International Journal of Computer Science Trends and Technology (IJCST) – Volume 7 Issue 3, May - Jun 2019

RESEARCH ARTICLE OPEN ACCESS

# An Elicit Elucidation of Machine Learning &Techniques

Preeti Karma [1], Dr. Harmeet [2]

Reasearch Scholar [1]

Sant Baba Bagh Singh University

India

**ABSTRACT**

Machine learning is the act of designing computer algorithm in such a manner so that computer can solve the real world algorithms as similar to the human brain. This study focuses on the basic concepts of machine learning and various types of algorithm used to develop the problem solving techniques. Moreover, this study compares the difference between various machine learning algorithms.

***Keywords:-*** Machine Learning, types of ML, Comparison of Machine Learning

## I. INTRODUCTION

Machine learning (ML) is a category of algorithm that allows software applications to become more accurate in predicting outcomes without being explicitly programmed. The basic premise of machine learning is to build algorithms that can receive input data and use statistical analysis to predict an output while updating outputs as new data becomes available.

A computer program is said to learn from experience E with respect to some class of task T and performance measure P, if its performance at the task in T, as measured by P. improves with experience E.

A machine can be considered to learn if it is able to gather experience by doing a certain task and improve its performance in doing the similar task in the future. Past experience refers to the past data related to the task. The data inputted to the machine collected from some source. In the context of playing checkers, E represents the experience of playing the game, T represents the task of playing the checkers, and P is the performance measure indicate by the percentage of games won by the player. The same mapping can be applied for any other machine learning problem, for example image classification problem. In context of image classification, E represents the past data with images having labels or assigned classes, T is the task of assigning class to new, unlablled images, and P is the performance measure by the percentage of the images correctly classified.

## II CATEGORIES OF MACHINE LEARNING

Machine learning is the process in which computer tries to learn so that it can be able to solve the problems as the human brain does. There are different types of machine learning as mentioned in the figure 1.

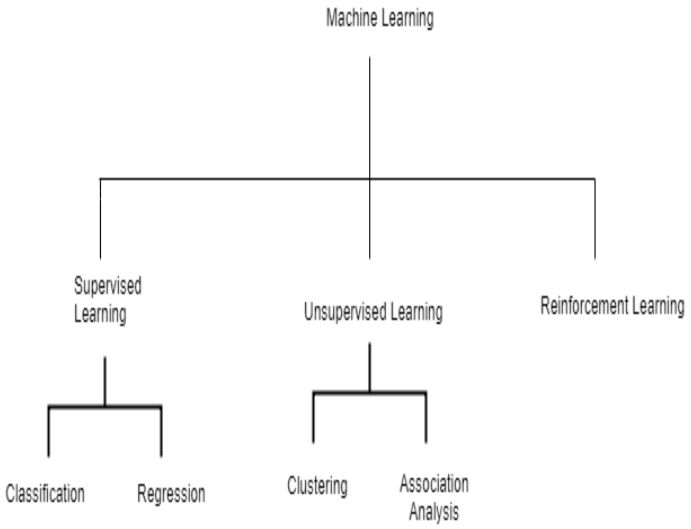
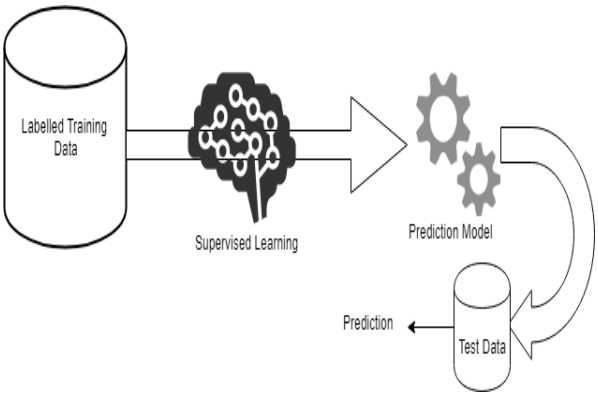



Figure 1: Machine Learning Techniques

***2.1 Supervised Learning:*** Supervised learning is also termed as predictive learning. A machine predicts the class of unknown objects based on prior class-related information of similar objects. The major motivation of supervised learning is to learn from previous experiences, or past available information. For example, a machine is

getting images of different objects as input and the task is to segregate the images by either shape or color of the object. If it is by shape, the images which are of round-shaped objects need to be separated from images of triangular-shaped objects etc. If the segregation needs to happen based on the color, images of blue objects need to be separated from images of green objects. The process of supervised learning is demonstrate by figure 2

Figure 2: Process of Supervised Learning Technique




In supervised learning a machine needs information to be provided to it. The basic input, or the experience in the paradigm or machine learning is given in the form of training data. Training data is the past information on a specific task. In the context of image segregation problem, will have past data on different aspects or feature on a number of images, along with the tag on whether the image is round or triangular, or blue or green in color. The tag is called "label" and we say that the training data is labeled in case of supervised learning. Supervised learning is categorized in two areas:

***2.1.1*** *Classification:* Classification is the process of assigning labels to the test data. The whole problem resolves around assigning a label or category or class to a test data based on the label or category or class information that is imparted by the training data. Since the target objective is to assign a class label, this type of problem as classification problem.

