Review on Machine Learning Algorithm

[1] Ajith B, [2] Shrinivas Sampath M, [3] Pankaj Nandhaa M R, [4] Nethaji T
Department of Electronics and Instrumentation Engineering,
Bannari Amman Institute of Technology, Tamilnadu, India

Abstract: - Nowadays machine learning has become a vital part of Artificial Intelligence. The main objective of Machine learning is to provide a solution to a computer to solve a specific problem based on the past data. It enables the machine to understand and take action according to the pattern recognized. It can be explained and practically seen in Robotics. Machine learning helps to take an efficient decision without any human concern. One of the existing applications includes categorizing the mail between spam and non - spam messages. The concept of machine learning can be applied to the analysis of Supervised Learning, Unsupervised Learning and Reinforcement Learning. This Special topic provides several contents of machine learning to enhance the security and to establish advanced technologies. The material presented here shows various methods to employ Machine learning effectively.

Key words: - Machine Learning, Supervised Learning, Unsupervised Learning.

I. INTRODUCTION

The efficiency and robustness of a human brain is noway near to the most advanced computer available today. For many years scientist have been trying to execute the intelligence of humans via Artificial intelligence. Humans are composition of variety of sensors integrated. Many inventions happened as a result of inspiration of human sensory sectors. The Machine Learning field has evolved from the vast space of Artificial Intelligence, which mainly target at acquiring intelligent ability of human to machines. Machine Learning is a unique method that helps in teaching computers to work exactly like humans (learn from experience). Machine Learning algorithms use computational ways to learn from raw data without depending on any pre-written equations. This improves the efficiency with the increase in sample data (greater the sample data, greater performance) [1] .These data captured is stored and will be useful in future for various purpose. One of the important factor for implementing Machine Learning is how to make machine to understand and learn by own. It automatically learns program and provides solution from the data acquired. Some of the existing application of machine learning include Spam, Web search, Filters in product based websites, designing drugs and so on. As per the report from McKinsey Global Institute that Machine Learning will be the frontier of big Innovation to come on. Machine learning is used to predict output or suggest calculations based on the input data provided. The following companies use machine learning technique to satisfies the user, they are Netflix,amazon..etc, these companies works on big data analytic's and they learn

based on the users input and prefer them better choices next time based on their previous choices. The main difficulty that comes particularly in the context pattern classification application, is that learning complexity grows exponentially with linear dimensional increase. Machine Learning has got information from information stored in the building blocks to make the computers learn to behave far intelligent by generalizing rather by storing and retrieving data's from database or other applications. The separation of data is done to predict the unknown values, using the number of known values. With addition to the spatial dimentionality of real life data, the temporal component which is very much important. A sequence of data conveys a context to the receiver or observer. The content delivered in is irregular format, which is difficult to identify the content. As said temporal component helps in representing information. In machine learning algorithms every set of certain data set is shown by same set of features. It may be categorical, continuous or even binary. As said earlier for known instance variables for which predicted values are gathered is known as supervised learning. For unsupervised learning the instances are unlabeled. By applying the unsupervised algorithms, researchers will be benefited t discover the unknown and various other classes of items [3]. Inventing methods to understand the deeper representations of data is one of the major research frontiers in machine learning [2].And the last algorithm is reinforcement learning here the information are acquired by learning systems by environment which is in the form of signals to communicate. And by which learner does not have any command to specific task instead experimenting different actions to equalize the rightful command [4]. In this paper the main

idea is to provide and contribute upcoming researchers to get up-to-date about the machine learning approaches.

