International Journal of Research Publication and Reviews

Machine Learning Algorithms – A Review

[1]Akashdeep Singh, [2]Suman Lata, [3]Harpreet Kaur

Guru Kashi University, Talwandi Sabo

ABSTRACT:

The study of algorithms and statistical models that computer systems employ to complete a task without being explicitly programmed is known as machine learning (ML). There are several daily-used programmes that incorporate learning algorithms. One of the reasons an online search engine like Google works so well every time it is used to search the internet is because of a learning algorithm that has learned the skill of ranking web sites. These algorithms are used for a number of different applications, including data mining, image processing, predictive analytics, etc. The main benefit of machine learning is that once an algorithm understands how to use data, it can follow out its work manually. This paper provides a quick overview and outlook on the multiple uses of machine learning techniques.

Keywords: Algorithms, Supervised learning, Unsupervised learning, Machine learning, SVM (Support vector machine), KNN

1. Introduction:

The fundamental idea of machine learning will be a great place to start for this essay. machine learning (ML) a form of artificial intelligence (AI) that enables software systems to become better at making predictions about the future. In machine learning a computer programme is given some tasks to complete; if the computer program's measured performance on these tasks improves as it obtains more and more experience completing these tasks, it is claimed that the machine has learned from its experience. Then the machine takes decisions and do prediction based on data. Image recognition is a best and well known example of machine learning. Based on the intensity of the pixels black and white or color photographs it can identify objects as a digital image. Machine learning is widely used in real world like Google assistants, traffic prediction, Google maps, robotics, data mining, banking etc.
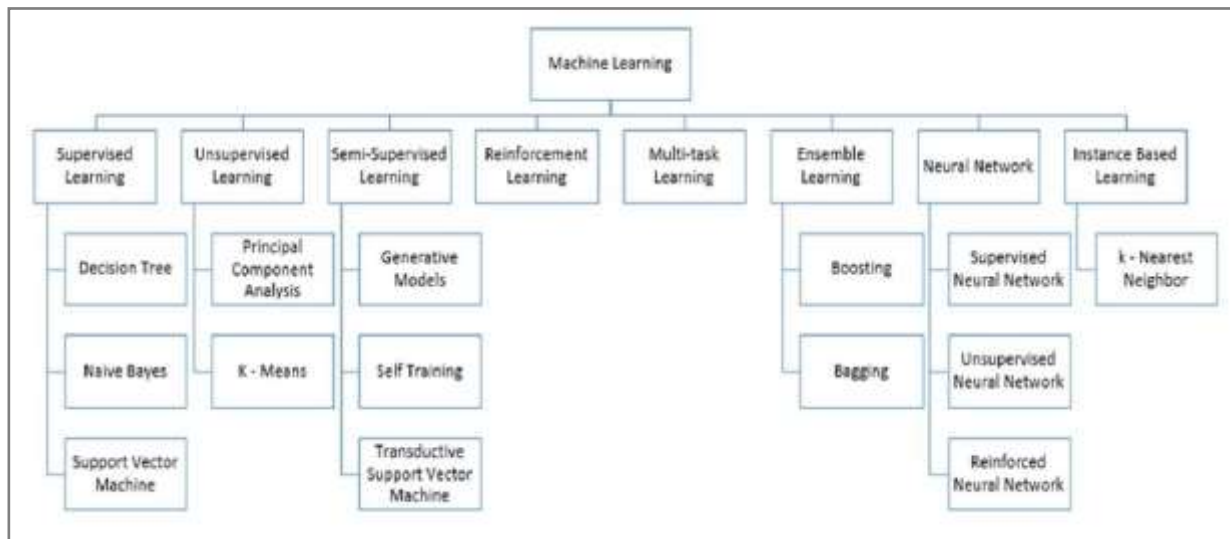

