International Journal of Recent Development in Engineering and Technology
Website: www.ijrdet.com (ISSN 2347-6435(Online) Volume 10, Issue 8, August 2021)

Taxonomy on Machine Learning Algorithms

Dr. Meghna Utmal
HOD-MCA, GGITS, Jabalpur, (M.P.), India

Abstract **– In the present study we compare different supervised machine learning techniques. Today several domains make efficient use of machine learning therefore our proposed paper discusses the various algorithms employed in machine learning. The primary goal of machine learning algorithms is to make machine do work automatically once it learns and complete the task as per the user requirements.**

Keywords **– Machine learning, algorithms, Supervised Machine learning, Data Mining, image processing, predictive analytics**

I. INTRODUCTION

Machine learning is an area of artificial intelligence (AI) and computer science that focuses on using data and algorithms to mimic the way humans learn, with the goal of steadily improving accuracy.

Machine learning is a crucial part of the rapidly expanding discipline of data science [1]. Algorithms are trained to generate classifications or predictions using statistical approaches, revealing crucial insights in data mining initiatives. Following that, these insights drive decision-making within applications and enterprises, with the goal of influencing important growth KPIs. As big data expands and grows, the demand for data scientists will rise, necessitating their assistance in identifying the most relevant business questions and, as a result, the data needed to answer them.

The goal of machine learning is to gain knowledge from data. There have been numerous studies on how to teach robots to learn on their own [2] [3]. To solve this problem, several mathematicians and programmers use a variety of ways. Figure 1 illustrates a few of them.

The following headings are used to categories machine learning:

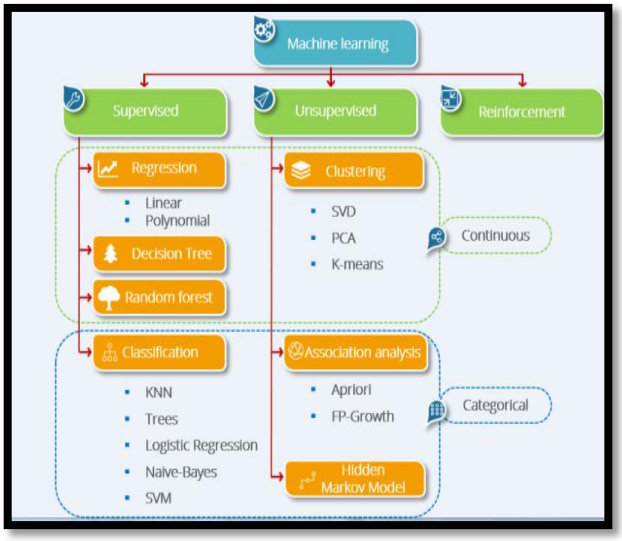


Figure 1: Classification of Machine Learning

1.1 Types of Classification Algorithms

The three basic types of machine learning algorithms are Supervised Learning, Unsupervised Learning and Reinforcement Learning.

In supervised ML we predict from a set of predictors or attributes which can also be termed as independent variables. Mapping of input to specific output is done via set of set of variables. Training on model takes place until desired accuracy is achieved on the training data [4][5].

We don't have any objective or result variable to forecast or estimate in the Unsupervised Learning process. It is commonly used for segmenting clients into separate groups for specialized intervention [6]. It is utilized for clustering population into different groups.

International Journal of Recent Development in Engineering and Technology
Website: www.ijrdet.com (ISSN 2347-6435(Online) Volume 10, Issue 8, August 2021)

The machine is educated to make certain decisions in the Reinforcement Learning algorithm. Trial and error is the key for making the machine learn in this method [7][9]. Previous experience forms the basis for the learning and attempts to capture the most relevant information in order to make appropriate business decisions. The flow of supervised learning is depicted in the diagram.