***2.1.2*** *Regression:* In linear regression. The objective is to predict numerical features like real estate or stock price, temperature, marks in an examination, sales revenue etc. the underlying predictor variable and the target variable are continues in nature. In case of linear regression, a straight line relationship is fitted between the predictor variables and the target variables, using the statistical concept of least square method. In case of simple regression, there is only one predictor variable whereas in case of multiple regression, multiple predictor variables can be included in the model.

### *2.2 Unsupervised Learning*

In unsupervised learning, there is no labeled training data to learn from and no prediction to be made. In unsupervised learning, the objective is to take a dataset as input and try to find natural groupings or patterns within the data elements or records. Therefore, unsupervised learning often termed as descriptive model and the process of unsupervised learning is referred as pattern discovery or knowledge discovery. One critical application of unsupervised learning is customized segmentation. Figure 3 shows the process of unsupervised learning.

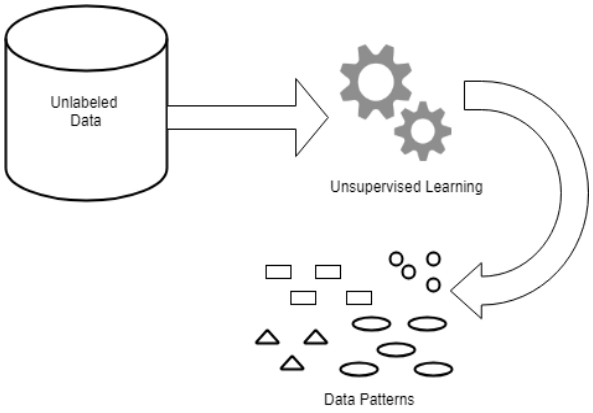



Figure 3:Process of Unsupervised Learning

Unsupervised learning is categorized in two areas:

***2.2.1*** *Clustering:* Clustering is the main type of unsupervised learning. It intends to group or organize similar objects together. For that reason, object belongs to the same cluster are quite similar to each other whereas object belongs to different clusters are quite dissimilar. Hence, the objective of the clustering is to discover the intrinsic grouping of unlabeled data and form clusters. Figure 3 shows the typical clustering scenario inside the data elements

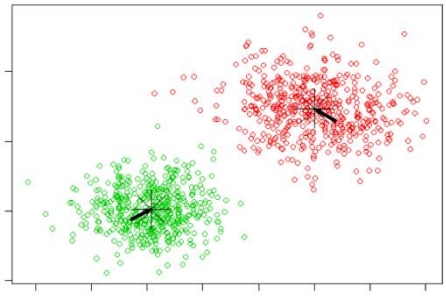


Figure 4 Clustering of data elements

*2.2.2 Association analysis:* In association analysis, the association between the data elements is identified. For example, from the past transaction data in a grocery store, it may be observe that the most of the customers who have bought item A, have also bought item B and item C or at least one of them. This means that there is strong association of the event purchase of item A with the event purchase of item B or purchase of item C.

### *2.3* Reinforcement Learning

Reinforcement learning is used when there is no idea about the class or label of a particular data. The model has to do the classification by itself. Self-driving car is the example of reinforcement learning. The critical information which it needs to take care of are speed and speed limit in different road segments, traffic conditions, road conditions, weather conditions etc. The task that have to be taken care of are start/stop, accelerate/decelerate, turn to left/right, etc.

## III. COMPARISON OF TYPES OF MACHINE LEARNING

| Supervised Learning | Unsupervised Learning | Reinforcement Learning |
|---|---|---|
| This type of machine learning is used when you know how to classify a given data, or in other words classes or labels are available | This type of learning is used when there is no idea about the class or label of a particular data. The model has to find pattern in the data. | This type of learning is used when there is no idea about the class or label of a particular data. The model has to do classification- it will get rewarded if the classification is correct, else get punished |
| Labeled training data is needed. Model is built based on training data. | Any unknown and unlabeled data set is given to the model as input and records are grouped | The model learns and updates itself through reward/punishment. |