Figure 1. Preview of Machine Learning

2. Supervised Learning:

Supervised learning is the machine learning task of learning a function that maps an input to output based of example input output pairs. It uses labelled training data made up of a collection of training examples to infer a function. Supervised algorithms are those algorithms which needs external assistance. Models are trained using labelled datasets in supervised learning, where the model learns about various types of input. By the completion of the training phase, the model is evaluated using test data (a subset of the training set), and it then makes output predictions. It is easy understanding by flow chart. [1]
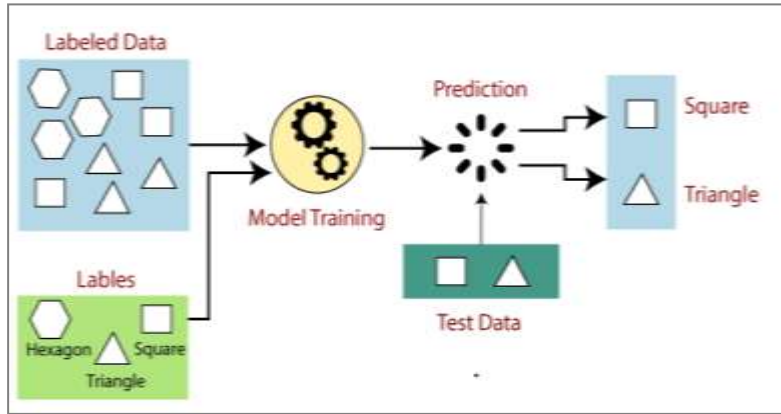
Figure 2. Supervised learning flowchart

3. Decision Tree:

A decision tree is a graph that displays options and their outcome as a tree. It is a type of supervised learning. In the decision tree we both perform classification and regression. The edge of the graph is describing the rules for making decisions, while the nodes in the graph represent a choice. There had node and branches in every tree. Nodes has represent attributes in a group that is to be a for classified and branch represent a value that node can take.[2]
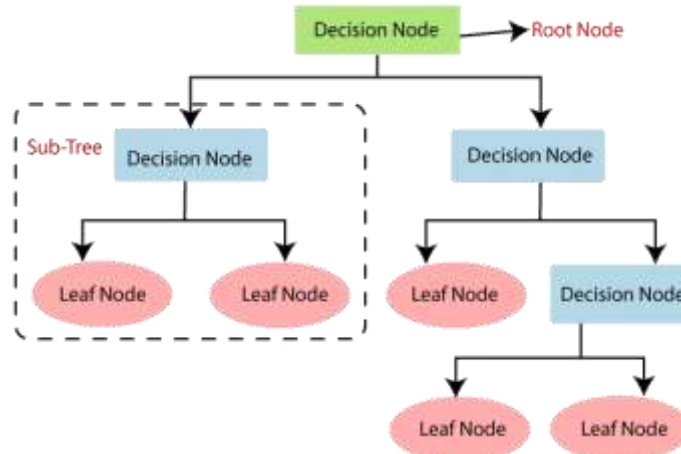
Figure 3 Decision tree graph

4. Support Vector Machine:

SVM is another widely used machine learning technique. It is also a part of supervised learning that study data used for regression and classification analyses. Finding a hyper plane in an N-dimensional space that effectively classifies the data points is the goal of the SVM method. SVM is effective in high dimensional cases. Its memory efficiency comes from the decision function's use of support vectors, a subset of training points.[3]
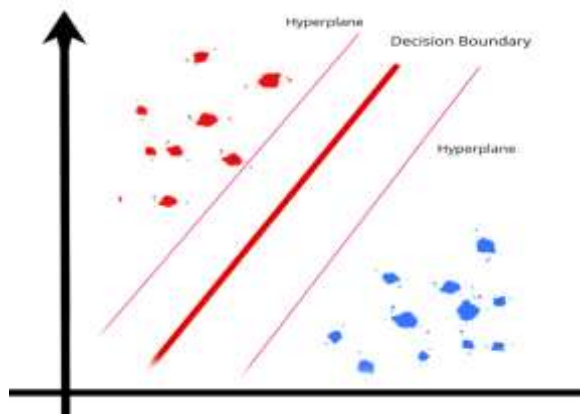
Figure 4 Support Vector Machine