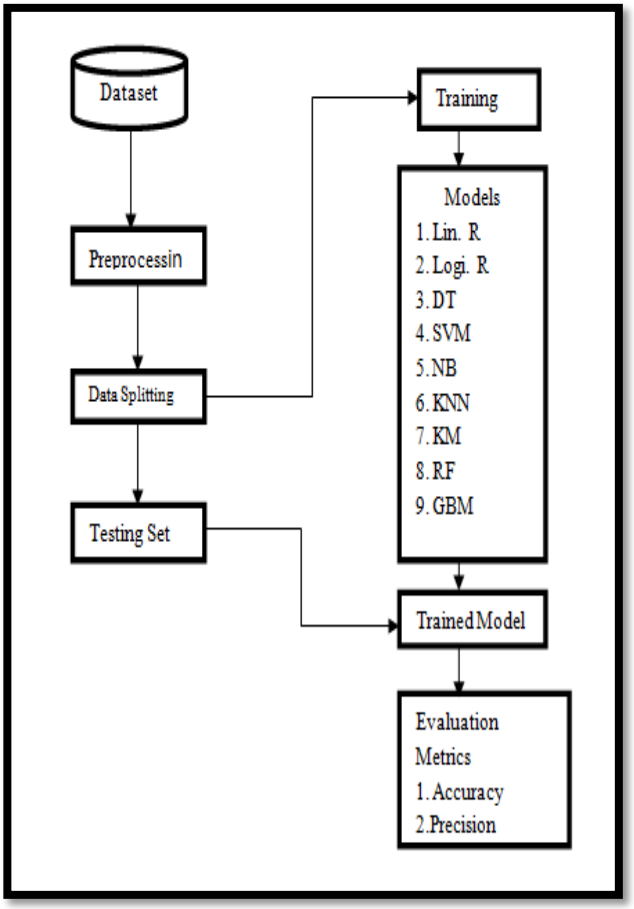


Figure 2: Flow of Supervised Machine Learning

II. Related Work

S.B. Kotsiantis (2007) [4], In this study, he discusses numerous supervised machine learning classification techniques. He also stated that a single paper could not possibly include all supervised machine learning classification algorithms.

P. Harrington (2012) [7], the supervised machine learning algorithms which deals more with classification includes the following: Linear Classifiers, Logistic Regression, Naïve Bayes Classifier, Perceptron, Support Vector Machine; Quadratic Classifiers, K-Means Clustering, Boosting, Decision Tree, Bayesian Networks and so on.

M. Praveena,V. Jaiganesh (2017) [8], this paper focuses on two well-known supervised machine learning algorithms: decision trees and support vector machines, and it presents recent research findings[8]. Additionally, new advancements in Adaboost algorithms (boosting process) are acknowledged. This survey research discovered that combining a supervised machine learning algorithm with a boosting method improved prediction efficiency, and this research aspect has a lot of potential.

Yogesh Singh, Pradeep Kumar Bhatia & Omprakash Sangwan (2007) [9], this study presents a comprehensive assessment of works on expert software development estimation using machine-learning techniques (MLT). In this new era, machine learning is proven its ability to produce consistently correct projections. The most often used machine learning approaches for expert estimates in the field of software development are regression trees, rule induction, genetic algorithm, and genetic programming. We discovered that the outcomes of various machine-learning approaches vary depending on the application areas to which they are applied in each of our studies.

Jason Brownlee (2013) [10], Author Jason Brownlee covered several machine learning approaches in this paper. Also discussed are the benefits and drawbacks of various machine learning algorithms in various techniques [10].

III. Discussion On Supervised Machine Learning Algorithms

In this section, we'll go through the many types of supervised machine learning algorithms. The following are some of them:

3.1 Linear Regression

This is the simplest algorithm for learning. It is an example of a statistical procedure, and it can be used in predictive analysis as well. It makes predictions based on both continuous and quantitative variables such as age, pricing, and sales.

International Journal of Recent Development in Engineering and Technology
Website: www.ijrdet.com (ISSN 2347-6435(Online) Volume 10, Issue 8, August 2021)

We name this method linear regression because it illustrates a linear connection between the dependent variables and one or more independent variables [11]. It shows how the dependent variable changes with the value of the independent variable. Figure 3 shows the linear regression line and the data points plotted near it

