AG Publishing House

Handbook On Fundamentals And Methods Of Machine And Deep Learning(VOLUME-1)

Volume 1
Year: 2024

AG PH Books

Evolution of Machine Learning Algorithms: A Comprehensive Review

Dr. D. Venkata Lakshmi[1*]

[1]*Professor, School of Computer Science and Engineering, VIT - AP University, Andhra Pradesh.*

Abstract

Machine learning (ML) enables machines to process and interpret large datasets, providing efficient solutions where traditional methods fall short. As data availability increases, the demand for ML grows across various industries to extract valuable insights. ML is a multidisciplinary field with diverse research areas aimed at teaching machines to learn from data without explicit programming. This review paper discusses the evolution of machine learning algorithms, including supervised, unsupervised, and reinforcement learning techniques. It also explores the applications of machine learning in different sectors. The paper highlights the progress of ML algorithms, their types, and their growing impact on data-driven decision-making.

Keywords: Machine Learning, Algorithms, Evolution, ML, Applications.

1 Introduction

Computers may learn to analyse data more effectively using machine learning (ML). Occasionally, users are unable to identify the information after they have seen the data. They use machine learning in these circumstances. The necessity for machine learning is growing as a result of the abundance of datasets. Machine learning is used in many sectors to retrieve pertinent data. Because it incorporates ideas from a wide range of academic fields, machine learning is a field that is interdisciplinary. Learning from the data is the aim of machine learning. The question of how to train robots to learn independently of human

[*] ISBN No. - 978-81-974433-9-8

instruction has been the subject of much research. Numerous mathematicians and programmers tackle this issue using a variety of techniques, including massive data sets.(Dey, 2016)

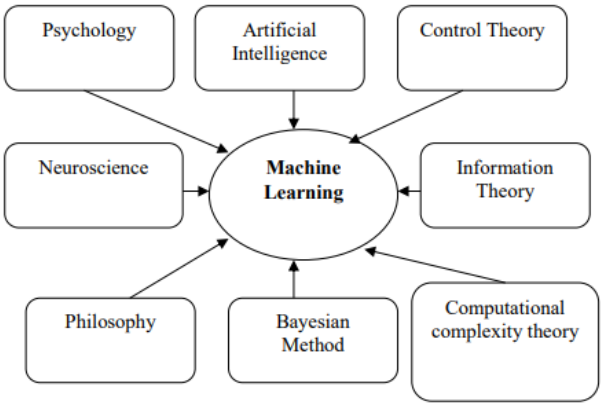

Figure 1 The Multi-disciplinary ML(Alzubi et al., 2018)

Development of Machine Learning

Instead of being an emerging concept, machine learning has a prototype that was developed in the mid-1950s to mid-1960s. Realising different self-organising and adaptable systems was the primary research objective at the time. To put it another way, it might also be seen as an unsupervised machine learning approach, which many academics have been interested in. By the mid-1970s, concept learning had become the main focus of machine learning research. Semantic networks and predicate logic were the instruments used at the time, and academics' excitement also waned.(Xu, 2019)

About ten more years later, idea learning was still the focus of the study. Studies were expanded from learning a single idea to learning many concepts, examining different learning approaches and tactics; in other words, advancements were achieved in both breadth and depth. The 1980s, which may be considered a time of intense progress in this discipline, saw the beginning of a new era in development of machine learning. The resurgence of neural networks has undoubtedly fuelled the investigation of different machine learning techniques and systems, as well as the transition from theoretical study to real-world machine learning applications. The advancement of machine learning is anticipated to be more important in the future. The use of machine learning will expand as big data continues to evolve.

