary and age. The main aim of this work is to examine using data driven machine learning, whether specific objective and readily measurable factors influence user's decisions. Machine Learning consist of two types -Supervised Learning and Unsupervised Learning .Supervised Learning consist of Classification and Regression. This paper employs two representative classification methods: KNN and SVM algorithm.

Keywords- Purchase Behavior, Machine Learning, Classification, K- Nearest Neighbor (KNN) and Support Vector Machine (SVM)

II.INTRODUCTION

Prediction is the process of making future prediction based on past and present data and most commonly by analysis of trends. Machine Learning is that the science of obtaining computers to act while not being explicitly programmed. It is the subset of Artificial Intelligence (AI) and it provides to develop the acceptable solutions on the basis of existing algorithms and datasets. In a way, Machine Learning works comparatively to human learning.It is applied in various fields in everyday life like using Facebook, Netflix, Amazon, and E-mail and so on.
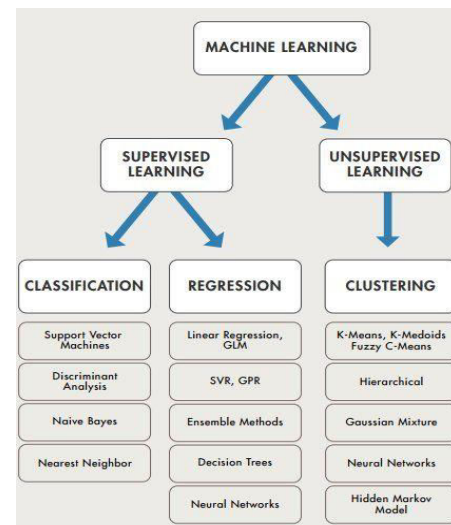

Fig1. Types of Machine Learning

A large number of prediction methods are available. Machine Learning Approaches is one of the major method. The device getting to know technique makes use of area of computational intelligence which assemble computer packages that routinely enhance with experience. In actual existence most of the statistics available for evaluation is of nonlinear and non-stationary kind. The majority of practical machine learning uses supervised learning. In this supervised learning X variables are takes as input and Y variables are taken as output the machine learning algorithm is used to learn the mapping function from the input to the output.

Y=f(X)

The main aim to comparative the mapping function so that the new input data (X) can predict the output variables (Y) for that data. Classification is one of the technique which is used in machine learning, and it has broad application sentiment analysis, ad targeting, risk assessment, diagnosis etc. The main objective of classification is to predict a category or class Y from some inputs X. There are number of classification models. Classification models include Decision Tree, Gradient-boosted Tree, K-Nearest Neighbor, Random Forest, Naive Bayes, and Support Vector Machine etc. Among these the project is done through K-Nearest Neighbor classification algorithm.

III.METHODOLOGY

I) K-NN ALGORITHM

The K-NN algorithm is the simplest among all other machine learning algorithms'-NN is one of the most known supervised learning algorithm in pattern classification. In KNN, the output is a class membership. For each data point, the algorithm finds the K nearest observation, and then classifies the data point to the maximum. K nearest neighbors are found near the query point element. Calculate distance between query point and each element of matrix. The distance which is less that is considered to be nearest neighbor. To calculate the distance, Euclidean Distance method is used. The advantage of KNN is that the cost of learning process is zero, no assumptions about the characteristics of the concepts to learn have to be done and the complex concepts can be learned by local approximation using simple procedures. It is a non-parametric and lazy learning algorithm. In KNN makes prediction using the training dataset directly. In KNN k is the number of nearest neighbor and it is the deciding core factor. Generally most of the k values are in odd numbers because it contains classes that are in even numbers. The given below graph contains two classes namely class A and class B. For this k value is one. Consider C is the point, for which label needs to predict. Find the closest point to the C and then classify points by majority vote of its k neighbors. Each object votes for their class and the class with the majority votes is considered as the prediction. For finding the most nearer point calculating the distance between the points using distance measures such as Euclidean distance, Manhattan distance, Hamming distance and Minkowski distance. KNN has some three basic steps

first step is to calculate the distance second step is to finding the neighbor elements and final step is to vote for labels.
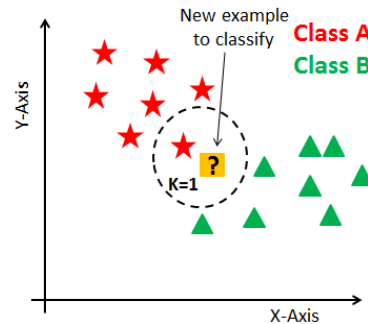

Fig2. Model Graph of KNN

II) SUPPORT VECTOR MACHINE

¨Support Vector Machine (SVM) ¨ could be supervised machine learning algorithmic program which may be used for each classification or regression challenges. However, it's principally utilized in classification issues. During this algorithmic program, it tend to plot every knowledge item as a degree in n-dimensional with worth of every feature being the worth of a selected coordinate. Then, it tend to perform classification by finding the hyper - plane that differentiate the two categories.
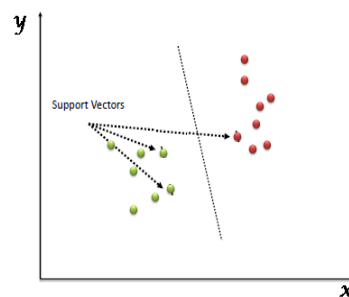

