A Review of Machine Learning Algorithms and its Applications

Vivekananda Journal of Research
January - June 2022, Vol. 12, Issue 1, 132-145
ISSN 2319-8702(Print)
ISSN 2456-7574(Online)
Peer Reviewed Refereed Journal

https://vips.edu/journal/

Mitanshi Rastogi*
Neha Goel**

Abstract

Machine Learning is a science that was developed in artificial intelligence. Machine Learning and its algorithms are used for various purposes like pattern recognition, image processing, predictive analysis, deep learning etc. Machine Learning works on the principle of determining the most appropriate model for the data. The research work highlights the work done by other authors in similar fields. A conclusive study of the work done has been discussed. The contribution of machine learning techniques in different sectors like computer vision, healthcare, construction, etc. along with their purpose and applications is presented. The idea behind this study is to spread awareness about machine learning algorithms and its application areas.

Keywords: *Machine Learning, supervised, semi-supervised, reinforcement.*

INTRODUCTION

Today, we are living in the era of data, where anything or everything is related with each other. Most of the data is documented and recorded on digital platforms. The data is increasing tremendously. To manage and analyze the data more effectively in a specific area, machine learning techniques and algorithms are implemented on the data.

The focus is to identify the best suitable machine learning algorithm for the available data. Many organizations with enormous amount of data are using machine learning algorithm to retrieve the valuable information. The purpose of examining the data gathered in diverse

* Research Scholar, K.R Mangalam University. Email: rmitanshi@gmail.com
** Assistant Professor, VIPS, Delhi. Email: nehagoel123@gmail.com

areas changes with the different business domains such as stock market, military, image processing, speech recognition, handwriting recognition, education, and construction.

CLASSIFICATION OF MACHINE LEARNING ALGORITHMS:

The machine learning algorithms are broadly categorized into four parts: Supervised Learning, Unsupervised Learning, Semi-supervised Learning and Reinforcement Learning.

a) **Supervised Learning:** An algorithm where external assistance is required. The given input data is further categorized into train dataset and test dataset (labeled data). The outcome variable is forecast from the train dataset. The supervised learning is further classified into classification and regression learning as shown in Fig1.1. Some of the most popular classification and regression techniques are Decision Tree, Naive Bayes, Support Vector Machine.

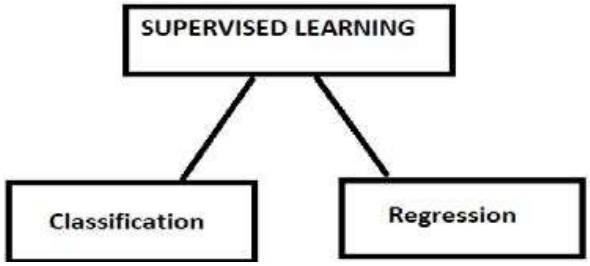

Fig 1.1: Types of Supervised Learning

b) **Unsupervised Learning:** An output variable is not specified or stated in unsupervised learning. When novel data is given, it emphasizes on the existing features to identify the data (unlabeled data). Fig 1.2 shows the types of Unsupervised Learning. As shown in the fig, the unsupervised learning is further classified into clustering and association techniques. Some of the popular techniques are: K-Means Clustering, Density based clustering, Dimensionality reduction- Principal Component Analysis, K-nearest neighbors.

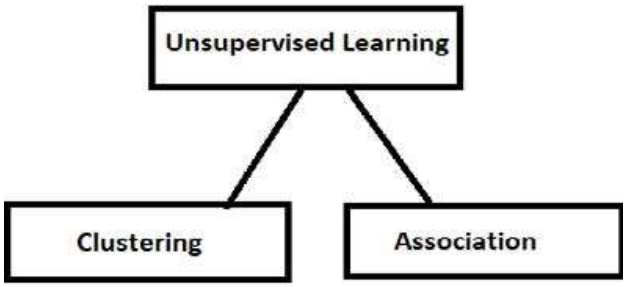

Fig 1.2: Types of Unsupervised Learning

c) **Semi-supervised Learning:** This learning method can be defined as a composition of supervised and unsupervised learning as shown in fig 1.3. The objective of this learning model is to produce a better output for forecast than that provided already with labeled data. Semi-supervised learning can be implemented both on labeled and unlabeled data. Few popular techniques are: Self-training, Generative Models.

