

A Review on Machine Learning

Saima Rafat Bhandari¹, Zarina Begam K Mundargi²

¹PG Student, Dept. of CSE, SIET, Vijaypur, KA, India

²Professor, Dept. of CSE, SIET, Vijaypur, KA, India

Abstract - The machine learning is a field which allows the computers to make predictions from the past experiences. Recently machine learning has made a impressive development in processing power as well as increased in the storage capacity. Machine learning is the subfield of the Artificial Intelligence (AI). Machine Learning is derived from the content that says the learning of the data make the system to be constructed. The building block makes the computer to learn and behave more intelligently. Machine learning using the data learn, though it and implement where it is necessary. Machine learning has two techniques supervised and unsupervised where the data can be learning by the input data or by the hidden data.

Key Words: Machine Learning, Data Learning, Supervised, Unsupervised.

I. INTRODUCTION

Every useful program in computer while doing the task, learn something. The things can be learned by Memorization (where the individual facts are collected) and Generalization (gather the new facts from old facts). Here learning means identifying and understanding the input data and making wise decisions to it. There are different definitions of Machine Learning such as "Machine Learning is a field of study that gives computers the ability to learn without being explicitly programmed" by Arthur Samuel (1959). "Machine Learning Said to learn from experience with respect to some class of tasks, and a performance measure P, if (the learner's) performance at tasks in the class as measured by P, improves with experience" by Tom Mitchell (1997). "Machine learning is a method of data analysis that automates analytical model building" by a North Carolina-based analytics software developer. In Traditional programming the data and the program is given to the computer and the result will be the output. But in Machine Learning the data and the output is given to the computer and the result will be the program. Hence, Machine Learning has been widely used and become the central part of some courses such as Natural Language Processing, Computational Biology, Computer Vision, robotics etc. Machine Learning is define as a core sub area of the artificial intelligence will enables the computer to self learn without being explicitly programmed. The programs are allowed to learn, grow, change and develop according to new data exposed. There is possibility of the program to learn from the previous observations and the pattern recognition to produces the valid results. For Machine Learning various task can be consider such as classification, Regression, and cluster.

II. LITERATURE SURVEY

To study and analyze more about Machine Learning, the following literature survey has be done.

In [1] the authors present an overview on machine learning: trends and future prospects. Where it says that Machine Learning field expands from the broad field of Artificial Intelligence, which aims to copy the intelligence of humans by machines. In Machine Learning it is important to know how computers are made to learn. The learning can be a inductive interface where the statistical data is been gathered and observed. The learning can be done by supervised learning where the data is associated with each example and in unsupervised learning where data is hidden or uncovered.

In [2] the authors present the discipline of machine learning. Machine Learning is the natural swelling of the synergy of computer science and statistics. Computer science is focused on how computer program should be standard, where as statistics is focused on data to be concludes. Not only this how Machine Learning should be organized bestow on what Computational architectures and algorithms which can be used accurately in capture, store, index, retrieve and merge of the data. The machine learning is also closely related to study of human and animal learning in Psychology, Neuroscience, and related fields. Machine Learning has extended from biology, economics to control theory. Machine learning has become the successful application in computer vision, speech recognition, Bio-surveillance, robot control etc.

In [3] the authors present implementation of machine-learning classification in remote sensing: an applied review. In remote sensing literature has become the major focus for the classification of machine learning. Machine learning algorithms used for complex class of signatures, accept the input predictor for varieties of data, and assumption are not made depending on data distribution. The different algorithms such as SVM, RF and boosted DTs has shown the powerful methods for classification of remotely sensed data and these methods produces the overall precision that compared to different machine classifiers such as single DTs and k-NN.

In [4] the authors present a brief introduction into Machine Learning. The result of the machine learning is a large number of very correct and efficient algorithms that make quite easy to use for a practitioner. Machine Learning can help in the automation of the tasks and predictions where

the human effect has failed to apprehend large amounts of data. The Machine learning has the wide range of examples such as: Text classification and categorization, network intrusion detection, bioinformatics, monitoring of electric appliances, optimization of hard disk caching strategies and disk spin-down prediction, drug discovery, high energy physics particle classification, recognition of hand writing, natural scene analysis etc.

In [5] the authors present an introduction to machine learning methods. Where it speaks about the background information about the scope and application of machine learning methods for ecological analyses. Machine learning techniques are extended to analysis the options and provide the different methods and alternatives to standard statistical analyses, particularly when it is used for data exploration and pattern recognition. Not only machine learning, different tools can be for the ecological data.

In [6] the authors present a Study of the Behavior of Several Methods for Balancing Machine Learning Training Data. Learning systems usually assume that training set used for learning is balanced. In real world it is not important that the data is represented for same wide class of examples, while some represent other class. This is called as the class imbalance problem. The problem of the class imbalanced data sets occurs when one class represents a limit concept and other class represent the counterpart of that concept. There are several try for dealing the problem of class imbalance in the field of data mining and knowledge discovery in database to which machine learning is a sustained contributor.

In [7] the authors consider Machine learning: Trends, perspectives, and prospects. In today's world where the technical fields are growing rapidly, lying at the intersection of computer science and statistics and at the core of artificial intelligence and data science. Machine learning answers the question of how to build computers that improve automatically through experience. Recent progress of machine learning has been used both by the development of new learning algorithm, theory, and ongoing explosion in the availability of online data and low cost computation.

