International Journal of Advance Research in Science and Engineering
Volume No. 13, Issue No. 04, April 2024
www.ijarse.com
IJARSE
ISSN 2319 - 8354

Machine Learning Techniques: A Review

G. Arunalatha

Assistant Professor, Department of Computer Science and Engineering,
Perunthalaivar Kamarajar Institute of Engineering and Technology (PKIET), Karaikal.
Puducherry.
Email: vigneshgayu1121@gmail.com

Abstract:

A subfield of computer science and artificial intelligence called machine learning (ML) is concerned with using data and algorithms to help AI mimic human learning processes and progressively become more accurate. It is a technological tool that works on the theory that information may be learned by a computer without the need for human intervention. It examines vast amounts of data or training data using algorithms to find distinctive patterns. These patterns are examined by this system, which then classifies and predicts the data. In traditional machine learning, data is labeled by humans and the computer learns how to interpret it. This paper discusses the various machine learning methods.

Keywords: Machine learning, Supervised learning, UnSupervised Learning, Semi Supervised learning, Reinforcement learning

1. INTRODUCTION

The field of artificial intelligence (AI) known as machine learning (ML) gives computers the capacity to automatically learn from data and prior experiences in order to recognize patterns and make predictions with the least amount of human intervention. Without explicit programming, computers can function independently with the use of machine learning techniques. Applications for machine learning are fed fresh data and have the ability to learn, grow, evolve, and adapt on their own. Machine learning is embedded in an array of business applications and software. It is commonly used in search engines, emails for spam filters, banking software for fraud detection, chatbots and apps in the form of speech recognition and predictive text. It can also be used for security purposes like analyzing email communications or internet usage. Organizations can benefit from machine learning as well with process automation — freeing up time and resources.

It turns out that machine learning is helpful, particularly in the big data world of today. It can recognize voice commands on our phones, suggest songs on Spotify or Amazon to buy next, and even use Waze to find the quickest route to your location, to name a few of the technologies it supports. In the healthcare industry, machine learning aids physicians in diagnosing illnesses more accurately and quickly. In the

banking sector, it helps identify fraudulent activity that deviates from customers' typical spending patterns.

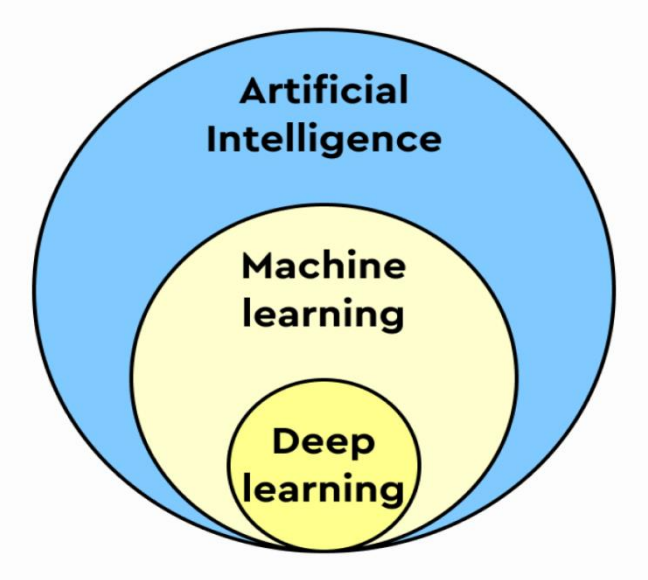

Fig:1 Machine Learning

The process of machine learning involves teaching algorithms on data sets to produce predicted results, like recognizing an object or seeing a pattern. Machine learning is the process of optimizing the model so that it can predict the correct response based on the training data samples. The more training samples the machine learning algorithm receives, assuming the training data is of high quality, the more accurate the model will become. During the "fitting process," which is part of the training process, the algorithm adapts the model to the data. The algorithm is repeatedly retrained until it produces the correct response if the result does not match the intended result. Essentially, the algorithm makes decisions based on its learning from the data and determines whether the response and the input fit into a line, cluster, or other statistical correlation.