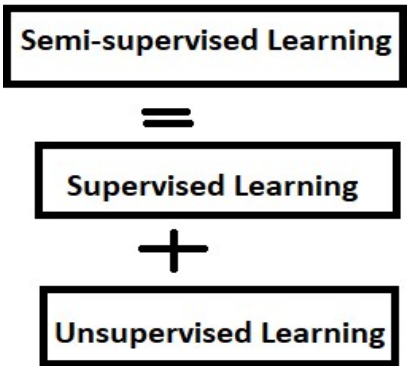

Fig 1.3: Semi-supervised Learning

d) **Reinforcement Learning:** In this learning model works on environment-driven approach. The agent has no prior information or knowledge which steps to be taken till been in a particular scenario. Positive and negative rewards are given depending on the path chosen by the agent.

LITERATURE REVIEW

Annina Simon [2015] demonstrated deep learning, their applications and its relationship with machine learning, big data, artificial intelligence. Many organizations are exploring and working on this area as it has evolved as one of the most leading fields of technology in modern era.

Murat Gunduz et al [2015] demonstrated on basis of accuracy, neural network is more appropriate than multiple regression analysis.

Ayon Dey [2016] demonstrated different types of machine learnings, their advantages and application areas. The benefit of applying machine learning is that when an algorithm acquires a knowledge of what to do with data, then it can operate individually.

Tanvi Sharma et al [2017] briefed that heart diseases are very crucial and costs human lives if not diagnosed in early stages. Based on human health conditions and required parameters the machine learning can help in predicting the diagnosis of heart disease from patient's data.

Ozer Celik et al [2018] explained about machine learning and their application areas. Machine Learning works on the phenomenon of identifying the suitable model for the latest data along with the existing available set of data.

Heejae Ahn et al [2018] explained the scope and importance of machine learning in different application areas to enhance the safety measures and minimize the waste of labour at construction sites.

G. Mahalakshmi et al [2019] demonstrated that the machine learning technologies can be helpful in minimizing the cost and effort by early estimating the costs of highway projects in India. There are many parameters that play a vital role in predicting cost at an early stage. Raw material, time duration, soil and topographical conditions, assembly and analysis of data is done.

Ahmed Gondia et al [2020] briefed that machine learning and its algorithms plays an important role and are very efficient in minimizing delay in risk prediction. Decision Tree and Naïve Bayes classification algorithms were selected for decision making in forecasting delay risk prediction.

Amy X. Du et al [2020] demonstrated the relevance of machine learning technologies in forecasting clinical end-result in dermatology with a variety of skin conditions. The author also states when and where which machine leaning should be applied to produce more accurate results.

Kevin Fauvel et al [2020] briefed that individual GPS stations and seismometers are not sufficient for predicting both medium and large earthquakes. Earthquakes are very disastrous. If we predict earthquake at an early stage, we can save lives of many people.

Chhaya Gupta et al [2020] demonstrated that breast cancer is one of the most common and dangerous disease these days which costs many lives. Machine learning and its techniques are very helpful in forecasting breast cancer at an early stage. So far, extreme machine learning gives better results with accuracy and time.

Sana Tayefeh Hashemi et al [2020] explained that price estimation and forecasting is necessary in construction projects. Construction price forecasting minimizes the risk, human effort, extra costs involved in construction projects like highway, roadway, hydroelectric and power plants and many more. Machine learning techniques play a vital in cost estimation and prediction and suggests which machine learning technique is suitable in a particular scenario.

Neha Sharma et al [2021] demonstrated the importance of machine learning in

different application areas are in trend these days. Further in detail discussed about how deep learning came into existence and their application areas like computer vision, financial sector, agriculture, cybersecurity, healthcare and many more. Also, deep learning and machine learning played a crucial role during Covid-19.

Iqbal H. Sarker [2021] demonstrated the importance of machine learning and techniques in real world applications and research. The author helps in finding suitable model for the situation.