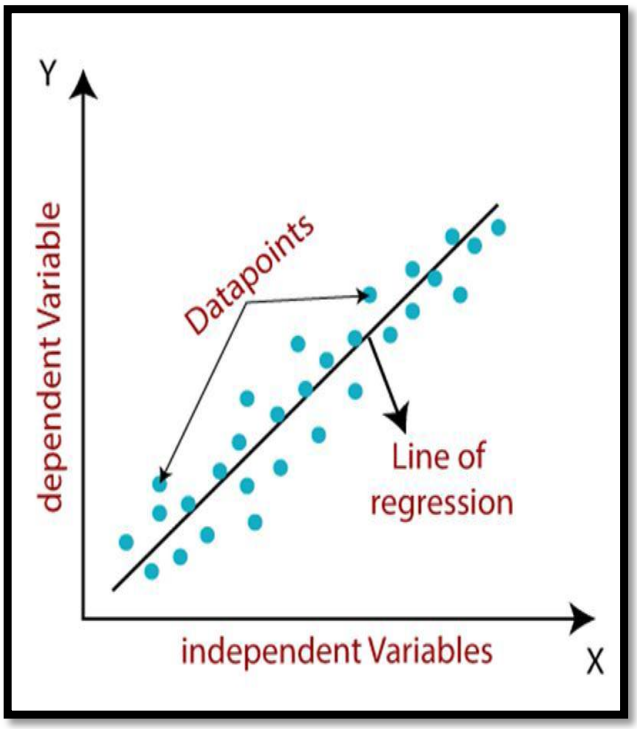


Figure 3: Graph of Linear Regression

Simple linear regression and multiple linear regressions are the two different types of linear regression. Multiple LR is characterized by one or more independent variables while Simple LR is characterized by one independent variable. Polynomial or curvilinear regression is the term for this type of analysis.

3.2 Logistic Regression

It's a classification algorithm and based on set of independent variables it estimates discrete values (like 0/1, yes/no, and true/false). In simple terms, it fits data to a logit function to forecast the probability of an event occurring.

As a result, it's also called logit regression. As a result, it's also called logit regression [12]. Its output values are between 0 and 1 because it forecasts probability.

The likelihood of the presence of the characteristic of interest, p, is given above. Rather of minimizing the sum of squared errors, it sets parameters that maximize the likelihood of witnessing the sample values (like in ordinary regression).

Why would you want to take a log, you might wonder? For the sake of simplicity this is the greatest mathematical way to mimic a step function. I could go on and on, but that would defeat the purpose of this article.

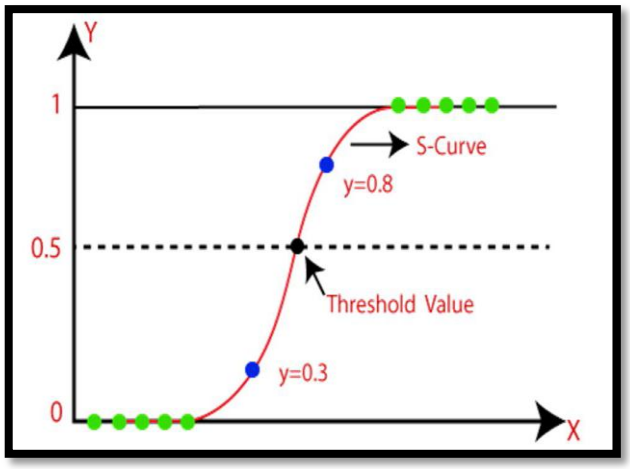


Figure 4: Graph showing Logistic Regression

3.3 Decision Tree

It is a supervised learning strategy that may be used to solve both classification and regression problems, however it is most commonly used to solve classification issues. It's a classifier with a tree structure. The attributes of a dataset are represented by the inner nodes, and the decision rules are represented by the branches, with each leaf node representing the ultimate result. The decision Node and the leaf Node are the two nodes of a decision tree. Leaf nodes are the results of the decisions and have no further sub branches. Decision nodes are used to make conclusions with multiple branches. The figure given below shows the structure of decision tree.

International Journal of Recent Development in Engineering and Technology
Website: www.ijrdet.com (ISSN 2347-6435(Online) Volume 10, Issue 8, August 2021)

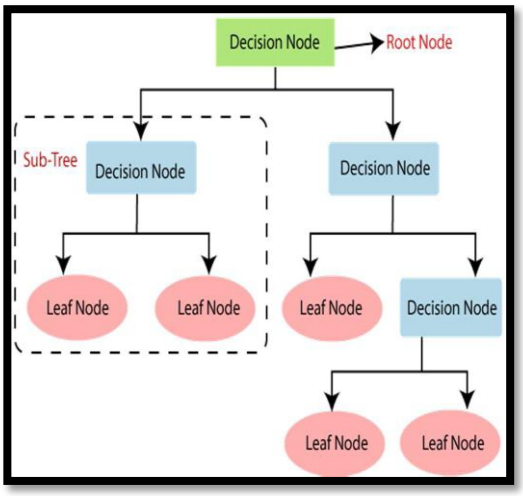


Figure 5: Flow chart of Decision Tree

It works for both categorical and continuous dependent variables, which is surprising. This is done to create as many separate groups as feasible based on the most important attributes/independent variables [13]. For further information, see Decision Tree Simplified.

