Machine Learning and its various Concepts

[1]Alisha Sikri, [2] Surjeet Dalal, [3]N.P Singh
PhD Scholar, SRM University, Computer Science Department[1]
Associate Professor, SRM University, Computer Science Department[2]
Professor and Dean, MDI Gurugram, Information Management[3]
[1]*alisha.sikri92@gmail.com,*[2]*profsurjeetdalal@gmail.com,*[3]*knpsingh@mdi.ac.in*

Abstract: - Machine learning can be defined as field of artificial Intelligence and it has been extremely active the last few years. The types of various machine learning algorithms are decision trees, naïve bayes, C4.5 or CART which are divided into categories like supervised, unsupervised and reinforcement learning. The other techniques which can be used in combination to machine learning are ensemble which includes bagging, boosting and stacking, classification technique, feature selection etc. This paper is divided into 7 sections. The First Section is the introduction part which gives a brief introduction of the topics described. Section-2 gives the introduction about Machine Learning, followed by section-3 which explains Classification of Machine Learning. Section-4 gives an introduction about Feature Selection followed by section-5 which introduces Ensemble Learning. The next Section is the literature review part. The paper is concluded by section-7 which summarizes the conclusion and the future scope.

Keywords: Machine Learning, Medical Domain, E-commerce

1. INTRODUCTION

Enterprises accumulates large amount of data. The data could be online generated, offline generated, qualitative, quantitative, graphic, audio or video in nature. Companies are using statistical tools, analytical tools, machine learning algorithms, deep learning algorithms, natural language processing algorithms, etc. for the analysis of this data. Machine Learning algorithms are a sub-field of Computer Science and have overlapping properties with artificial intelligence and prove to be useful on any kind of e-commerce data. Machine learning algorithms can be divided into different categories such as supervised, unsupervised and reinforcement Learning (Khall.2015) (Hu et al.2018) (Pandey et al. 2015) (Muhammad & Yan. 2015) (Gmyzin. 2017) are used by companies in problem relating to classification, clustering, association rule mining, prediction/ prescription, and estimation. Due to all the above mentioned factors the utility of the The machine learning algorithms increase every day. In the section below, we will discuss the training details of the machine and many other factors related to it.

2. MACHINE LEARNING: -

Machine training can be called a model that may involve learning from previous experience that enhances performance in the future. Machine research has been very active in recent years (Ratsch). Learning can also be defined as a revision or improvement of an automated "past experience" algorithm without external help to people (Das and Behera 2017). Machine training also shows a number of statistical algorithms that can be used for mining exploration and forecasting.This includes a variety of techniques, such as decision making, trees, regression, aids, images, images, and other mining techniques. During the period of 1995-2000, many attempts were made to improve the practice of guardian algorithms through the program (Upadhyaya 2011). Machine learning can be defined as:-

"It is said that computer software learns from experience E related to several classes of tasks T and P performance measure.It is assumed that if its performance in T measured with P with experience E It uses a computer for simulating people's biases and allows the computer to identify and acquire real world knowledge and improve sub-deployment. A specific task pane based on this new knowledge (Portugal et al.2017).