Table 1: Literature Survey of Machine Learning

S. No	Author's Name	Year of publication	Title	Objective	Tools/Tech	Discussion/ Conclusion	Limitations	Future Scopes
1	Annina Simon	2015	An overview of Machine Learning & its Applications	To make the researchers more aware in the area of deep learning	Machine learning, Deep Learning, Artificial intelligence	Many organizations are working on real-world appliactions using deep learning.	Deep learning is difficult to understand for beginners	Further advancement in the area of deep learning and machine learning
2	Murat Gunduz, Haci Bayram SAHIN	2015	An early cost estimation model for hydroelectric power plant projects using neural networks and multiple regression analysis	To identify the most appropriate model for early cost estimation in HEPP	NN (Neural Network), MRA (Multiple regression analysis)	Neural network gives more accurate result as compared to multiple regression analysis.		
3	Ayon Dey	2016	Machine Learning Algorithms: A Review	To identify different machine learning algorithms in diverse application areas	numerous machine learning techniques	Depicts the importance of machine leaning in our day-to-day life in different areas.		

4	Tanvi Sharma, Sahil Verma, Kavita	2017	Prediction of heart disease using cleveland dataset: A machine learning approach	To identify which ML technology is used to forecast heart disease in early stage	DT, MARS, RF, TMGA	Decision Tree is gives better results with respect to accuracy and time		More research needs to be done in area of health sector
5	Ozer Celik	2018	A research on Machine Learning and its applications	To find the best model for the new data along with the previous data	various machine learning algorithms	Machines successful in the jobs that cannot be done by human will affect lots of business sectors and humans.		
6	Heejae Ahn, Dongmin Lee et. Al.	2018	Applications of Machine Learning for contruction site	To identify where machine learning algorithms required in different application areas of contruction sites	ANN, SVM , Random Forest , GA	By using ML , reduction in waste of labor and safety risk	effective-ness not verified	Appropriate model for the potiential problems and effectiveness of the model would be defined
7	G. Mahalakshmi and C. Rajasekaran	2019	Early cost estimation of highway projects in India using Artificial Neural network	To develop a model for early cost estimation of Highway projects in India	ANN	multi percep-tron network with back-propagation algorithm is capable of predicting construc-tion cost of highway with reasonable accuracy		

8	Ahmed Gondia, Wael El-Dakhani , Ayman H.Nassar	2020	Machine Learning algorithms for Construction Projects Delay Risk Prediction	To identify which Machine learning technology gives most accurate result in delay risk prediction of construction projects	Decision Tree and Naïve Bayesian Classification	Bayesian Model provides a better predictive performance.	performance of the model is dependent on the dataset available	external project variable influence should also be studied
9	Amy X. Du, Sepideh Eman and Robert Gniadecki	2020	Review of Machine Learning in Predicting Dermatological outcomes	To identify the relevance of ML in forecasting clinical end-result in dermatalogy with a variety of skin conditions	Naïve Bayesian, ANN, Support vector machine, fuzzy logic	when and where Machine learning technology need to be applied to retrieve the more accurate results in dermatology	to identify how much data needed to apply ML algorithms and the human operators are not capable enough to brief how ML technology draw their conclusion	To know if ML predicted end-results are genuinely required for clinical trials
10	Kevin Fauvel et al	2020	A distributed Multi-Sensor Machine Learning Approach to earthquake Early Warning	To predict and detect medium and high on Richter scale earthquakes	GPS, Seisometers, DMSEEW	Based on the performance analysis DMSEEW gives more accurate results than individual GPS stations and seismometers		Aim to apply DMSEEW on different seismic network through collaboration with NIED