II. METHODOLOGY

There are two methods that are involved in machine learning. Figure 1 shows the types of methods in Machine Learning.
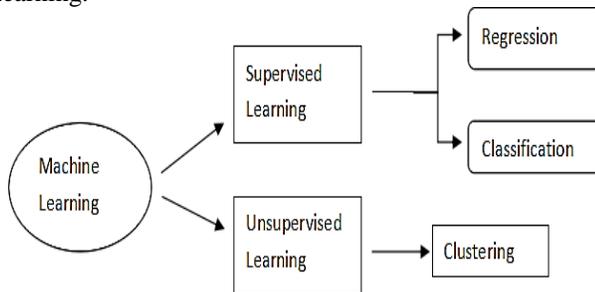

Figure 1: Methods of Machine Learning

A. Supervised Learning

Supervised learning is the one in which the input and the output is given and it gives the output of the future samples based on the known inputs and outputs. A sample is built that predicts based on the evidences without being sure. The supervised learning technique takes the inputs and outputs of many samples and builds a model that can predict the response for the new data. There are two different processes in supervised leaning based on the type of response required. Classification is the process which gives discrete responses. Example, whether an email is spam or not. Regression is the process that gives continuous responses. Example, flow control checking in water tanks.

B. Unsupervised Learning

Unsupervised learning is the one which predicts solutions based on the hidden pattern and structures of the input (this technique does not require any information it predicts the output based on the patterns in the input). Unsupervised learning uses clustering technique to predict responses. This is done when a cluster analysis (it predicts the distribution of a cluster of input so that the response is efficient). The classification technique is used when the data can be categorized. Example, voice recognition uses classification to recognize words. Unsupervised classification of pattern recognition technique is used in image processing and computer vision for object recognition and image segmentation. There is also even further subset in machine Learning types (i.e) Semi supervised Learning and Reinforcement Learning. For Semi Supervised Learning The files or data's required to process is less here, when compared to supervised and unsupervised Learning. The learning by trial and error

method, solely from rewards and punishments is known as Reinforcement learning.

III. ALGORITHMS

Machine Learning algorithms are divided based on their three classifications Supervised Learning, Unsupervised Learning and Reinforcement Learning. The following algorithms are the basic and important algorithms for the machine learning described in a simplest way; they are Decision Trees, Least Square Regression, Naïve Bayes, Logistic Regression, Support Vector Machines, Clustering Algorithms and Principle Component Analysis.

A) Decision Tree
It uses a tree like structure and explicitly calculate based on the decisions represented and by decision making steps. These steps are mostly used in data mining areas. It's like predicting the output based on the input data's provided mostly like Yes/no type. According to authors [7], the decision tree induction, was originally developed to find and study the statistical problems, later on it was developed to handle single or multidimensional regression.

B) Least Square Regression
You can take a set of points and draw a line across those points in a graph. Then measure the distance between the points and the line and sum up all the points distance. The fitted line would be the one where the sum of distance is small as possible.

C) Support vector machines
SVM is called as support vector machines, say there are group of two different points arranged in a graph (N), SVM generates a (N-1) hyper plane to separate those points, and it will separate those points with a linear line and make sure the points are far away as possible.

D) Naïve Bayes
Naïve bayes is a probability classifier, they classify all the features extracted from the data easily, which are which another word for naïve bayes theorem is, and it allocates the input data into described place using naïve bayes theorem.

E) Logistical Regression
Logistical regression is a powerful statistical way of predicting the output using logistic functions. It differentiates the relationship between dependant and independent variable using Cumulative logistic functions. We can even predict the earthquake.

F) Clustering algorithms

Clustering Algorithm can be easily understand by its name itself, it groups the data based on the clustering probability, each group is same to each other and different from other clusters, clustering differs from what application we apply, there are different types of clustering algorithm, say neural and deep learning.

G) Principal Component Analysis

PCA is a statistical analysis of separating a correlated variable into uncorrelated variables using orthogonal transformation. This transformation is called principle components. Some of the examples are compression, visualizing and simple learning.

