MACHINE LEARNING ALGORITHM AND DEVELOPMENT: AN OVERVIEW

[1]S.V.Kamble, [2]V.C. Sakhare, [3]B.G.Warvante
[1,2,3]Assistant Professor
[1]Information Technology Department, [2,3]Electronics Engineering Department
[1,2]DKTE'S Society's Textile & Engineering Institute, [3]TKIET Warananagar
Dist:Kolhapur, Maharashtra.

Abstract— In this paper, various machine learning techniques (MLT) have been discussed. Machine learning algorithms are used for many applications such as data mining, image processing, predictive analytics, data classification, prediction etc. The primary goal of machine learning is to automate human assistance by training an algorithm on relevant data. Once an algorithm learns what to do with data, it can do its work automatically.

Keywords—Machine Learning Technique (MLT)

1. INTRODUCTION

Machine learning Technique (MLT) is the part of computer algorithms that usually automatically improve performance through experience by use the data [1], it is seen as the part of artificial intelligence. The main purpose of machine learning is to learn from the data. Machine learning algorithms consist of a build a model using sample data this data preferred as "training data". Mostly machine learning algorithm are used in different types of applications includes as medicine, computer vision, and email filtering here with level of difficult or unfeasible to develop conventional algorithms to perform the needed tasks.

Machine learning is used to teach machines how to handle the data more efficiently. At the time after viewing the data, we cannot interpret the pattern or extract information from the data. In this case, we apply machine learning [3].

There are two types of machine learning: a. Inductive b. Deductive. In the inductive machine learning methods consist of computer programs that to be extracting rules and patterns out of massive data sets. Inductive learning takes examples and generalizes rather than starting with existing knowledge one major subclass of inductive learning is concept learning. Deductive learning works on existing facts and knowledge and deduces new knowledge from the old [2].

Machine learning algorithms use computational methods to "learn" information directly from data without relying on a predetermined equation as a model. The algorithms adaptively improve their performance as the number of samples available for learning increases. Machine learning algorithms find natural patterns in data that generate insight and are usually help to take better decision as well as better predication. Nowadays it is used to make critical decision in medical diagnosis, stock trading, energy load forecasting and more.

The use of a machine learning approach in lieu of a more conventional engineering design should be justified on a case-by case basis on the basis of its suitability and potential advantages [7]

All the techniques of machine learning are explained in Section II, in Section III explained machine learning algorithm, Section IV discuss the machine learning development and Section V concludes this paper.

2. MACHINE LEARNING TECHNIQUES

Main aim of Machine learning algorithms to optimize the performance of a certain task by using examples or past experience. Machine learning has two types of techniques namely supervised learning, it has trains a model on known input and output data so that it can predict future outputs, and unsupervised learning, it finds hidden patterns or intrinsic structures in input data.

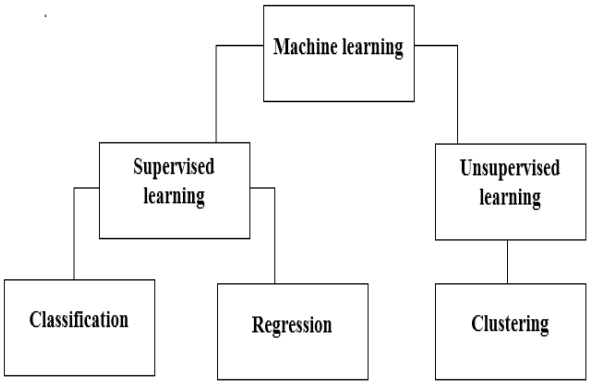

Fig. 1 A Machine Learning Technique

- Supervised Learning

Supervised learning belongs to relatively basic learning method. Supervised machine learning is based on the same principles that as to be a standard fitting procedure, it tries to find the unknown function that connects known inputs to unknown outputs. This desired result for unknown domains is estimated based on the extrapolation of patterns found in the labeled training data [4], during the initial training of the machine, the machine relies data to learn the needs of learning, In the process of collect basic data information, complete the required learning content in a supervised

environment. as compared to other methods of learning supervised learning stimulate the generalized learning potential of the machine [5].

Most of current machine learning applications comes in the supervised learning category, and hence aim at learning an existing pattern between inputs and outputs. Supervised learning is relatively well-understood at a theoretical level [7].

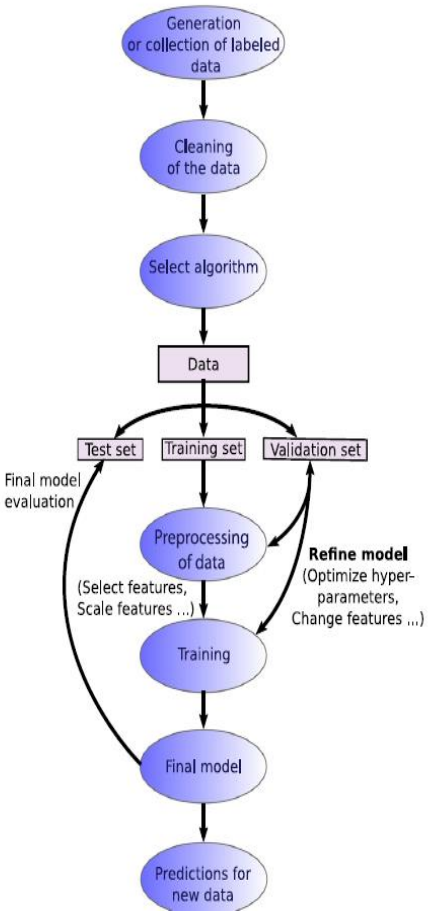

Fig.2 Supervised learning workflow [4]

Supervised learning uses classification and regression techniques to develop predictive models. Regression techniques predict continuous responses.