Fig4. Model Graph of SVM

Support Vectors are merely the co -ordinates of individual observation. Support Vector Machine could be a frontier that best separates the two categories (hyper-plane/line). Application of SVM includes text and machine-readable text categorization, as their application will considerably cut back the necessity for labeled coaching instances in each the quality inductive and transductive settings. The advantage of SVM is that it works really well with clear margin of separation and it effective in high dimensional spaces. The SVM algorithm is to find the

hyperplane in an N dimensional space. Hyperplanes are the decision boundaries which classifies the data values. The hyperplane dimension fully depends upon the number of features.

$$f(x) = B(0) + sum(ai * (x,xi))$$

This is an equation that involves calculating the inner products of new input vectors(x) with all support vectors in training data. The coefficients BO and ai (for each input) must be estimated from the training data by the learning algorithm.

IV.EXPERIMENTAL RESULT

We see that K-NN algorithm have performed better than others. The execution involves the following steps:

i) Import the required Python machine learning libraries

The library is built upon SciPy that is installed before use of scikit-learn. This includes:

- Pandas: Data structures and its analysis
- NumPy:base n-dimensional array package
- Matplotlib:Comprehensive 2D/3D plotting
- SciPy: Scientific computing

ii) Load the dataset

Before start the machine learning project, must load the dataset. The most common format for machine learning data is CSV files. Python has more number of ways to load a CSV file. We import our dataset called Social_Network_Ads.csv, has [400 rows x 5 columns].Assign the variables x and y called dependent and independent variables respectively.

iii) Splitting the dataset into training set and test set

The loaded data is usually split into training data and test data. The training set comprises of known output and therefore the model learns on this data so as to be generalized to alternative data shortly. The test data is to test the model prediction.Split our dataset which contains 75% of training data and 25% test data.
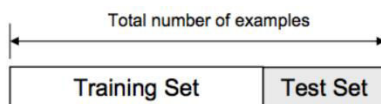

Fig5.Splitting the data

iv) Feature Scaling

Feature Scaling is a methodology which is used to standardize the range of independent variables or other data features. It general work to normalize the data about to a particular range and also helps in fast calculation of data in an algorithm.

v) Fitting the training set

- Fitting K-NN to the training set

K-NN makes prediction directly by the training dataset. Predictions for new instance by searching the complete training set for the K instance neighbors and result the output variable for those K instances. In classification, this the common class value. The K in training dataset are similar to new input a distance measure is used.For,the real-valued input, Euclidean distance is the popular method. Where k=sort(n).In this project, we have taken a K value as 5.

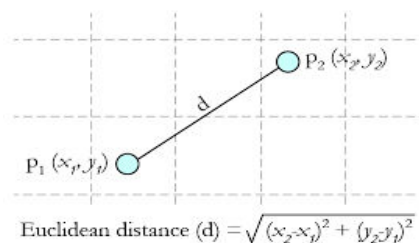$$\text{Euclidean distance (d)} = \sqrt{(x_2 - x_1)^2 + (y_2 - y_1)^2}$$

Fig6.Euclidean distance

- Fitting SVM to the training set

Support Vector Machine, the learning of the hyperplane in linear SVM is done by transforming the problem using some linear algebra. It is implementing in practice. A powerful insight is that the kernel='linear', SVM can be rethought utilizing the input product of any two given perceptions, as

opposed to the perceptions themselves. The input product between two vectors is the entirety of the increase of each pair of info esteems. The dot product is used for linear SVM or kernel because the distance is linear combination of inputs.

The dot-product is called the Kernel,

$$K(x, \ xi) \ = sum(x*xi)$$

vi) Predicting the test set result

A class prediction in which the given finalized model and one or more data instances, predict the class for the data instances. Thus we can predict the class for new data instances using the finalized classification model.

vii) Visualizing the training set and test set results

To visualize the distribution of a variable or the relationship between the multiple variables separately within subsets of a particular dataset. The visualization is also done in high-dimensional datasets in 3D using color, shape, size, and alternate properties of 3D and 2D objects. We predict about the product purchasing as independent data accuracy by taking Salary and Age as dependent data.

