

A Survey on Machine Learning Algorithms

¹Hetal Z. Bhaidasna, ²Zubin C. Bhaidasna

¹Assistant Professor, ²Assistant Professor & H.O.D.

¹Computer Science & Engineering,

¹Parul Institute of Engineering & Technology, Vadodara ,India.

Abstract : Nowadays machine learning algorithms are widely used in Industries, academics, and government institutions. Machine Learning is an emerging research area in artificial intelligent and pattern reorganization field. Machine learning in divided into three category like supervised, unsupervised and reinforcement learning. In this paper various machine learning algorithms have been discussed. These algorithms are used in different application like artificial intelligence, robotics , data mining, image processing, Data Science , predictive analytics, etc. The main advantage of using machine learning is that, once an algorithm learns what to do with train data, it can do its work automatically without explicit input or we can say without human interaction.

IndexTerms – Machine Learning, SVM, Neural Network, Decision Tree, PCA

I. INTRODUCTION

Machine Learning is the application of artificial intelligence which creates system that has ability to automatically learn and improve from experience without any external programming. In recent years many successful machine learning applications have been developed, like fraud detection, Object detection, weather forecasting, robotics etc. Machine learning interest on the extensions of computer programs which is capable enough to modify when new data arrive. Machine Learning based on computational learning and pattern reorganization. It is the most effective approach in the field of data science. Based on the model of machine learning, we can predict result and decision. Based on that result we can produce reliable result for the system. And the result will be taken input in the model for the improvement of the system. So past observation or result will improve the system and model for the future.

For example, prediction of the cricket match will be identifying with the help of machine learning algorithm using model. We can create a model on which we can train the data based on previous match. And on the basis of the train data we can predict result of the current match. And decision of the current match will be added in train database for the next match. It is very complex task to create an accurate model based on rule for the any system.

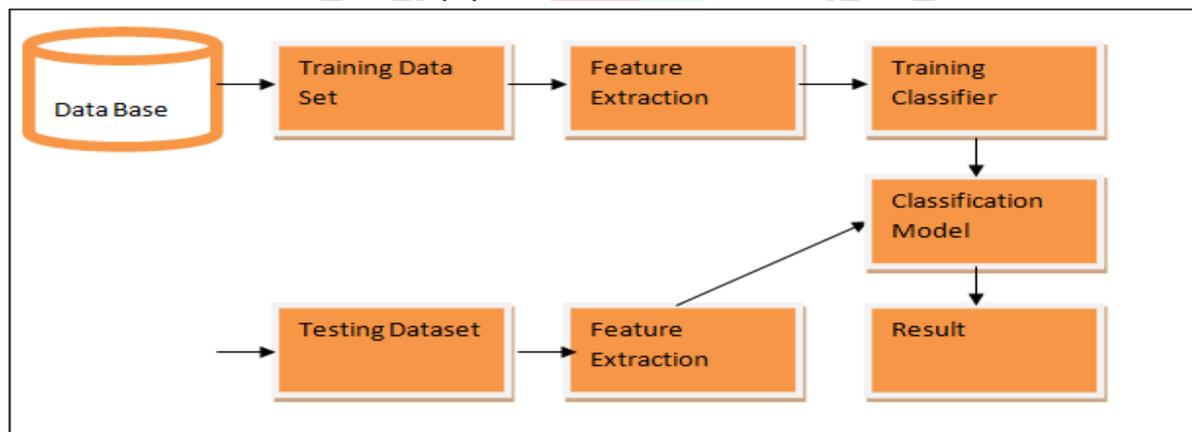


Figure 1 Machine Learning Architecture

Figure 1 shows the flow of machine learning model. In this approach, first we have to prepare training data set based on previous experience and then extract the features and then apply those features on Classification model. After that whenever the new data arrived we have to extract the features and then same classify the using classification algorithm. And after that result will be predict based on the classification algorithm. Machine learning is a paradigm that may refer to learning from past experience means previous experience to improve future performance.

II. MACHINE LEARNING

A more formal definition was given by Tom Mitchell as a computer program is said to learn from experience (E) with respect to some task (T) and some performance measure (P), if its performance on T, as measured by P, improves with experience E then the program is called a machine learning program[1]. In Traditional programming approach, we give input and logic , run it on machine and we will get the output. It will look like,



Figure 2 Traditional Programming

But In machine Learning approach we give input data and also expected output , run it on machine during training and the machine creates own logic , which can be evaluated while testing. It will look like,



Figure 3 Machine Learning Programming

Based on learning capacity, ML divide into three categories: supervised Learning, Unsupervised Learning and Reinforcement Learning. This learning process is based on the comparison of computed output and expected output, that is learning refers to computing the error and adjusting the error for achieving the expected output. Unsupervised learning is termed as learned by its own by discovering and adopting, based on the input pattern. Reinforcement learning is based on output with how an agent ought to take actions in an environment so as to maximize some notion of long-term reward. A reward is given for correct output and a penalty for wrong output [1].

III. MACHINE LEARNING ALGORITHM

In this section, several machine learning algorithms are classified based on learning techniques. Following figure shows the different techniques for the Machine Learning.