5. Unsupervised Learning:

Unsupervised learning, commonly referred to as unsupervised machine learning, analyses and groups unlabeled datasets using machine learning algorithms. These algorithms identify hidden patterns or data clusters without the assistance of a human. When new data is introduced, it identifies the class of the data using the previously learnt features. It is mostly implemented for feature reduction and grouping. Data labeling requires a lot of manual labor and costs money. Unsupervised learning resolves the issue by classifying the data without the use of labels. The ability to analyses raw data through unsupervised learning makes this the ideal tool for data scientists.[4]

6. K- Means Clustering:

K means is a clustering type. It is the easiest algorithm of unsupervised learning that solve the well known clustering problems. This method of clustering is iterative. According to this method, similar data points should be located close together. It gives us the ability to divide the data into various groups and provides a practical method for automatically identifying the groups in the unlabeled dataset without the need for any training.

The two major functions of the k-means clustering algorithm are:

- Uses an iterative technique to choose the best value for K centre points or centroids.

- Each data point is matched with the nearest k-center. A cluster is formed by the data points that are close to a certain k-center.

It is easy to implement and also difficult to predict the number of clusters.

7. KNN Algorithm:

The K Nearest Neighbor (KNN) technique attempts to categorise data points in a database that contains data points organised into various classifications. The sample piece of data that was provided to it as a classification issue. It uses a large number of labeled points to teach itself how to label new ones. It is a very adaptable classification system that works well for multimodal classes. It can be used for classification and regression to learn non-linear decision boundaries. This algorithm not suitable for large dataset because of time complexity.[5]

8. Applications of Machine Learning:

- Image Recognition

- Speech Recognition

- Traffic prediction

- Product Recommendations

- Self Driving Cars

- Email spam

- Virtual Personal Assistance

- Medical Diagnosis

- Automatic Language Translation[6]

Conclusion:

Machine learning has both supervised learning and unsupervised learning. If you have lesser amount of data then it must be for supervised learning but if you have large dataset then unsupervised learning algorithms had better performance. You also have learned about deep learning and reinforcement. Today every person using a machine learning whether he knows or not. The advantages and disadvantages had also discussed in the paragraphs. In future the Machine learning is the most important thing of our life. Artificial intelligence makes a robot's they are smarter or intelligent than the average person. That is beginning of AI.

References:

[1]. https://www.geeksforgeeks.org/machine-learning/

[2]. Batta Mahesh," A decision tree is a graph that displays options and their outcome as a tree", International Journal of Science and Research (IJSR) ISSN: 2319-7064 Research Gate Impact Factor (2018): 0.28 | SJIF (2018): 7.426

[3]. Batta Mahesh," SVM is another widely used machine learning technique", International Journal of Science and Research (IJSR) ISSN: 2319-7064 Research Gate Impact Factor (2018): 0.28 | SJIF (2018): 7.426

[4]. https://www.javatpoint.com/unsupervised-machine-learning

[5]. 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

[6]. https://www.javatpoint.com/applications-of-machine-learning

[7]. Lata, S., and R. Kumar. "A Hybrid Approach for ECG Signal Analysis." Proceedings - IEEE 2018 International Conference on Advances in Computing, Communication Control and Networking, ICACCCN 2018, 2018, doi:10.1109/ICACCCN.2018.8748858.

[8]. Lata, Suman, and Rakesh Kumar. "Disease Classification Using ECG Signals Based on R-Peak Analysis with ABC and ANN." International Journal of Electronics, Communications, and Measurement Engineering, vol. 8, no. 2, July 2019, pp. 67–86, doi:10.4018/IJECME.2019070105.

[9]. Lata, Suman, and Dheerendra Singh. "A Hybrid Approach for Cloud Load Balancing." In 2022 2nd International Conference on Advance Computing and Innovative Technologies in Engineering (ICACITE), pp. 548-552. IEEE, 2022.