3. CLASSIFICATION OF MACHINE LEARNING ALGORITHMS: -

Machine training is a technique used to train machines for efficient data management, with various machine learning algorithms. These algorithms are used for various purposes such as data mining, image manipulation, analysis, predictions, and more. (Dey 2016) (Hormozi et al., 2012) (Kotsiantis 2007) (Nasteski). Also, training machines can be classified into the following categories and they have: -
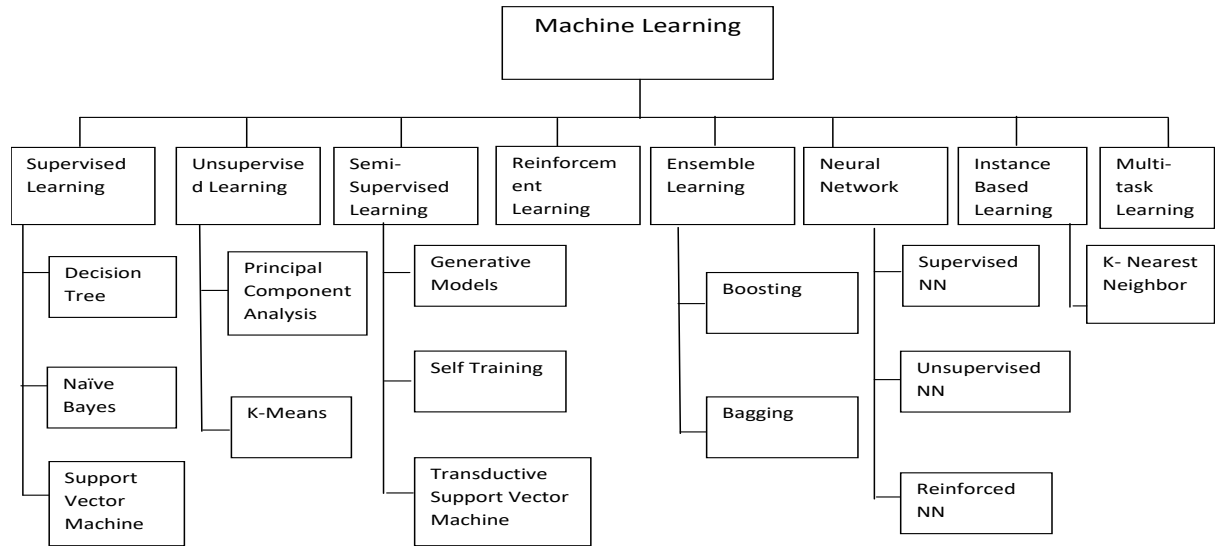
Fig 3. Classification of Machine Learning Algorithms (Dey 2016)

Table 2: Differentiation between various types of ML Algorithms

S.NO	SUPERVISED LEARNING	UNSUPERVISED LEARNING	SEMI-SUPERVISED LEARNING	REINFORCEMENT LEARNING
1.	It is provided with the training data and the correct answers (Portugal et al.2017)	It is only provided with the training set (Portugal et al.2017)	This type of learning is provided with the training set and the missing information (Portugal et al.2017)	It is provided with the feedback (Portugal et al.2017)
2.	These type of ML Algorithm learn with the training set (Portugal et al.2017)	These ML algorithms are presented with some data or information about the real world &then they have to learn from that data on their own (Portugal et al.2017)	These ML algorithms occur when they work with a training set with missing information but still need to learn from it (Portugal et al.2017)	These type of ML algorithms occur when the algorithms attempt to learn based on some external feedback given either by a thinking entity, or some environment (Portugal et al.2017)
3.	Priori is necessary (Dasgupta and Nath 2016)	Priori is not necessary (Dasgupta and Nath 2016)		Priori is required (Dasgupta and Nath 2016)
4.	It will also produce same output for a specific input (Dasgupta and Nath 2016)	It will produce different outputs on each run for a Specific input (Dasgupta and Nath 2016)		The output changes if the environment does not remain same for a specific input (Dasgupta and Nath 2016)

4. FEATURE SELECTION: -

With increased data content, large-scale data analysis has become a major challenge for researchers and engineers to learn machine as well as in data mining. The solution to the problem in functional selection can be classified into two groups: functional withdrawal and functional selection (Cao et al., 2007). So feature picks can be considered an effective way to provide a way to solve problems by removing inappropriate and unnecessary data. It reduces time, calculates, improves training accuracy and facilitates better understanding of learning patterns or data. The functional selection method can be evaluated by the machine learning model.The various commonly used machine learning models are Naïve Bayes, KNN, C4.5, SVM, BP-NN, K-means, Hierarchical Clustering, Density-based clustering (Cai et

al.2017) (Miao and Niu 2018).Function selection can be defined as "the process of obtaining the subset of the original set of the function according to the specific criteria for selecting functions that select relevant data" (Baumann et al., 2018).