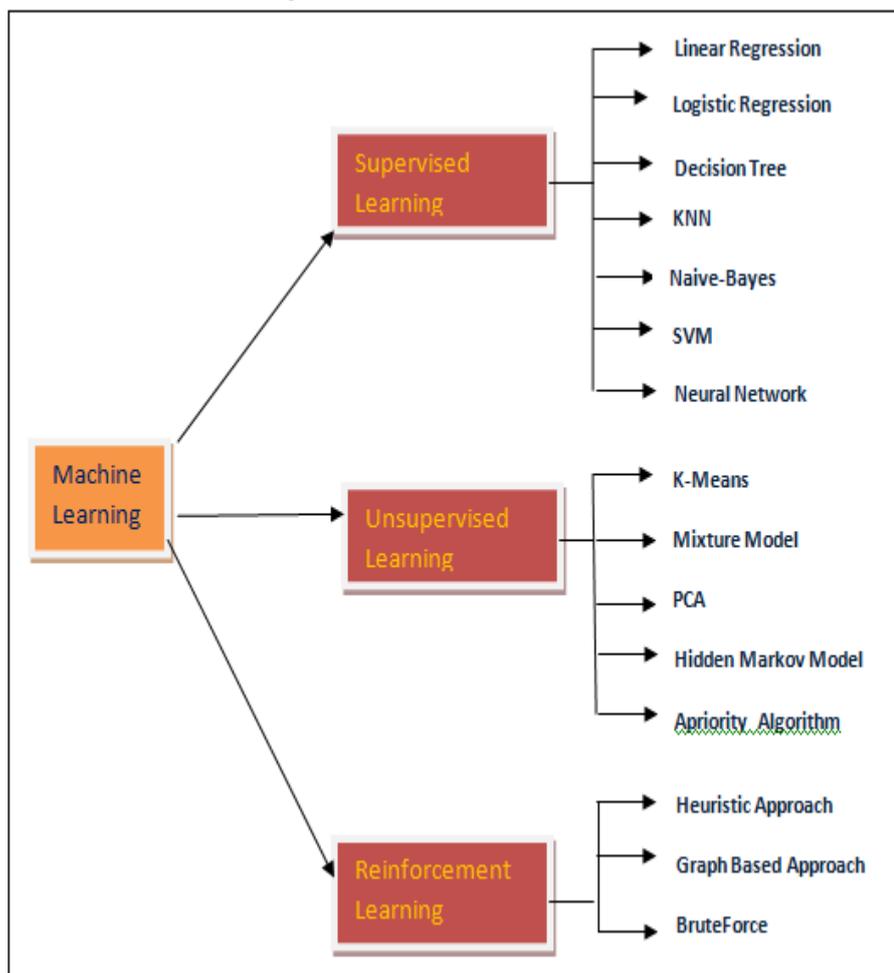


Figure 4 Machine Learning Algorithms

In Supervised Learning all data and input are predetermine with some label and after applying model we can finalize the label with training dataset. For example, we have train the dataset of fruits with label mango, apple banana etc.. And when we give input one fruit like mango then it will predict with mango. In Unsupervised Learning data will be categorized in the basis of similarity or grouping of data. Like based on similarity we can classify data male, female and kids. Unsupervised Learning deals with finding a structure in a collection of unlabeled data. The algorithm is trained to map action to situation so that the reward or feedback signal is maximised. Some methods of Machine Learning will be discuss here,

3.1 Linear Regression

Regression analysis is part of predictive analytics and exploits the co-relation between dependent (target) and independent variables [2]. The simplest mathematical relationship between two variables x and y is a linear relationship. Here from x and y , one is dependent and other is independent variable. Independent variable is also called as predictor and denoted as x. Dependent variable is also known as response variable and is denoted as y. In a cause and effect relationship, the independent variable is cause and the dependent variable is the effect. Linear Regression is a machine learning method for predicting the value of dependent variable y, based on the independent value x. Linear Regression is the best model for the prediction. It also known as predictive model which is used for prediction or forecasting.

3.2 Naive Bayes

Bayes theorem in ML mainly used for classification and clustering. Naive Bayes theorem based on conditional probability. Naive Bayes classifiers performed better in many real time application, like in text classification, spam detection, weather forecasting and Marketing. Only a small amount of training data is required to estimate certain parameters. Naive Bayes can solve diagnostic and predictive problems very efficiently. Following equation shows the Bayesian rule,

$$P(A|B) = \frac{P(B|A) \times P(A)}{P(B)} \dots\dots\dots(1)$$

3.3 Support Vector Machine(SVM)

SVM is the supervised Learning approach. SVM is so popular a ML technique that it can be a group of its own. It uses a separating a decision plane to demarcate decision boundaries among a set of data points classified with different labels. It will perform linear and non linear classification [2]. The main idea of SVM is to construct the hyper plane in a high dimensional space which can be used for classification. Hyper plane refers to a subspace one dimension less than its ambient space[5]. SVM work better on classification with training data. SVM is generally used in stock marketing in financial industries.