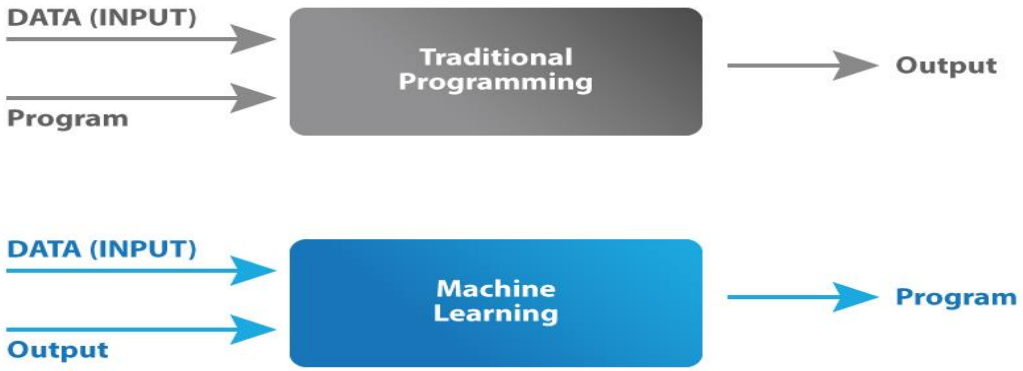

Fig:2 Machine Learning & Traditional Programming

Data—numbers, images, or text—is the foundation of machine learning. Examples of data include bank transactions, images of individuals or even specific bakery goods, repair records, time series data from sensors, and sales reports. To be utilized as training data, or the information the machine learning model will be trained on, the data is collected and prepared. Programmers then select a machine learning model to utilize, provide the data, and allow the computer model to learn on its own to identify trends or anticipate future events. As time goes on, the model's parameters can be altered by a human programmer to assist it provide results that are increasingly accurate.