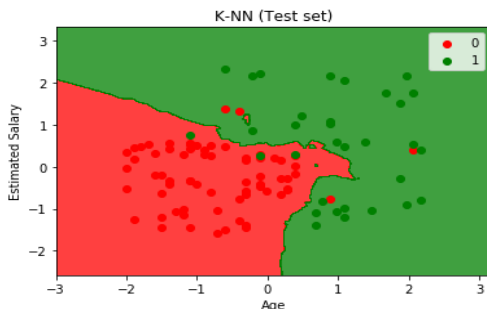- Visualize the test set result using K-NN

Fig7.KNN Test Set

The above plotting shows the accurate prediction of product purchasing by K-Nearest Neighbor Algorithm. By using this our test data accuracy is,

Test data Accuracy: 93%

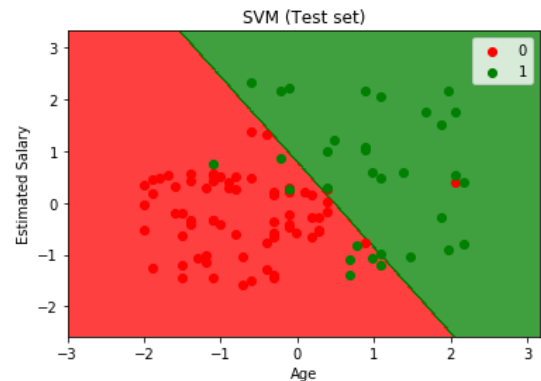- Visualizing the test set result using SVM

Fig8.SVM Test Set

The above plotting shows the accurate prediction of product purchasing by Support Vector Machine Algorithm. By using this our test data accuracy is,

Test data Accuracy: 90%

By estimating these two test set results, the test data accuracy by K-NN Algorithm shows more accurate prediction than SVM Algorithm. Our dataset accurately predict the product purchasing in K-NN Algorithm. In this paper, a prediction of product purchase is done using the concept of machine learning algorithms. It performs better as compared to traditional methods.

V. CONCLUSION

In this paper, we have predicted that the user's purchasing behavior based on the user's Id, age,

customer, salary and gender. By estimating the user's age and salary, which is one in every of the foremost necessary factors impacting one's purchasing decision. Thus we have used the K-NN and SVM Algorithms to predict the user's purchasing behaviour.The SVM to apply to forecast purchasing behavior, which was independent of the distribution and relationship of variables even though they were linear or nonlinear aspects of variables. Second, we use the K-NN Algorithm which makes the better prediction than the SVM Algorithm for our dataset. We hope to build upon this work to the highest accuracy level possible for user behavior extraction.

VI.REFERENCE

[1]Sunil Gupta,Ïmpact of Sales Promotions on When, What, and How Much to Buy¨, Journal of Marketing Research,Vol.25, No.4,pp.342-355,1988.

[2]Shinya Miyazaki,Takashi Washio and KatsutoshiYada,¨Analysis of Residence Time in Shopping using RFID data-An application of the Kernel Density Estimation to RFID¨,2011 IEEE 11th International Conference on Data Mining,pp.1170-1176,2011.

[3].I. H. Witten, and E. Frank, "Data Mining: Practical machine learning tools and techniques," Morgan Kaufmann, 2nd Ed., 2005.

[4].Yuta Kaneko, Katsutoshi Yada, "A Deep Learning Approach for the Prediction of Retail Store Sales", Barcelona, Spain, 2017

[5].Peter M. Guarani and John D. C. Little, "A Logit Model of Brand Choice Calibrated on Scanner Data", Marketing Science, Vol.2, No.3, pp.203-238, 1983.

[6].Jeffrey S. Larson, Eric T. Bradlow and Peter S. Fader, "An Exploratory Look at Supermarket Shopping Paths", International Journal of Research in Marketing, Vol.22, No.4, pp.395-414, 2005.

[7]. Thomas S. Robertson and James N. Kennedy, "Prediction of Consumer Innovators: Application of Multiple
Discriminant Analysis", Journal of Marketing Research, Vol.5, No.1, pp. 64-69, 1968.

[8].Shinya Miyazaki, Takashi Washio and Katsutoshi Yada,"Analysis of Residence Time in Shopping Using RFID Data – An Application of the Kernel Density Estimation to RFID", 2011 IEEE 11th International Conference on Data Mining, pp.1170-1176, 2011.

[9].Sam K. Hui. Eric T. Bradlow and Peter S. Fader, "Testing Behavioral Hypotheses using an Integrated Model of Grocery Store Shopping Path and Purchase Behavior", Journal of Consumer Research, Vol.36, No.3, pp.478-493, 2009.

[10].Herb Sorensen, "The Science of Shopping", Marketing Research, Vol.15, pp.30-35, 2003.

[11].Barracliffe, L.; Arandjelovic, O.; and Humphris, G. 2017. ´Can machine learning predict healthcare professionals' responses to patient emotions? In Proc. International Conference on Bioinformatics and Computational Biology 101– 106.

[12].Breiman, L. 2001. Random forests. Machine Learning 45(1):5–32.

[13].Suh, E.; Lim, S.; Hwang, H.; and Kim, S. 2004. A prediction model for the purchase probability of anonymous customers to support real time web marketing: a case study. Expert Systems with Applications 27(2):245–255.

[14].Yi Zuo, "Prediction of consumer purchase behavior using Bayesian network: an operational improvement and new results based on RFID data", Int. J. Knowledge Engineering and Soft Data Paradigms, Vol. 5, No. 2,pp.85-105,2016.

[15].Jurgen A. Doornik and Henrik Hansen, "An Omnibus Test for Univariate and Multivariate Normality", Oxford
Bulletin of Economics and Statistics, Vol.70, Issue Supplement s1, pp.927-939, 2008.