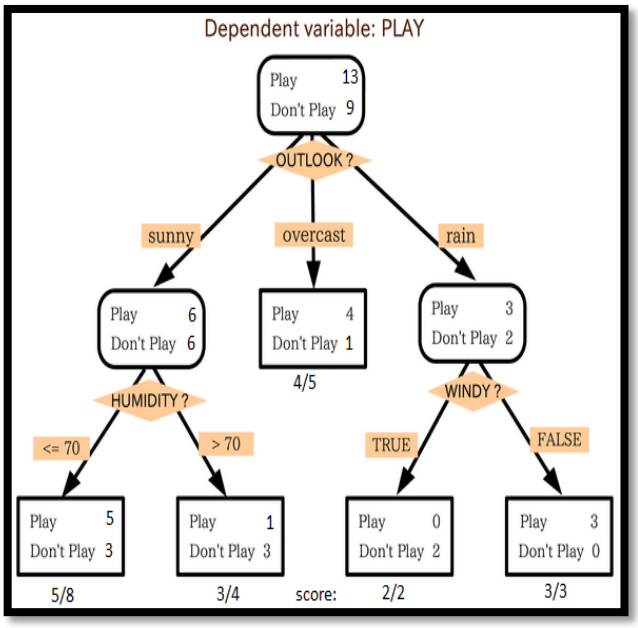


Figure 6: An Example of Decision Tree

In the figure above, you can see how the population is divided into four categories depending on a variety of factors in order to determine 'whether they will play or not.' It employs a variety of approaches, including Gini, Information Gain, Chi-square, and entropy, to divide the population into separate heterogeneous groups.

Playing Jezzball, a classic Microsoft game, is the greatest method to grasp how decision trees function (image below). In essence, you have a room with moving walls, and you must construct barriers so that the largest amount of space is cleared without the use of balls.

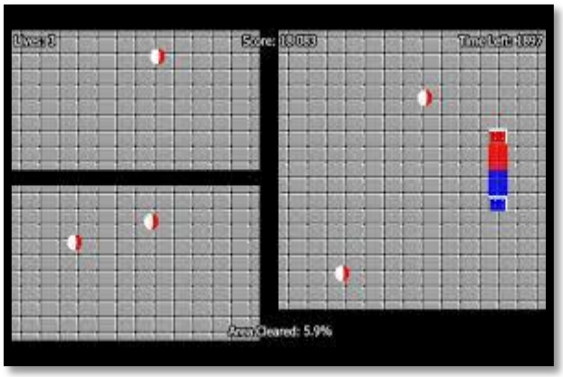


Figure 7: Jezzball Game Snapshot

As a result, every time you divide a room with a wall, you're attempting to establish two distinct populations within the same space. Decision trees act similarly to decision trees in that they divide a population into several groups.

3.4 SVM (Support Vector Machine)

In this every data item is plotted in an n-dimensional space [14].

For example, if we just possessed two characteristics of an individual, which is depicted in the plot below where we have height and hair length in a 2D space with two co-ordinates (these co-ordinates are known as Support Vectors)

International Journal of Recent Development in Engineering and Technology
Website: www.ijrdet.com (ISSN 2347-6435(Online) Volume 10, Issue 8, August 2021)

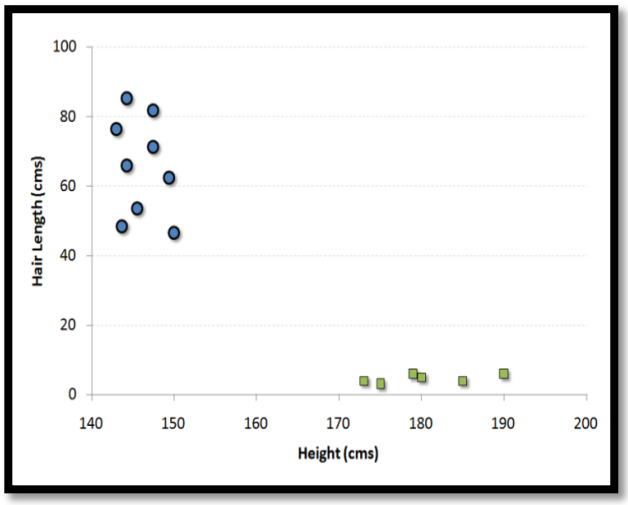


Figure 8: 2-D Plotting of Support Vector in SVM

Now we'll look for a line that divides the data into two groups with distinct classifications. This will be the line along which the distances between the two groups' nearest points are the greatest [15].