IV. DATA PREDICTION

A pattern must be inferred from the input data, random input data will affect the conclusion. The working of machine learning is shown in figure 2. The machine will only recognize the same pattern to arrive at the conclusion or it may be difficult for the machine to predict the output. Statistical learning is a tool for machine learning, analyzing the data from the ground of statistics and functional analysis [16]. Then there should be large amounts of data given to the input, if the data is less, then the machine would predict the output wrong.
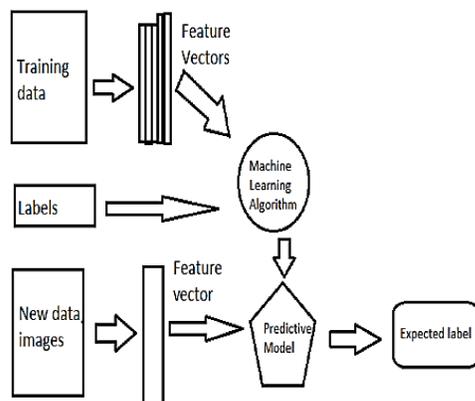

Figure 2: Working of Machine Learning

There must be enough data should be provided to train the machine, if the data is not enough it would affect the accuracy of the machine, for best machine ample amounts of data should be provided and it also hard for us to predict the behaviour of the problem mathematically which we predicted already and hence the data is taken and structured learning is applied to approximately arrive at the conclusion by the machine itself. If all the above concepts are applied correctly, then we are good to proceed.

V. SUMMARY

The research in machine learning has been very active since the last decade. Because of this there has been a lot of improvement in the efficiency and accuracy of the algorithm for the practitioner. As a result, it has become almost compulsory for Scientists and Engineers to learn Machine Learning. The wide range of examples of Machine Learning includes: Classification of Text (span in mailing system), network intrusion system (tissue classification), discovery of drug, hand writing recognition etc. The future is tend towards machine learning.It will having an increasing amount of impact in forthcoming years, it will change the world's way of thinking .Soon it will change the automobile industry, defense, telecommunication media and technology.The diagnosis in hospitals are turning pretty easy due to machine learning awareness. Driver less car has been the significant impact in the car industry.Tesla, Google and all the technological industries are changing the machines way of thinking. Alphago programme introduced by Google defeated the best chess player in the world. Normal technology is simply powered by artificial intelligence, which in turn given knowledge by machine learning techniques. The understanding is mostly derived from situations or experiments that are obtained close in time [5] There are many pitfalls in machine learning and yet it is learning to improve the error value by getting more input data and predicting the output with great accuracy, when every new input data arrives. In the coming years it is possible to solve the most pressing problems very easily and it will change the world to other level of thinking. The power of machine learning will only continue to grow in the near future, to be a part of something incredible. Machine learning helps us to predict the output and recognize patterns and create high performing useful data from the input.

REFERENCES

[1] J. Manyika, M. Chui, B. Brown, J. Bughin, R. Dobbs, C. Roxburgh, and A. Byers. Big data: The next frontier for innovation, competition, and productivity. Technical report, McKinsey Global Institute, 2011.

[2] Y. Bengio. Learning deep architectures for AI. Foundations and Trends in Machine Learning, 2:1–127, 2009

[3] Richard S. Sutton and Andrew G. Barto, "Reinforcement Learning: An Introduction", Cambridge, MA: MIT Press, 1998

[4] S. B. Kotsiantis, "Supervised Machine Learning: A Review of Classification Techniques", Informatica, Vol. 31, No. 3, pp. 249-268, 2007.

[5] G. Wallis and H. Bülthoff, "Learning to recognize objects," Trends Cogn. Sci., vol. 3, no. 1, pp. 23–31, 1999.

[6] L. Qiao, M. Sato, and H. Takeda, "Learning algorithm of environmental recognition in driving vehicle," IEEE Trans. Syst., Man, Cybern., vol. 25, no. 6, pp. 917–925, Jun. 1995.

[7] Pierre Geurts, Alexandre Irrthum, Louis Wehenkel, "Supervised learning with decision tree-based methods in computational and systems biology", Molecular BioSystems, Vol. 5, No. 12, pp. 1593-1605, 2009.

[8] Mehryar Mohri, Afshin Rostamizadeh and Ameet Talwalkar, "Foundations of Machine Learning", One Rogers Street Cambridge MA: The MIT Press, 2012.