11	Chhaya Gupta, Nasib Singh Gill	2020	Machine Learning Techniques and Extreme learning machine for early breast cancer detection	To predict and detect breast cancer in early stage	Decision Tree, K-nearest neighbor, Support Vector Machine, Random Forest, Extreme Learning Machine	PCA(Principal Component Analysis) with Extreme Learning Machine gives better result on the basis of accuracy and time comparison		to trace the benign and malignant cells in early span of time
12	Sanaz Tayefeh Hashemi , Omid , Harleen	2020	Cost estimation and Prediction in construction projects : a systematic review on ML techniques	To identify various machine learning algo-rithms and analytics techniques to forecast the cost estimation of con-struction oriented projects.	ANN, SVM , Fuzzy NN	Suggests which machine learning technique is suitable in a particular scenario		
13	Neha Sharma, Reecha Sharma, Neeru Jindal	2021	Machine Learning Deep Learn-ing Appli-cations-A vision	The aim is to give the under-standing of machine learning algorithms in diverse areas.	SVM, NLP,CNN, Logistic regression, deep learning	suggests different modeling techniques of ML and deep learning in covid-19 scenario		Deep learning is trending with latest applications in diverse segments in the future.

14	Iqbal H. Sarker	2021	Machine Learning: Algorithms, Real-world applications and Research directions	The purpose is to identify various machine learning algorithms applicable and suitable in real-world application areas	K-NN, SVM, Decision tree, naïve bayes, random forest, clustering	recommends where machine learning model is optimal in a particular scenario	poor quality and insufficient data gives low accuracy results & Finding a learning model appropriate for the particular application area	Further enhancement of hybrid learning model

APPLICATION AREAS

The earlier section defines the history of machine learning and its techniques. In current segment, the glimpse regarding the applications, sectors and which algorithms are used will be defined. Currently, popularity of machine learning has increased tremendously in majorly all fields. Some of the observations are given below in Table 2: Overview of algorithms used in different application area:

Computer Vision: is a broad segment of machine learning which learns the machines using machine learning algorithms. The few application areas are listed below:

a) Face recognition

b) Speech recognition

c) Handwriting recognition

Financial Sector: It is one of the most popular and focused area where the possibility of fraud detection and prevention are going viral. Based on pattern detection, frauds can be detected and prevented at an early stage.

a) Bank sector

b) Stock Market

Healthcare: In healthcare industry the diagnosis of illness at an early stage is very important. Detecting the regularities and irregularities from patient's data are done with the

help of machine learning algorithms. Some of the areas are listed below:

a) Medicine

b) Breast cancer

c) Covid-19

d) Dermatology

e) Tumor

f) Heart Disease

Construction: Machine learning is making an impact in the construction industry. Machine learning algorithms are being widely used in order to identify potential problems in the construction industry.

a) Highways, roadways, flyovers

b) Hydroelectric power plant

c) Building

Cyber Security: Cybersecurity is very crucial in current scenario of Industry 4.0. the data has increased tremendously.

Agriculture: Machine learning algorithms used in different phases of agriculture-predicting crops, soil properties, weather, weed detection by assessing satellite images.

NLP and sentiment analysis: Natural language processing (NLP) includes the analysis of verbal or written language. In sentiment analysis, machine learning identifies and analyses the text for "optimistic", "non-optimistic" or "equitable" along with good, bad, very good, very bad, happy, sad etc.

Earthquake: Earthquake early warning systems was developed with the help of machine learning to detect large earthquakes and to damage control the after-effects. But this method was insufficient to trace the medium earthquakes. In order to prevent and control the damages, a hybrid of GPS stations and seismometers was developed known as Distributed Multi-sensor Earthquake Early Warning (DMSEEW).

Table 2: Overview of algorithms used in different sectors

S. No	Sector	Algorithms Used	Used for	Purpose	Applications
1	Computer Vision	KNN, SVM, Naïve Bayes	Face Recognition,	helpful in contactless security	airports, security, aadhar cards
		ANN, Vector quantization, dynamic time wrapping	Speech Recognition,	processing, analysing, recognising	healthcare, military, car system
		KNN, SVM, Naïve Bayes	Handwriting Recognition	scanning and ditizing large documents	organizations with huge databases
2	Financial Sector	Classification, regression and Neural Network	Banking	identification of pattern	banks
		Artificial neural network, random forest	Stock Market	to control credits and assess the risks	market campaign
3	Healthcare	ANN, SVM, decision tree, random forest, deep learning	Dermatology	by analysing data give warning and diagnonis to patient	skin related disease
		Extreme machine learning, random forest	Breast Cancer	to predict breast cancer at an early stage	problems related to breast cancer
		CNN, regression models	Covid-19	To detect and predict number of cases around the world.	problems related to Covid-19