3.4 PCA

Karl Pearson invented the principal component analysis in 1991. Orthogonal transformation is the part of statically procedure which makes a set of uncorrelated linear variables known as principal components. In principal component analysis recognition, feature space given eigenfaces, the reduction in dimensionality of original image data takes place. But principal component analysis has large computation and low discrimination power problems which gets eliminated in linear discriminate analysis [3]. In Principal Component Analysis or PCA, the dimension of the data is reduced to make the computations faster and easier.

3.5 Neural Network

Neural Network is the supervised learning algorithm which is widely used algorithm in nowadays. Neural Networks is a field of Artificial Intelligence (AI) where we, by inspiration from the human brain. The neural network is derived from the biological concept of neurons. A neuron is a cell like structure in a brain. To understand neural network, one must understand how a neuron works. They are dendrites, nucleus, soma and axon. The dendrites receive electrical signals. Soma processes the electrical signal. The output of the process is carried by the axon to the dendrite terminals where the output is sent to next neuron. The nucleus is the heart of the neuron. The inter-connection of neuron is called neural network where electrical impulses travel around the brain [4]. In this approach data can be trained by adjusting the weights of the inputs with Supervised Learning. In this learning technique, the patterns to be recognised are known in advance, and a training set of input values are already classified with the desired output. Before commencing, the weights are initialised with random values. Each training set is then presented for the perception in turn. For every input set the output from the perception is compared to the desired output. If the output is correct, no weights are altered. However, if the output is wrong, we have to distinguish which of the patterns we would like the result to be, and adjust the weights on the currently active inputs. Neural Network are regarded as non-linear models as it tries to discover complex associations between input and output data. A neural network behaves the same way. It works on three layers. The input layer takes input. The hidden layer processes the input (like soma and axon). Finally, the output layer sends the calculated output (like dendrite terminals).

3.6 K-Means

K-Means clustering is one of the simplest algorithm which uses unsupervised learning method to solve known clustering. K-mean is a type of unsupervised algorithm which solves the clustering problem. Clustering means grouping objects based on the information which are similar to each other but differentiate from others. The objects which possess similar characteristics are put in the same cluster. This algorithm is called k-means because it creates k distinct clusters. The mean of the values in a particular cluster is the centre of that cluster [4]. This algorithm is easy to implement but the problem is it will only applicable if mean is defined.

3.7 Decision Tree

A classification tree, known as decision tree is one of the most successful supervised learning algorithm. Decision tree creates a graph or tree that design branching to demonstrate every probable result of a decision. It will be representing as a node. In a decision tree representation, every internal node tests a feature, each branch corresponds to outcome of the parent node and every leaf finally assigns the class label. This method work on top –down approach that means it will start with root of the tree [2].

IV. APPLICATION OF MACHINE LEARNING

In today's life we are use different machine learning application knowingly or unknowingly. Here some of them are describe as follow,

- Traffic Forecasting Service
- Face Detection System
- Weather Forecasting
- Stock Management
- Speech Reorganization
- Bio-Surveillance
- Recommended System
- Social Networking Site
- Pattern Reorganization

IV. Conclusion

This paper surveys various machine learning algorithms. Each and every methods has its own advantage and disadvantage. Today each and every person is using machine learning knowingly or unknowingly. From using smart phone, face detection , getting a recommended product in online shopping ,stock marketing, updating photos in social networking sites, weather forecasting. This paper gives an overview of the supervised and unsupervised machine learning algorithms.

REFERENCES

- [1] Sumit Das, Aritra Dey, Akash Pal, Mabamita Roy. 2015. Applications of Artificial Intelligence in Machine Learning: Review and Prospect, International Journal of Computer Applications (0975 – 8887), Volume 115 (9).
- [2] Kajaree Das, Rabi Narayan Behera, 2017. A Survey on Machine Learning : Concepts, Algorithms, and Application”, International Journal of Innovative Research in Computer and Communication Engineering.
- [3] Varsha D. Patil, 2018. A REVIEW ON FACE RECOGNITION TECHNIQUES”, NCETET, ISBN:978-93-87793-03-3
- [4] Ayon Dey, 2016. Machine Learning Algorithms: A Review, International Journal of Computer Science and Information Technologies, Vol. 7 (3) ,1174-117.
- [5] Sunpreet Kaur, Sonika Jindal, 2016. A Survey on Machine Learning Algorithms, International Journal of Innovative Research in Advanced Engineering (IJIRAE) ISSN: 2349-2763 Issue 11, Volume 3.
- [6] Dr. Rajesh A ,Gender Recognition from Face Image Based on Textural Analysis and Machine Learning Approach, , IJSRSET , Volume 4(1)ISSN: 2395-1990.
- [7] Ajinkya Kunjir, Basil Shaikh, 2017. A Survey on Machine Learning Algorithms for Building Smart Systems, International Journal of Innovative Research in Computer and Communication Engineering, Vol. 5(1).
- [8] Navdeep Singh Jaggi1 , 2017. Survey of Machine Learning Applications, International Journal of Innovative Research in Computer and Communication Engineering, Vol. 5(8).