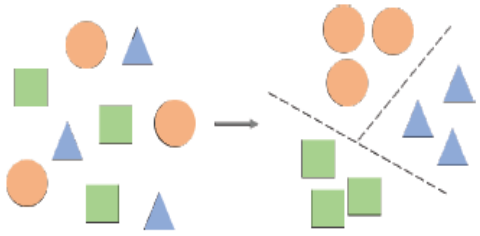

Fig.3 Overview of supervised learning. Input examples are categorized into a known set of classes[8]

- Unsupervised learning

Unsupervised learning finds hidden patterns or intrinsic structures in data. It is used to draw inferences from datasets consisting of input data without labeled responses. Clustering is the most common unsupervised learning technique. It is used for exploratory data analysis to find hidden patterns or groupings in data. In unsupervised learning, the training set consists of unlabelled inputs, that is, of inputs without any assigned desired output. Unsupervised learning is concerned with finding patterns in unlabeled data, as, e.g., in the clustering of samples [4].

Fig.4 Unsupervised learning

- Reinforcement learning

Reinforcement learning lies, in a sense, between supervised and unsupervised learning. Supervised learning and unsupervised learning are also application method of reinforcement learning in machine learning. reinforcement unlike unsupervised learning, some form of supervision exists, reinforcement learning algorithm takes the feedback from the environment only after selecting an output for a given input or observation. The feedback indicates the degree of the output, and its known as action in reinforcement learning, that fulfils the goals of the learner. reinforcement learning is one kind of method that expands data collection based on statistics and dynamic learning [5][7].

3. ALGORITHM FOR MACHINE LEARNING

1. Decision Tree Algorithm:

A decision tree lets we predict responses to data by following the decisions in the tree from the root (beginning) down to a leaf node. A tree consists of branching conditions where the value of a predictor is compared to a trained weight. The number of branches and the values of weights are determined in the training process. Additional modification may be used to simplify the model.

In the order for implementing decision number algorithm analysis of data information will be continue with split the branches and at the time, the branches will be trimmed to improve the integrity of the data content. From calculation point of view algorithm consists of top-down algorithm. During the analysis of content of the node is analyzed for the optimal attributes and after the node is expanded to more than two based on the node. Data information of the split and bacnching like trees can also increases the number of sample that can comprehensively analyze. At the same time it determines contents of in the order for implementing decision number algorithm analysis of data information will be continue with split the branches and at the time, the

branches will be trimmed to improve the integrity of the data content. From calculation point of view algorithm consists of top-down algorithm. During the analysis of content of the node is analyzed for the optimal attributes and after the node is expanded to more than two based on the node. Data information of the split and bacnching like trees can also increase the number of sample that can comprehensively analyze. At the same time it determines contents of sample in the classification according to the sample number of statistics [5].

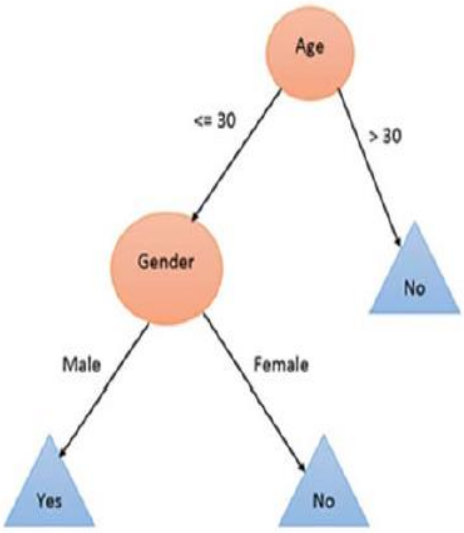

Fig.5 Decision Tree [3]

2. Support Vector Machine (SVM Algorithm)

SVM is also most widely used algorithm content, state-of-the-art machine learning technique is relies to Support on the Vector Machine to complete the data analysis work. It is mainly used for classification. SVM works on the principle of margin calculation, in order to improve final data analysis result in the actual analysis process multiple set analysis samples of data get collected to the determine the sample data of the boundary value. It basically, draws margins between the classes. The margins are drawn in such a fashion that the distance between the margin and the classes is maximum and hence, minimizing the classification error [3][5].