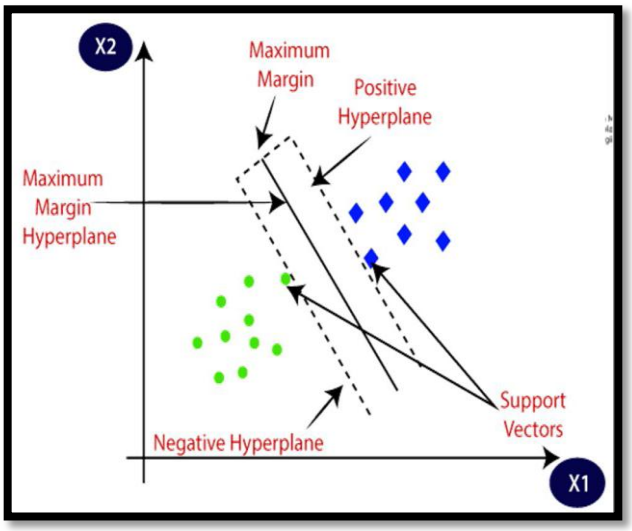


Figure 9: Plotting of SVM with Hyperplane

Because the two closest points are the furthest distant from the line in the example above, the line that divides the data into two differently classed groups is the black line.

3.5 Naive Bayes

The Bayes theorem and the attribute independence assumption are the foundations of Naive Bayes [16]. Because the conditional independence assumption on which it is built is rarely true in real-world applications, its competitive performance in classification is surprising. Some academics have investigated Nave Bayes extensively in text categorization tasks.

This model is well known for its simplicity in construction and can be well worked with huge datasets. It outperforms in even the most advanced classification systems due to its simplicity.

Consider the following equation:

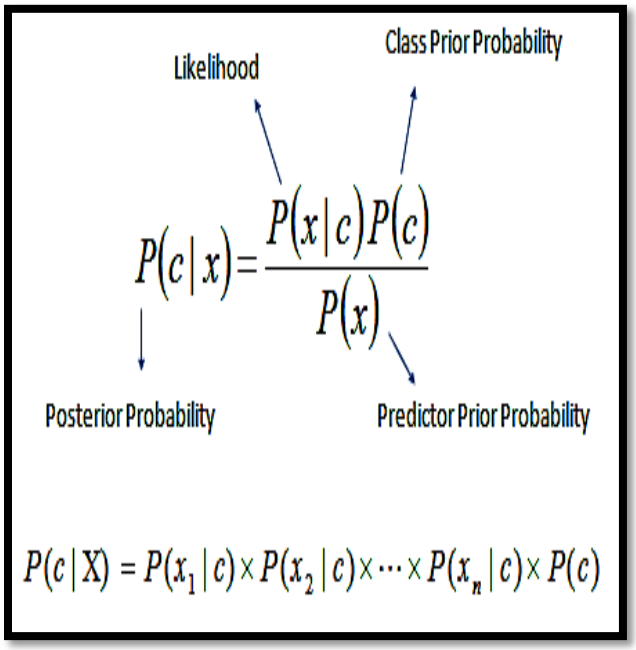
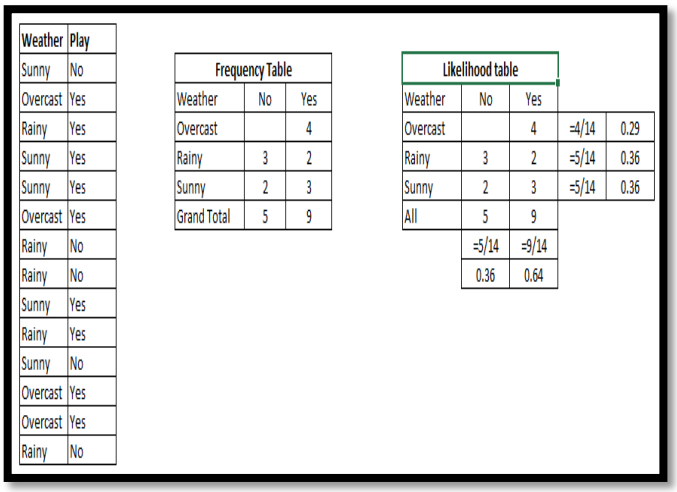


Figure 10: Probability Equation of Bayes theorem

We are taking well known example of sports dataset where we are going to predict whether a player on the basis of certain weather will play or not. For this our first step is construction of frequency table then we make its likelihood table to calculate the probability. The figure below depicts the calculate values for both overcast and playing probability. At the end for each class we find the posterior probability using Naïve Bayes equation and the one with highest probability is the prediction.