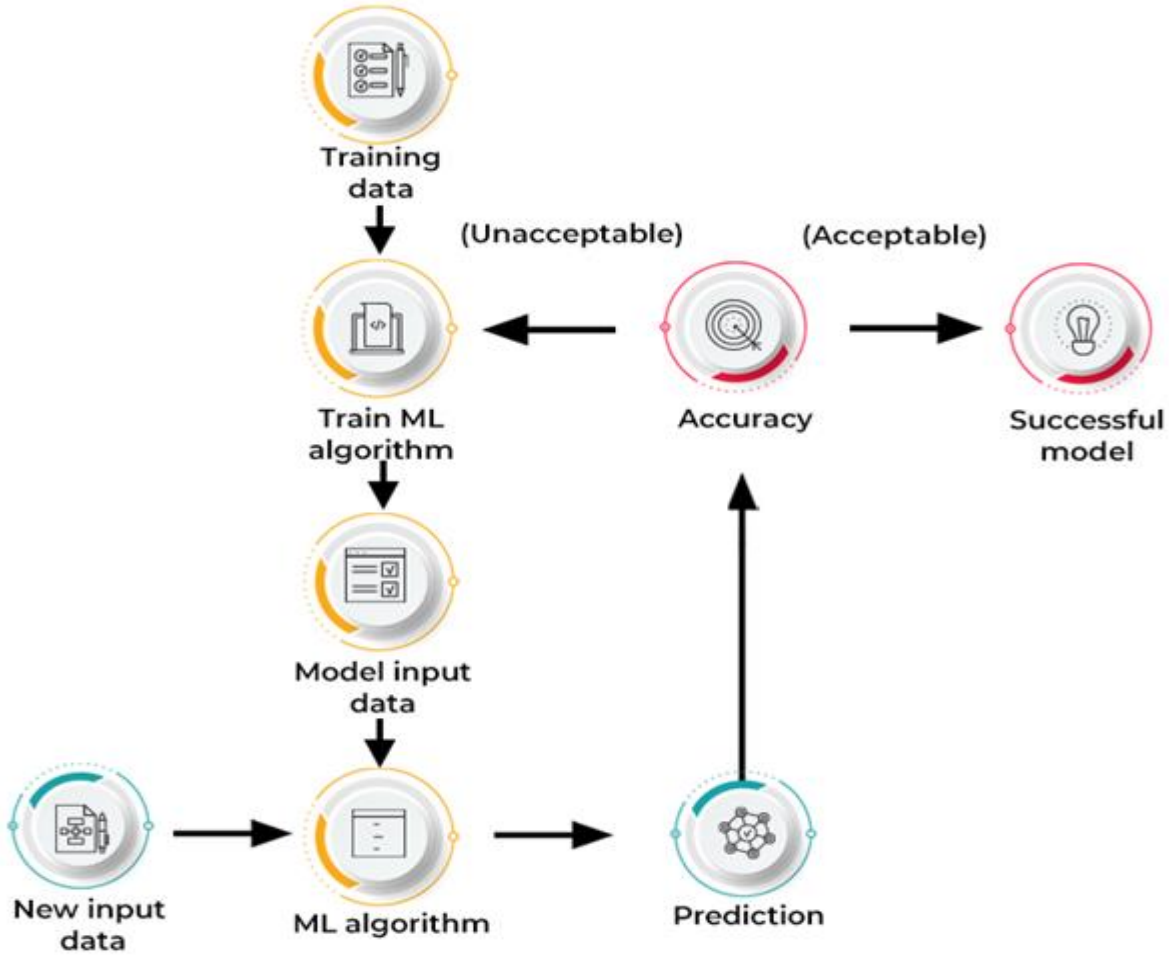

Fig:3 Work flow of Machine Learning Algorithm

2. SUPERVISED LEARNING:

Algorithms are trained on labeled data sets with tags that describe each item of data in supervised machine learning. The algorithms receive data that is accompanied by a "answer key" that provides guidance on how to interpret the data. For example to recognize fruits when presented with fresh photographs, an algorithm might be fed images of fruits that have tags for each type of fruit. Machine learning models for categorization and prediction are frequently developed via supervised machine learning.

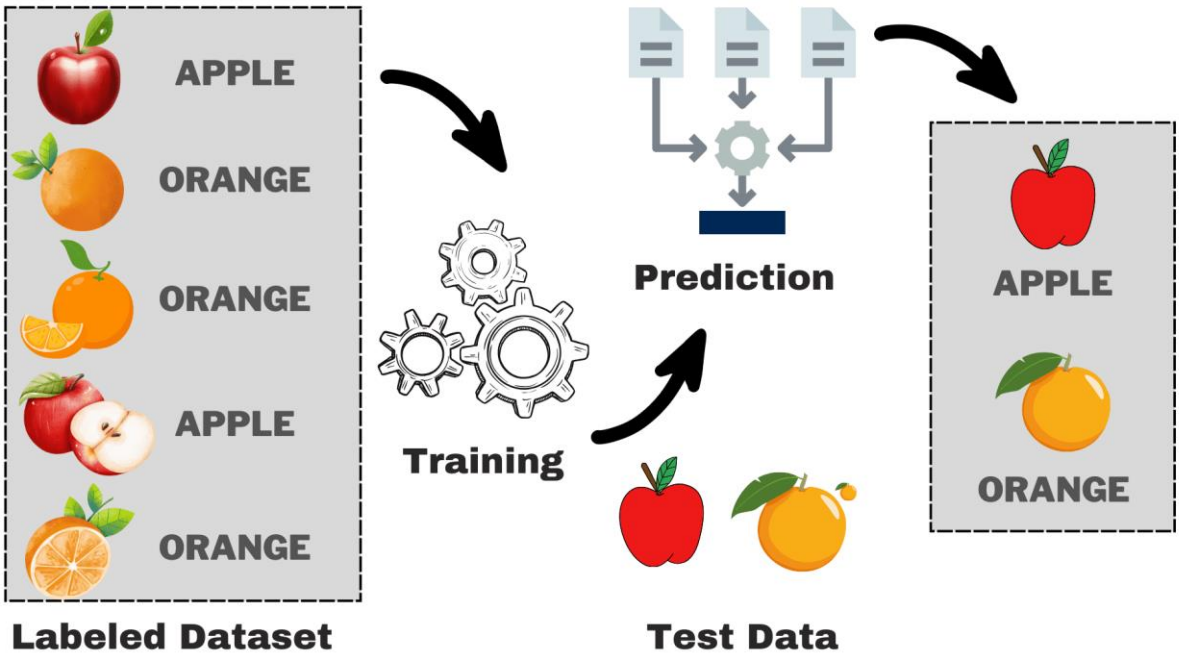

Fig:4 Supervised Machine Learning

3. UNSUPERVISED LERNING:

Unsupervised learning is a type of machine learning model that identifies patterns in unstructured, or unlabeled, data. In contrast to supervised learning, the output's "correctness" is unknown in advance. Instead, the algorithm classifies the data into groups according to attributes as it learns from it without human input (i.e., it is unsupervised). When presented with images of apples and bananas, for example, the algorithm can identify which image is an apple and which is a banana on its own. Pattern matching and descriptive modeling are two areas where unsupervised learning excels. Today's most often used unsupervised learning algorithms include partial least squares, fuzzy means, K-means clustering, and hierarchical clustering.