III. MACHINE LEARNING PROCESS

As there is constant evolution in the field, which has increased the importance, uses and demands of Machine Learning. As big data is has become the slang in last few years, and this has made machine learning to increase where the huge amount of the data can be consider. Machine learning has replaced the traditional programming where the data fetching has been changed and interpreted. The flow of machine learning process is showed in the fig 1.

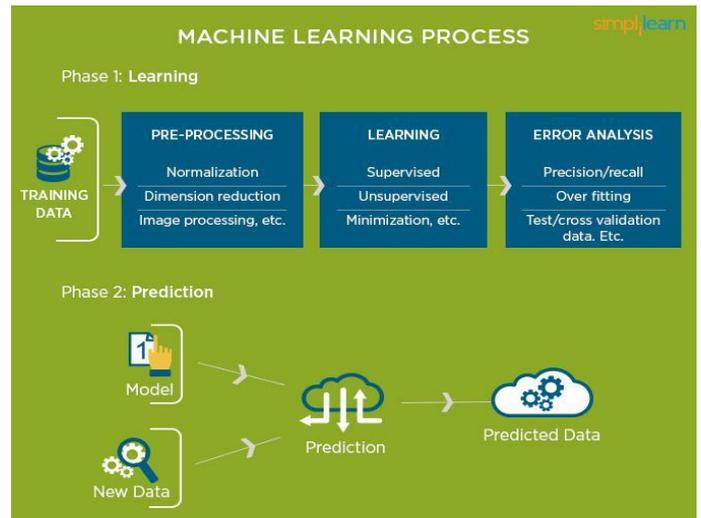


Fig 1: Machine Learning Process.

The Machine Learning process has two phases: Learning and Prediction.

In learning phase, the training data has to be pre processed by using Normalization- the process of arranging the data without any redundancy and in the simple form. Dimension reduction – the process of seeing random variables where it is divided into feature selection and feature extraction. And image processing- the process related to computer vision. Then the learning of the data is done by using supervised, unsupervised and minimization methods. After all the data processing error analysis is done by using different methods like precision/recall, over fitting, test/cross validation data etc. Second phase is prediction where the model and the new data are predicted to get the predicted data which can be used as a data for the computer. Machine learning uses two types of techniques: supervised learning, which predicts the outputs from the trained known input and output data and unsupervised learning, which find hidden patterns or structures in input data. The overview of machine learning techniques is shown the fig 2.

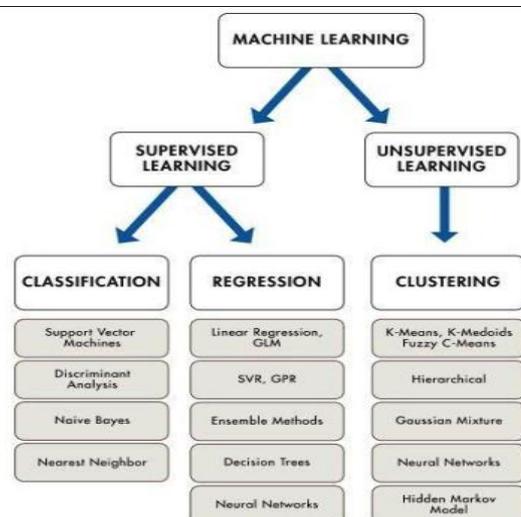


Fig. 2: Overview of Machine learning techniques.

The learning is further divided according to their techniques. The supervised learning is divided into two techniques such as classification and regression. Classification technique predicts discrete responses. It classifies the input data using different algorithms such as support vector machines, discriminate analyses, naïve bayes and nearest neighbour. Regression technique predicts continuous responses. It using different algorithms such as linear regression GLM, SVR, decision trees, neural networks. The Unsupervised learning is divided into clustering. Clustering is used for exploratory data analysis to find hidden patterns or groupings in data. It uses algorithms such as k-means, k-medoids, Fuzzy c-means, hierarchical, Gaussian mixture etc.

IV. CONCLUSION

Machine Learning methods provides the processes for the model of different data sets where the supervised learning uses tags and classify the separate data into known data classes and unsupervised Learning takes the latent variable and divide it to the group. Machine Learning has become the very active for last few years. Machine learning is being widely used according to the choice of features depending on the results and choice of distance measurement. Machine Learning has been used in several technologies application and improved it. Machine learning as made a crucial success in the artificial intelligence and made a way for huge and complex data form.

REFERENCES

1. Manish Kumar Aery, Chet Ram, "A REVIEW ON MACHINE LEARNING: TRENDS AND FUTURE PROSPECTS", 2017.
2. Tom M. Mitchell, "The Discipline of Machine Learning", 2006.
3. Aaron E. Maxwell, Timothy A. Warner & Fang, "Implementation of machine-learning classification in remote sensing: an applied review", 2018.
4. Gunnar R`atsch, "A Brief Introduction into Machine Learning", 2016.
5. Alan H. Fielding, "An introduction to machine learning methods", 1999.
6. Gustavo E. A, P. A. Batista, Ronaldo C. Prati, Maria Carolina Monard, "A Study of the Behavior of Several Methods for Balancing Machine Learning Training Data", 2015.
7. M. I. Jordan and T. M. Mitchell, "Machine learning: Trends, perspectives, and prospects", 2016.
8. Wikipedia.
9. <http://slideshare.net/>.