| | | |
|---|---|---|
| The model performance can be evaluated based on how many misclassification have been done based on the comparison between predicted and actual values. | Difficult to measure whether the model did something useful or interesting. Homogeneity of records grouped together is the only measure. | Model is evaluated by means of the reward function after it had some time to learn. |
| There are two types of supervised learning problems-classification and regression. | There are two types of unsupervised learning problem-clustering and association. | No such types. |
| Simplest one to understand | More difficult to understand and implement than supervised learning | Most complex to understand and apply. |
| Standard algorithms include<br>1.Naïve Bayes<br><br>2.K-nearest neighbor(kNN)<br><br>3.Decision Tree<br><br>4.Linear Regression<br><br>5.Logistic Regression<br><br>6.Support Vector Machine(SVM) | Standard Algorithms include:<br>1.K-means<br><br>2.Principal Component Analysis(PCA)<br><br>3.Self-Organizing Map(SOM)<br><br>4.Apriori Algorithm<br><br>5.DBSCAN | Standard Algorithms include<br>1.Q-Learning<br><br>2.Sarsa |
| Practical applications include<br><br>1.Hand writing recognition<br><br>2.Stock market prediction<br><br>3.Disease prediction<br><br>4.Fraud Detection | Practical Applications include:<br><br>1.Market basket analysis.<br><br>2.Recommender systems<br><br>3.Customer Segmentation | Practical Applications include:<br><br>1.Self-driving cars<br><br>2.Intelligent robots<br><br>3.AlpahGo Zero |

Table1: Comparison of Types of Machine Learning Algorithms

## IV. CONCLUSION

In nutshell, we can say that machine is the field of study in which we can apply some mathematical models to the computer system so that they can solve real time problems in more effective manner as done by human being. Machine learning includes different set of algorithms such as supervised learning, unsupervised learning and reinforcement learning algorithms that can be applied to different data set to capture results of problem.

## REFERENCES

[1] Robin R. Murphy (2019) Computer vision and machine learning in science fiction. Science Robotics 22 May 2019:Vol. 4, Issue 30, eaax7421

[2] ROBIN R. MURPHY(2019) .Robot learning in science fiction.SCIENCE ROBOTICS30 JAN 2019

[3] HackerEarth. (2018). Practical Guide to Clustering Algorithms & Evaluation in R Tutorials & Notes | Machine Learning | HackerEarth.

[4] Dataaspirant. (2018). classification and clustering algorithms. [online]

Available:http://dataaspirant.com/2016/09/24/classification-clustering-alogrithms

[5] Analytics, B. and clustering, A. (2018). An Introduction to Clustering & different methods of clustering. [online] Analytics Vidhya. Available at: https://www.analyticsvidhya.com/blog/2016/11/an-introduction-to-clustering-and-different-methods-of-clustering.

[6] Kajaree Das1, Rabi Narayan Behera (2017)A Survey on Machine Learning: Concept, Algorithms and Applications. International Journal of Innovative Research in Computer and Communication Engineering, Vol. 5, Issue 2, February 2017

[7] Zhan, J., Chen, J. and Xing, J. (2017). Research advance of clustering ensemble algorithm. 2017 International Conference on Wavelet Analysis and Pattern Recognition (ICWAPR).

[8] Erik Brynjolfsson,Tom Mitchell (2017).What can machine learning do? Workforce implications.SCIENCE22 DEC 2017 : 1530-1534

[9] Abdi, Asad. (2016). Three types of Machine Learning Algorithms. 10.13140/RG.2.2.26209.10088.

[10]. D., N. and Vidyavathi, B. (2015). A Survey on Applications of Data Mining using Clustering Techniques. International Journal of Computer Applications, 126(2), pp.7-12.

[11] M. I. Jordan and T. M. Mitchell (2015) Machine learning: Trends, perspectives, and prospects. Science 349, 255 (2015), DOI: 10.1126/science.aaa8415.

[12] M, N. (2012). A Comprehensive Overview of Clustering Algorithms in Pattern Recognition. IOSR Journal of Computer Engineering, 4(6), pp.23-30.

[13] COMPARISON OF VARIOUS CLASSIFICATION ALGORITHMS ON IRIS DATASETS USING WEKA. (2014). International Journal of Advance Engineering and Research Development, 1(01).

[14] Greenlaw, R. and Kantabutra, S. (2013). Survey of Clustering. International Journal of Information Retrieval Research, 3(2), pp.1-29.

[15] Ackermann, M., Märtens, M., Raupach, C., Swierkot, K., Lammersen, C. and Sohler, C. (2012). StreamKM++. Journal of Experimental Algorithmics, 17(1), pp.2.1.

[16] Ayodele, Taiwo. (2010 )Types of Machine Learning Algorithms 10.5772/9385

[17] Ayodele, Taiwo. (2010). Types of Machine Learning Algorithms. 10.5772/9385.

[18] Ankerst, M., Breunig, M., Kriegel, H. and Sander, J. (1999). OPTICS. ACM SIGMOD Record, 28(2), pp.49-60.