International Journal of Advance Research in Science and Engineering
Volume No. 13, Issue No. 04, April 2024
www.ijarse.com
IJARSE
ISSN 2319 - 8354

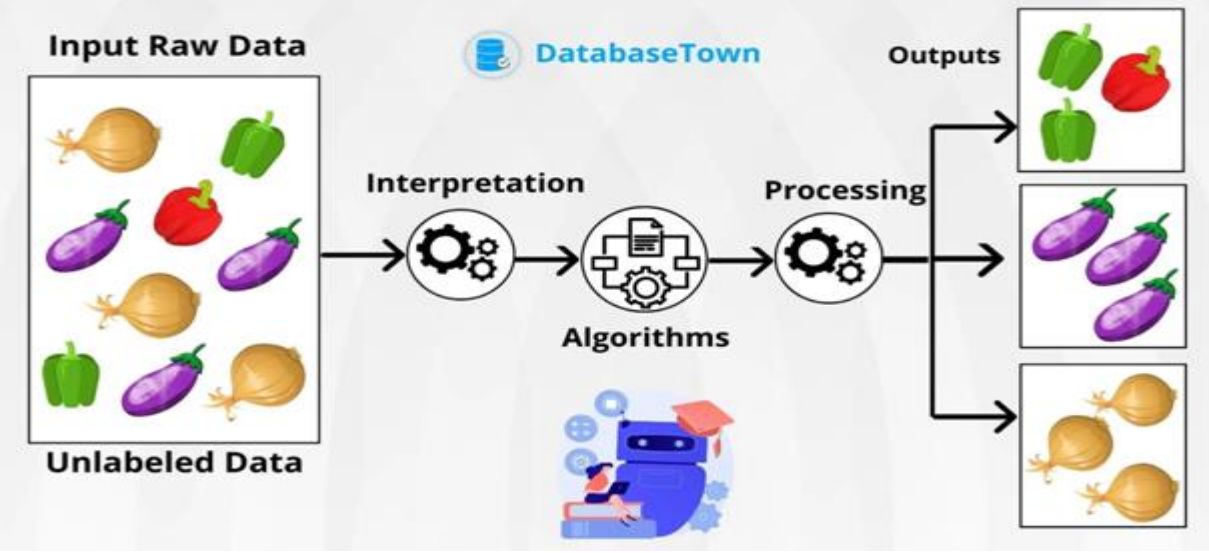

Fig:5 UnSupervised Machine Learning

4. SEMI SUPERVISED LEARNING:

Semi-supervised learning is the third of four machine learning models. In a perfect world, all data would be structured and labeled before being input into a system. But since that is obviously not feasible, semi-supervised learning becomes a workable solution when vast amounts of raw, unstructured data are present. This model consists of inputting small amounts of labeled data to augment unlabeled data sets. Essentially, the labeled data acts to give a running start to the system and can considerably improve learning speed and accuracy. A semi-supervised learning algorithm instructs the machine to analyse the labeled data for correlative properties that could be applied to the unlabeled data.

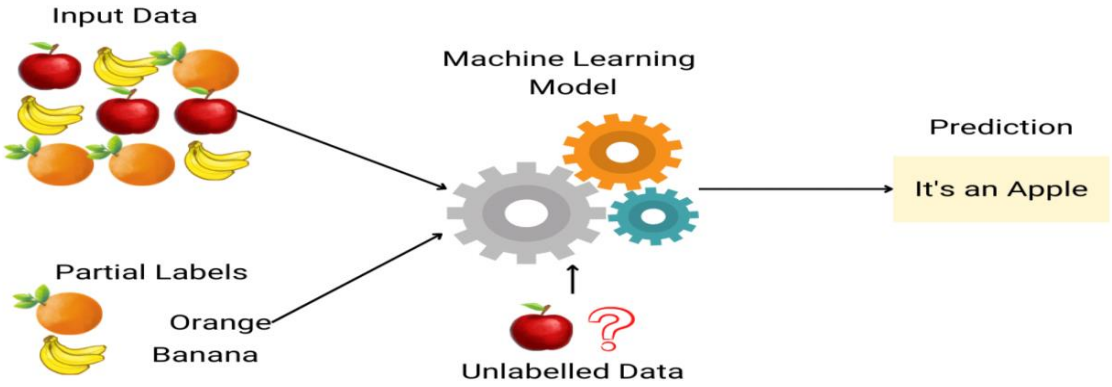

Fig:6 Semi Supervised Machine Learning

5. REINFORCEMENT LEARNING

A machine learning training technique called reinforcement learning works by rewarding positive behaviors and penalizing undesirable ones. The thing being taught, or a reinforcement learning agent, can typically sense and interpret its surroundings, act, and learn by making mistakes. An agent must explore an unknown environment in order to accomplish a goal in the Reinforcement Learning issue. The foundation of reinforcement learning (RL) is the idea that the maximization of predicted cumulative reward can characterize any goal. In order to obtain the greatest possible reward, the agent must develop the ability to perceive and alter the condition of the environment.