It is considered an important undertaking in the field of inappropriate machine learning that is used as part of the training process of different forecasting systems, which can lead to increased spending and Time of system process. In its general form, the FSP can be used to learn a problem

Data. It can be defined as a sample of the sample data (x1, y1), ..., (xl, yl) where xiRn and yi∈ R (or yi∈ {±} 1} in the case of partitioning problems .) Select a subset of math characteristics (m <n) that achieve the lowest possible error in the evolution of the yi variable (Sanz et al., 2018).
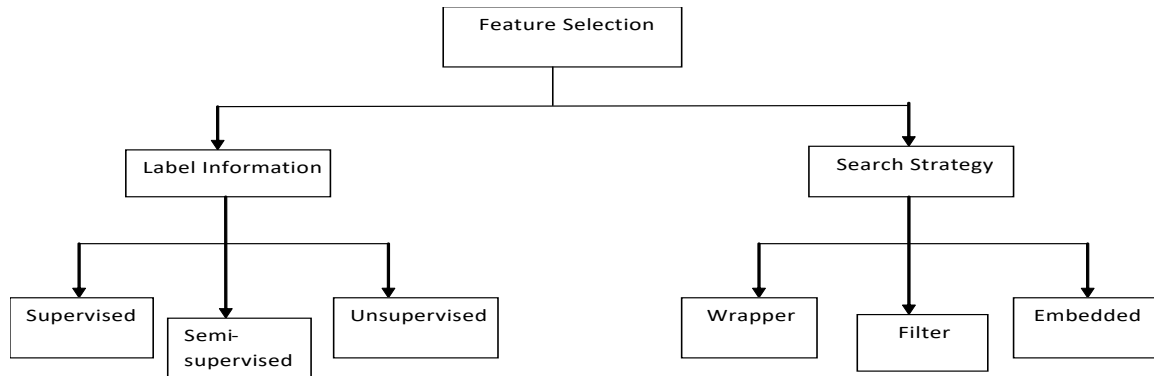

Fig 2. Feature Selection Category (Miao et al.2016) (Zhang et al. 2018)

V. ENSEMBLE: -Entire learning is the main topic of machine learning in recent years (Zhang et al., 2018). Machine learning techniques can be learned in a number of ways, such as logging and forestry accidentally (January 2006), diversion of the forest (Joshi and Srivastava 2014) and Random Subspace (Dahiya et al.). 2015).

The set integrates a series of k samples called the basic classifiers - M1, M2, ... Mk, to create an improved component class template. To create the set of training k D1, D2, ... DK, the D dataset is used when Di (1≤i≤ k-1) is used to create the Mi emulator. A joint decision on the handling of a new and unseen case will be conducted by elections. Elections may be weighty or not weighty (Dahiya et al., 2017).

The definition of assembly in the context of machine learning is "the technique of assembly, where the machine learning system is created by the model of a consistent, productive individual, combined with a merger strategy to create One Response (Huang et al., 2009) (Tiwari and Prakash 2014) (Koutanaei et al., 2015) (Patil et al., 2016) (Guinness and Agwala, 2010).

6. LITERATURE REVIEW

Portugal et al. (2017) "Using machine learning algorithms in transmission systems: systematic model. The article also shows that the MLP neurotransmitters are better than other distributors in the

monitoring". The conclusion of this article is that the ML algorithm is one of the most important and progressive steps in terms of system arbitration because it allows the computer to learn customer-based ideas as well as personalized recommendations. It can also be seen that uncontrolled and untreated training is also being taken into account. Symbolic clusters, as well as groups and SVM (vector vectors) are among the most used. This study has contributed a great deal of research into the expert system and intelligent systems.

Cai et al. (2017) "Select Training Machine Functionality: New Perspective". This article discusses some common measures to evaluate samples or data. This article seeks to explore the methodology to select the semi-oversized and semi-controlled controls that are widely performed and are used in machine learning such as classification, classification and preparation. The authors also try to describe a feature selection framework and describe future challenges related to feature selection.

Dahiya et al. (2015) "Loan Model Using Hybrid Engine Training Technique". Experimental results from the report show that the hybrid modeling model is more than the sample

same model and hybrid. The MLP rating best suited among 83% of all classified ads, and the MLP-MLP

hybrid is best amongst other high-end ads with the highest accuracy of 86%. It can be observed that more breeding experiments can be performed using machine learning techniques, such as SVM and some new tree dividers. Also, multiple methods of Matching such as stacking, amplification, for hybridization as well as to achieve better accuracy and minimize errors in distribution.