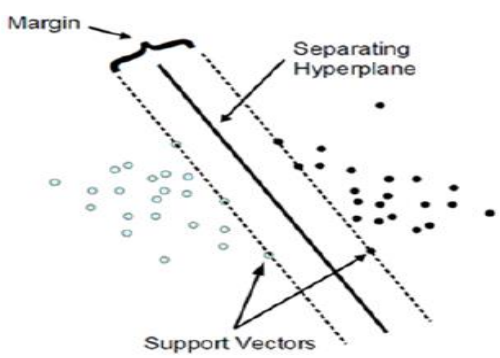

Fig.7. Support Vector Machine [3]

3. K-Means Clustering:

Clustering or grouping is a technique that involves for creates groups automatically. Therefore the item having same 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 center of that cluster. Clustering comes in unsupervised learning technique [3].

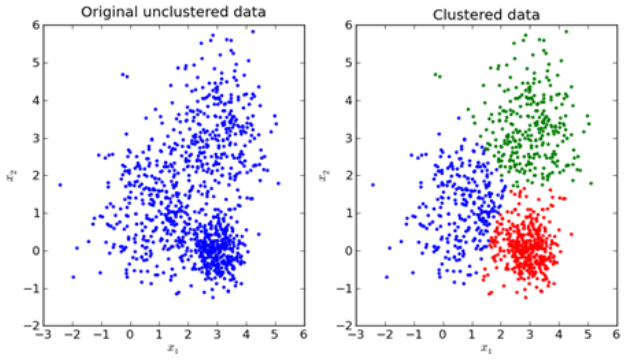

Fig. 6 K-Means Clustering [3]

4. Neural Network Learning

The neural network (or artificial neural network or ANN) is derived from the biological concept of neurons. A neuron is a cell like structure in a brain. It refers to imitating the process of human information transmission, classifying different data into one neuron and connecting the different data neuron to pass information with the help of the internet to achieve the complex memory activities. The inter-connection of neuron is called neural. The neural network has three layers the input layer takes input, the hidden layer processes the input and the output layer sends the calculated output. In order to formulate the analysis and calculation of the data we can set the weighting coefficient in advance and then set the output threshold [3][5].

5. Instance-Based Learning

In instance-based learning techniques that learner learns a particular type of pattern. It consists of applying the same pattern to the newly fed data. Hence the name is assign instance-based. It is a type

of lazy learner which waits for the test data to arrive and then act on it together with training data. The complexity of the learning algorithm increases with the size of the data [3].

6. Boosting Algorithm

Boosting is a technique in ensemble learning which is used to decrease bias and variance it is new type of machine algorithm content. The main advantages of this algorithm consist of accurate processing of data information and improve the accuracy of the final processing result. Boosting creates a collection of weak learners and converts them to one strong learner. In this algorithm function prediction optimized therefore speeding up the processing of data information and at the same time AdaBoost is also an important guarantee for the expansion of the boosting algorithm [5].

4. MACHINE LEARNING DEVELOPMENT

In the future of development process the mechanical theory will be also further optimized and its contents branches will also be expanded. In the formulation of machine learning content is mainly applicable for automation industry. In the context of the rapid development of internet technology the autonomous learning ability of machine will be further made as strengthened. Using supervised learning and unsupervised learning the autonomy that machine learning master will continue to increase. In the future development process IoT, digital technology, cloud computing technology can provide many conditions in the process of data calculation. This technology will be combined with algorithms to form a new technology of application.

5. CONCLUSION

This paper surveys various machine learning algorithms. Supervised, unsupervised, and reinforcement learning paradigms lend themselves to different tasks depending on the availability of examples of desired behaviour or of feedback. Machine learning is mainly relies on supervised learning and does not fully overcome weak artificial intelligence.

REFERENCES

[1] Yogesh Singh,Pradeep Kumar Bhatia,Omprakash Sangwan,"A REVIEW OF STUDIES ON MACHINE LEARNING TECHNIQUES", International Journal of Computer Science and Security, Volume (1) : Issue (1)

[2] W. Richert, L. P. Coelho, “Building Machine Learning Systems with Python”, Packt Publishing Ltd., ISBN 978-1-78216-140-0

[3] Ayon Dey,"Machine Learning Algorithms: A Review ",(IJCSIT)Vol. 7 (3) , 2016, 1174-1179

[4] Jonathan Schmidt,Mário R. G. Marques "Recent advances and applications of machine learning in solidstate materials science",npj Computational Materials (2019) 5:83 ; https://doi.org/10.1038/s41524-019-0221-0

[5] Wei Jin,"Research on Machine Learning and Its Algorithms and Development",Journal of Physics: Conference Series 1544 (2020) 012003,doi:10.1088/1742-6596/1544/1/012003

[6] W. Richert, L. P. Coelho, “Building Machine Learning Systems with Python”, Packt Publishing Ltd., ISBN 978-1-78216-140-0

[7] Osvaldo Simeone, "A Very Brief Introduction to Machine Learning With Applications to Communication Systems",arXiv:1808.02342v4 [cs.IT] 5 Nov 2018

[8] Shagan Sah,"Machine Learning: A Review of LearningTypes",DOI:10.20944/preprints202007.0230.v1