6. Applications of Machine Learning :

Strategies are required in generalized epidemic situations to give priority to people who are more likely to contract the human immunodeficiency virus (HIV) for prevention interventions. HIV risk scores were constructed using population-level HIV testing data from rural Kenya and Uganda, and their capacity to detect seroconversions was evaluated. When comparing the classification of people at risk of HIV acquisition with a model-based method, machine learning[1] produced better results. For estimating patients' chances of survival for colorectal cancer, a method[2] was developed by utilizing case-based reasoning and nearest neighbor analysis. This study included 216 patients with complete clinic pathologic records and a five-year follow-up. They were separated into a test group consisting of 54 cases and a core database of 162 cases, with follow-up on every patient.

There is an evidence that being obese and overweight raises one's risk of developing some types of cancer. Studies conducted on adults provide evidence for the involvement of oestrogens and insulin-like growth factors (IGFs) in the pathogeneses of many malignancies. The prepubertal obese children have hormone changes that are linked to cancer risk in fat adults. For the study[3], a group of 40 obese children ages 6 to 9 and a control group matched for age and sex were used. Innovative techniques for detection and prevention have been developed as a result of the threat that financial transaction fraud poses to both individuals and companies. This research study[4] examines the application of machine learning algorithms and real-time monitoring systems to enhance fraud detection and prevention in financial transactions. Two proactive methods of preventing fraud that are currently being studied are dynamic risk scoring and adaptive thresholds. Data security and legal compliance, along with scalability and deployment considerations, are also discussed.

A Swarm-Optimized Fuzzy Instance-based Learning approach[5] was done to predict slope collapses by optimizing the FKNN hyper-parameters using the firefly algorithm. The suggested model outperformed alternative techniques for comparison.

In order to enhance the classifier's performance for the IDSs, a unique fuzziness-based semi-supervised learning strategy[6] that makes use of unlabeled examples aided by supervised learning algorithms was discussed. The fuzzy quantity is used to perform sample categorization (low, mid, and high fuzziness categories) on unlabeled samples after a single hidden layer feed-forward neural network (SLFN) is trained to produce a fuzzy membership vector.

7. Conclusion

Algorithms used in machine learning are trained to identify patterns and relationships in data. They reduce dimensionality, classify information, cluster data points, make predictions, and even assist in the creation of new content by using past data as input. This paper reviews the various machine learning techniques and their applications. Machine learning is used in healthcare to make diagnosis and recommend treatments. Predictive maintenance, corporate process automation, malware threat detection, fraud detection, and spam filtering are further popular machine learning use cases.

References:

[1]. Balzer LB, Havlir DV, Kamya MR, Chamie G, Charlebois ED, Clark TD, Koss CA, Kwarisiima D, Ayieko J, Sang N, Kabami J, Atukunda M, Jain V, Camlin CS, Cohen CR, Bukusi EA, Van Der Laan M, Petersen ML. Machine Learning to Identify Persons at High-Risk of Human Immunodeficiency Virus Acquisition in Rural Kenya and Uganda. Clin Infect Dis. 2020 Dec 3;71(9):2326-2333. doi: 10.1093/cid/ciz1096. PMID: 31697383; PMCID: PMC7904068.

[2] Hamilton PW, Bartels PH, Anderson N, Thompson D, Montironi R, Sloan JM. Case-based prediction of survival in colorectal cancer patients. Anal Quant Cytol Histol. 1999 Aug;21(4):283-91. PMID: 10560505.

[3]. Gascón, F., et al. “Childhood Obesity and Hormonal Abnormalities Associated with Cancer Risk.” European Journal of Cancer Prevention, vol. 13, no. 3, 2004, pp. 193–97. JSTOR, http://www.jstor.org/stable/45051398. Accessed 5 Apr. 2024.

[4] Kazeem, Oladimeji. (2023). FRAUD DETECTION USING MACHINE LEARNING. 10.13140/RG.2.2.12616.29441.

[5] Cheng, Min-Yuan & Hoang, Nhat-Duc. (2015). A Swarm-Optimized Fuzzy Instance-based Learning approach for predicting slope collapses in mountain roads. Knowledge-Based Systems. 10.1016/j.knosys.2014.12.022.

[6] Raza, Rana Aamir & Wang, Xi-Zhao & Huang, Joshua & Abbas, Haider & He, Yu-lin. (2016). Fuzziness based semi-supervised learning approach for Intrusion Detection System. Information Sciences. 378. 10.1016/j.ins.2016.04.019.