Evolution of machine learning algorithms

Artificial intelligence (AI) is a fascinating area that has been impacted by major turning points in machine learning history. The inception of machine learning could be linked to significant events that occurred in the mid-1900s and paved the way for subsequent advancements. When Warren McCulloch and Walter Pitts created the first neural network using an electric circuit in 1943, it was the first attempt at machine learning. This innovation paved the way for further advancements in machine learning by tackling the

problem of allowing computers to interact with one another. Alan Turing developed the seminal Turing Test in 1950, which has since become a standard for evaluating artificial intelligence's capacity to mimic human behaviour. This groundbreaking idea established the framework for assessing machine intelligence by measuring its ability to replicate human reactions. (Zaidi et al., 2018)

Arthur Samuel invented machine learning in 1952 when he created first computer program that could play checkers at the championship level. The software changed the field of AI that plays games by using strategies like alpha-beta pruning. When Cover and Hart first presented the Nearest Neighbour Algorithm in 1967, it quickly emerged as a crucial technique for automatically spotting patterns in big datasets. This system helped with classification and pattern recognition tasks, showcasing machine learning's ability to handle intricate datasets. Back propagation, a major development in neural network training, was introduced in 1974. This method's basis was established by Paul Werbos, who sought to increase predictive power and model accuracy by modifying weights.

The AI winter, which lasted from late 1970s to the 1990s, was a time of lower funding and less zeal as a result of unfulfilled expectations. Machine learning persisted in spite of the difficulties, which resulted in a comeback in the latter half of the 20th century. According to Moore's Law, the exponential increase in processing power is responsible for the emergence of machine learning in twenty-first century. IBM's Deep Blue beat chess champion Garry Kasparov in 1997, demonstrating how machine learning may outperform humans at challenging tasks.

In 2002, Geoffrey Hinton, Pedro Domingos, & Andrew Ng developed the Torch software package, which offered a scalable platform for data science and machine learning. Torch made machine learning technologies more accessible by laying the foundation for later libraries and frameworks. The promise of deep learning within pattern recognition tasks was shown by Geoffrey Hinton's seminal work on deep belief networks in 2006. The creation of Google Brain in 2011 further cemented the importance of deep learning, resulting in innovations like as AlphaGo.

Types of Machine Learning Algorithm

Numerous algorithms are used in machine learning to analyse and forecast data. It has several methods to read data. What you are specifically looking for from the data, how many variables are involved, which model is ideal, and other factors to take into account all influence the sort of approach that is used.(Negi & Bhavsar, 2023)

1. **Supervised learning algorithms**

Two subsets of supervised learning exist in data mining: classification and regression.

- **Classification:** Classification successfully separates test results into distinct groups using an algorithm. It makes an effort to make inferences about the proper definition or labelling for certain items in the collection. Common classification methods include random forests, K-nearest neighbour, decision trees, support vector machines, and linear classifiers.

- **Regression:** Regression is used to understand relationship between variables that are both independent and dependent. Forecasting things like company's sales revenue is one of its common uses. Polynomial, logistic, as well as linear regression are common regression techniques.

Supervised machine learning processes use a variety of calculation methods and algorithms, which are often computed using Python and other similar programs. Among the supervised learning algorithms are:

AdaBoost or gradient boosting: In order to generate a stronger algorithm with fewer mistakes, AdaBoost, also known as gradient boosting/ adaptive boosting7, combines regression algorithm which performs badly with weaker ones. The predictive capability of many base estimators is combined in boosting.

Artificial neural networks: Artificial neural networks, often known as neural networks or simulated neural networks, are the foundation of deep learning techniques. A subset of machine learning methods are artificial neural networks (ANNs). To find patterns in incoming data, the learner algorithm makes use of neural building blocks, that are gradually trained and modified to resemble the neurones in the human brain. Study up on "neural networks."

Algorithms for decision trees: Decision trees, which may be shown with a tree diagram, use a branching series of connected choices to categorise data into categories and forecast numerical values (regression issues). Unlike the neural network's dark box, decision trees have the benefit of being simple to audit and verify.

Dimensionality reduction: A data collection is said to have high dimensionality if it contains a large number of features. The most significant ideas or information are then left behind after dimensionality reduction reduces the number of characteristics. Principal component analysis is one example.