International Journal of Recent Development in Engineering and Technology
Website: www.ijrdet.com (ISSN 2347-6435(Online) Volume 10, Issue 8, August 2021)



Weather	Play
Sunny	No
Overcast	Yes
Rainy	Yes
Sunny	Yes
Sunny	Yes
Overcast	Yes
Rainy	No
Rainy	No
Sunny	Yes
Rainy	Yes
Sunny	No
Overcast	Yes
Overcast	Yes
Rainy	No

Frequency Table

Weather	No	Yes
Overcast		4
Rainy	3	2
Sunny	2	3
Grand Total	5	9

Likelihood table

Weather	No	Yes		
Overcast		4	=4/14	0.29
Rainy	3	2	=5/14	0.36
Sunny	2	3	=5/14	0.36
All	5	9		
	=5/14	=9/14		
	0.36	0.64		

Figure 11: An example of Naïve Bayes Classification

On the basis of different attributes a similarity based approach is used by Naive Bayes to forecast the likelihood of various classes. The same technique is also adopted for dealing with multiclass problem and text classification.

3.6 KNN (K- Nearest Neighbors)

It can be used to solve problems involving classification and regression. However, for the purpose of categorization it is mostly used in industries [18]. It is the simplest algorithm that maintains all available examples and classifies new cases based on a majority vote of its k neighbours. Distance function is one of the most important parameter for determining the k number of nearest neighbours in kNN method.

Several distance finding functions like Euclidean, Manhattan, Minkowski, and Hamming are used in this approach. For continuous functions, the first 3 functions are utilised, while for categorical variables, the fourth function (Hamming) is used. When performing kNN modelling, selecting K can be difficult at times.

More: Introduction to k-nearest neighbors: Simplified.

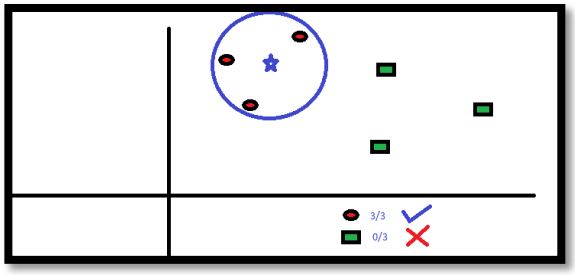


Figure 12: k-nearest neighbors (where k=3)

KNN is easily relatable to our daily life. To know more about a particular person we try to know about him from his closed ones to get more knowledge about him.

Keynotes to be taken for kNN:

- Choosing the appropriate value of K is a major issue.
- Higher variables can bias so its is important to choose normalized variables.
- Before proceeding to kNN it focuses more on pre-processing.

3.7 Random Forest

Random forest comes under the category of ensemble based approach [9]. We have a collection of decision trees (also known as "Forest") in Random Forest. In this technique majority of votes are an important measure for prediction using a majority of votes therefore it is also called as majority voting classifier.

Each tree is planted & grown as follows:

1. For N number of samples in the training data with random sample is taken.
2. For M input variables, a number mM is supplied so that m variables are randomly chosen from the M at each node, and the best split on these m is used to split the node.
3. Each tree is grown to its maximum potential. Pruning is not an option.

3.8 Gradient Boosting Algorithms

3.8.1 GBM

GBM is a boosting algorithm that is utilized when there is a lot of data and we need to generate a prediction with a lot of accuracy. Boosting is a collection of learning techniques that combine the predictions of several different base estimators to increase resilience over a single estimator[19][20]. It builds a strong predictor by combining many weak or mediocre predictors. In data science competitions like as Kaggle, AV Hackathon, and CrowdAnalytix, these boosting methods consistently perform well.

IV. Conclusion

This study examines a number of different machine learning algorithms. Today, everyone, intentionally or unknowingly, employs machine learning. From online shopping for a recommended product to uploading images on social networking sites, there's a lot to do. This document provides an overview of the most widely used machine learning algorithms.