4	Construction	Neural network, multiple regression analysis	Hydroelectric Power plants	to predict cost estimation at an early stage	HEPP
		ANN, SVM , Random Forest, GA	construction sites: buildings, projects	to minimize the risk	Buildings, Projects
		Neural network	highways, roadways	to predict the cost, safety measures at an early stage	highways, roadways
5	Cybersecurity	classification, regression, neural network & Clutering techniques	cyber-anomalies, policy violations	to identify numerous cyber threats, attacks or intrusions	Preventing system, hardware, networks or digital data
6	Agriculture	NN, decision tree, regression techniques	crops, weather, soil	to increase overall productivity	weather prediction, soil properties & management, irrigation requirements
7	NLP and sentiment analysis	CNN, RNN & Reinforcement learning	syntax, semantics, expressions,	makes it easier to understand for human	Facebook, Twitter, Movie Reviews
8	Earthquake	DMSEEW (combination of GPS stations and seismometers)		to detect medium and large earthquakes	

CONCLUSION

Data is growing at a tremendous rate. This paves the way for different algorithms to handle this voluminous data. Machine learning is one such technique that can handle this voluminous data. Machine learning is growing rapidly and so is its application area. This paper gives an insight into various machine learning algorithms which are available along

with the areas where it is being applied. It will help in understanding the basics of different Machine learning algorithms, and which algorithm will be suitable in a particular application area. A comprehensive literature review has been carried out which helps in knowing the work done to date by various authors in this area. A tabular representation is presented at the end which gives an overview of algorithms used in different sectors on the basis of different parameters like sector, algorithms used for, purpose and applications. This will help in understanding in detail machine learning techniques.

REFERENCES

Simon, Annina et al (2015), "An Overview of Machine Learning and its Applications", *International Journal of Electrical Sciences & Engineering (IJESE*), Vol 1, Issue 1,22-24.

Gunduz Murat et al (2015), "An Early Cost Estimation Model for Hydroelectric Power Plant Projects Using Neutral Networks and Multiple Regression Analysis", *Journal of civil engineering and management, 21(4):* 470-477.

Dey Ayon (2016), "Machine learning algorithms: A review", *International Journal of Computer Science and Information Technologies,7 (3),*1174-1179.

Sharma Tanvi et al (2017), "Prediction of Heart disease using Cleveland Dataset: A Machine Learning Approach", *International journal of recent research aspects, Vol (3),* 17-21.

Celik Ozer et al (2018), "A Research on Machine Learning Methods and its Applications", *Journal of Educational Technology & Online learning, 1(3).*

Ahn Heejae et. Al (2018)., "Applications of Machine Learning for Construction Site", *35th International Symposium on Automation and robotics in construction (ISARC 2018).*

G. Mahalakshmi et al (2019), "Early Cost Estimation of Highway Projects in India using Artificial Neural Network", *Springer nature Singapore pvt ltd.*

Gondia Ahmed et al (2020), "Machine Learning Algorithms for Construction Projects Delay Risk Prediction", *J. Constr.Eng. Manage., 146(1):04019085.*

Du Amy et al (2020), "Review of Machine Learning in Predicting Dermatological outcomes", *Mini Review, 7, Article 266.*

Kevin Fauvel et al (2020), "A distributed Multi-Sensor Machine Learning Approach to earthquake Early Warning", *The thirty-fourth AAAI Conference on Artificial Intelligence (AAAI-20).*

Gupta Chhaya et al (2020), "Machine Learning Techniques and Extreme learning machine for early breast cancer detection*", International Journal of Innovative Technology and Exploring Engineering (IJITEE), 9(4).*

Hashemi Sanaz et al (2020), "Cost estimation and Prediction in construction projects: a systematic review on ML techniques", *SN Applied Science- a springer nature Switzerland, 2:1703*.

Sharma Neha et al (2021), "Machine Learning Deep Learning Applications-A vision", *Global Transitions Proceedings 2,* 24-28.

Iqbal H. Sarker (2021), "Machine Learning: Algorithms, Real-world Applications and Research Directions", *SN Computer Science- a springer nature, 2:160.*