7. CONCLUSIONS AND SCOPE OF THE FUTURE

From the above study, it can be assumed that machine learning has found a wide variety of applications in everyday life. Other machine learning algorithms can be studied and in the future the learning style of the learning machine can be evaluated in the context of the selection of the features and techniques of the set. Also, these algorithms can be applied to programs such as R or Weka

REFERENCES

[1] Khall.KM, Aziz.A, Nazmy. TT and Salem.AB. (2015) "Machine Learning Algorithms for Multi-Agent Systems". IPAC'15, Proceedings of the International Conference on Intelligent Information Processing, Security and Advanced Communication.doi: 10.1145/2816839.2816925.

[2] Hu. Y, Da. Q, Zeng. A, Yu. Y and Xu.Y. (2018) "Reinforcement Learning to Rank in E-Commerce Search Engine: Formalization, Analysis, and Application" .doi: 10.1145/3219819.3219846.

[3] Hormozi. H, Hormozi. E and Nohooji.H.R. (2012) "The Classification of the Applicable Machine Learning Methods in Robot Manipulators", International Journal of Machine Learning and Computing, Vol.2, No. 5.

[4] Dey. A. (2016) "Machine Learning Algorithms: A Review", International Journal of Computer Science and Information Technologies. Vol. 7 (3), pp: 1174-1179.

[5] Portugal. I, Alencar. P and Cowan.D. (2017) "The Use of Machine Learning Algorithms in Recommender Systems: A Systematic Review", Expert Systems with Applications, Vol. 97, pp: 205-227. doi: 10.1016/j.eswa.2017.12.020

[6] Kotsiantis. S. B. (2007) "Supervised Machine Learning: A review of Classification Techniques", Informatics 31, pp: 249-268.

[7] Nasteski. V. "An Overview of the Supervised Machine Learning Methods", Vol. 160. doi: 10.20544/HORIZONS.B.04.1.17.P05.

[8] Dasgupta. A, Nath. A. (2016) "Classification of Machine Learning Algorithms", International Journal of Innovative Research in Advanced Engineering (IJIRAE), Vol. 3, Issue 03

[9] Zhang. R, Nie.F, Li.X and Wei.X. (2018). "Feature Selection with Multi-view Data: A Survey", Information Fusion. doi:10.1016/j.inffus.2018.11.019

[10] Wang. C.W. (2006) "New Ensemble Machine Learning Method for Classification and Prediction on Gene Expression Data", International Conference of the IEEE Engineering in Medicine and Biology Society. doi: 10.1109/IEMBS.2006.259893.

[11] Dahiya. S, Handa.S.S and Singh. N.P. (2015) "Credit Scoring Using Ensemble of Various Classifiers on Reduced Feature Set", Vol. 43, No.4, pp: 163-174. doi: 10.5937/industrija43-8211

[12] Joshi.N and Srivastava.S. (2014) "Improving Classification Accuracy Using Ensemble Learning Technique", International Journal of Computer Science and Mobile Computing, Vol.3, Issue.5, pp : 727-732.

[13] Dahiya.S, Handa.S.S and Singh.N.P. (2017) "A Feature Selection enabled hybrid-bagging algorithm for credit risk evaluation", WILEY Expert Systems, Vol. 34, Issue 06. doi: 10.1111/exsy.12217

[14] Tiwari. A and Prakash. A.(2014). "Improving classification of J48 algorithm using bagging, boosting and blending ensemble methods on SONAR dataset using WEKA", International Journal of Engineering and Technical Research (IJETR), Volume-2, Issue-9.

[15] Patil. P, Aghav. J.V and Sareen. V. (2016) "Overview of Classification Algorithms and Ensemble Methods in Personal Credit Scoring", International Journal of Computer Science and Technology (IJCST), Vol. 7, Issue 2.

[16] Koutanaei. F.N, Sajedi. H and Khanbabaei. M. (2015) "A hybrid data Mining model of feature selection algorithms and ensemble learning classifiers for credit scoring", Journal of Retailing and Consumer Services 27, pp: 11-23. doi: 10.1016/j.jretconser.2015.07.003.

[17] Gandhi. P and Aggarwal. V. (2010) "Ensemble Hybrid Logit Model", Journal of Machine Learning Research x, pp: 1-18