REFERENCES

[1] Mariette Awad, Rahul Khanna. "Efficient Learning machines: Concepts and Applications". Aspress Publishers, 2015.

[2] M. Welling, "A First Encounter with Machine Learning"

[3] M. Bowles, "Machine Learning in Python: Essential Techniques for Predictive Analytics", John Wiley & Sons Inc., ISBN: 978-1-118-96174-2

[4] S. B. Kotsiantis."Supervised Machine Learning: A Review of Classification Techniques", Informatica. pp 249-268, 2007.

[5] Yogesh Singh, Pradeep Kumar Bhatia & Omprakash Sangwan "A review of studies in machine learning technique". International Journal of Computer Science and Security, vol.1, pp 70 – 84, 2007.

[6] Kajaree Das, Rabi Narayan Behera. "A Survey on Machine learning: Concept, Algorithms and Applications", International Journal of Innovative Research in Computer and communication Engineering . vol. 5, 2017.

[7] P. Harrington, "Machine Learning in action", Manning Publications Co., Shelter Island, New York, 2012.

[8] M. Praveena,V. Jaiganesh, "Literature Review on Supervised Machine Learning Algorithms and Boosting Process". International Journal of Computer Applications, ISSN No. 0975 – 8887, vol. 169, 2017.

[9] L. P. Kaelbing, M. L. Littman, A. W. Moore, "Reinforcement Learning: A Survey", Journal of Artificial Intelligence Research, 4, Page 237-285, 1996

[10] Jason Brownlee, "A Tour of Machine Learning algorithms" ,November 25, 2013.

[11] Kajaree Das, Rabi Narayan Behera. "A Survey on Machine learning: Concept, Algorithms and Applications", International Journal of Innovative Research in Computer and communication Engineering . vol. 5, 2017.

[12] Kuan-Yu Chen, Cheng-Hua Wang, "Support vector regression with genetic algorithms in forecasting tourism demand" , Tourism Management 28, pp 215–226, 2007.

[13] L. Rokach, O. Maimon, "Top – Down Induction of Decision Trees Classifiers – A Survey", IEEE Transactions on Systems, , Man, and Cybernetics, Part C (Applications and Reviews) (Volume: 35, Issue: 4, Nov. 2005)

[14] D. Meyer, "Support Vector Machines – The Interface to libsvm in package e1071", August 2015

[15] S. S. Shwartz, Y. Singer, N. Srebro, "Pegasos: Primal Estimated sub -Gradient Solver for SVM", Proceedings of the 24th International Conference on Machine Learning, Corvallis, OR, 2007

[16] D. Lowd, P. Domingos, "Naïve Bayes Models for Probability Estimation", ICML '05: Proceedings of the 22nd international conference on Machine learning August 2005 Pages 529–536https://doi.org/10.1145/1102351.1102418

[17] X. Zhu, A. B. Goldberg, "Introduction to Semi – Supervised Learning", Synthesis Lectures on Artificial Intelligence and Machine Learning, 2009, Vol. 3, No. 1, Pages 1-130

[18] J. M. Keller, M. R. Gray, J. A. Givens Jr., "A Fuzzy K-Nearest Neighbor Algorithm", IEEE Transactions on Systems, Man and Cybernetics, Vol. SMC-15, No. 4, August 1985 Ayon Dey / (IJCSIT) International Journal of Computer Science and Information Technologies, Vol. 7 (3) , 2016, 1174-1179

[19] R. S. Sutton, "Introduction: The Challenge of Reinforcement Learning", Machine Learning, 8, Page 225-227, Kluwer Academic Publishers, Boston, 1992

[20] R. Caruana, "Multitask Learning", Machine Learning, 28, 41-75, Kluwer Academic Publishers, 1997

[21] P. Harrington, "Machine Learning in Action", Manning Publications Co., Shelter Island, New York, ISBN 9781617290183, 2012, www.ijcsit.com.

[22] http://www.simplilearn.com/what-is-machine-learning-and-why-itmatters- article

[23] X. Zhu, "Semi-Supervised Learning Literature Survey", Computer Sciences, University of Wisconsin-Madison, No. 1530, 2005

[24] D. Opitz, R. Maclin, "Popular Ensemble Methods: An Empirical Study", Journal of Artificial Intelligence Research, 11, Pages 169-198, 1999