K-nearest neighbour: This non-parametric method, sometimes referred to as KNN, groups data points according to their connections and closeness to other publically accessible data. It is predicated on the idea that comparable data points may be discovered nearby. It initially tries to determine the distance between data points, usually using the Euclidean distance, in order to categorise the data according to average or most common category.

Linear regression: Determines the connection among a dependent variable as well as one or more independent factors. It is often used to guess what will happen in the future. In basic linear regression, there is only a single dependent variable and no independent variables.

Logistic regression: Logistic regression is used for categorical variables that are dependent, such as those having binary outputs including "true" and "false" or "yes" and "no," whereas linear regression is used for continuous dependent variables. Despite the fact that both regression models aim to comprehend the connections among data inputs, the logistic regression approach is mostly used to address binary classification issues, including spam identification.

Neural networks: Neural networks are mostly used for deep learning algorithms that evaluate training data utilising layers of nodes, simulating the connections seen in the human brain. Inputs, weights, bias, and a output comprise each node. Node "fires," or activates, and transmits data to network's next layer if the output value rises beyond a certain threshold. Neural networks employ gradient descent to learn from changes, depending on the loss function. You can be certain that the model is accurate when cost function is at or close to zero.

Naïve Bayes: This proves that each predictor had an equal impact on a certain result and that the presence of one trait does not change the chance of another occurrence. The three types of Naïve Bayes classifiers are Bernoulli Naïve Bayes, Multinomial Naïve Bayes, and Gaussian Naïve Bayes. The primary uses of this technology are recommendation systems, spam detection, and text classification.

Forests at random: To forecast a value or category, a random forest machine learning algorithm aggregates the results from several decision trees. Uncorrelated decision trees are referred to as the "forest" because they are put together to lower variance and provide forecasts that are more accurate.

SVMs, or support vector machines: By constructing a hyperplane where the distance between two classes of the data points is largest, this method is most often used for regression problems, however it may also be used to classification difficulties. The decision boundary is a hyperplane that divides data point classifications (oranges vs. apples) on each side of the plane.

2. Unsupervised learning algorithms

Unsupervised learning makes use of unlabelled data, in contrast to supervised learning. By finding patterns in the data, the program assists in resolving association or clustering issues. This is particularly helpful when subject-matter specialists are unclear about usual data collecting features. K-means, Gaussian mixture models, dimensionality reduction, and hierarchical Popular clustering techniques include PCA and t-SNE.

Clustering: These algorithms make it possible to organise data by identifying patterns in it. The capacity of computers to identify distinctions among data items that users would have overlooked might be useful to data scientists.

Hierarchical Clustering: This turns the data into a tree of clusters. Considering each data point as a separate cluster is the initial stage in the hierarchical clustering process. Then, it does these actions repeatedly: 1) Determine which two clusters are the closest to one another, and 2) combine the two most similar clusters. Until every cluster is combined, these procedures are repeated.

K-means clustering: This finds groupings of data that are similar to each other, therefore identifying groups within data with labels9 into distinct clusters. The centroids it employs to define clusters are the source of the term "K-means." If a point is closer to the centroid of a given cluster than to any other centroid, it is allocated to that cluster.

3. Reinforcement Learning

Reinforcement learning algorithms learn to make a sequence of decisions. The algorithm acquires the ability to achieve an objective in an environment that is potentially complex and unpredictable. An agent makes decisions by following a policy during reinforcement learning, and it learns from the outcomes of these decisions by getting rewards or penalties.

Q-learning: Finds the value of a certain action in a certain state using model-free reinforcement learning method.

DQNs, or deep Q-networks: The technique may directly train efficient policies using high-dimensional sensory inputs by combining deep neural networks with Q-learning.

Policy Gradient approaches: These methods specifically optimise the parameters of a policy rather than assessing the worth of activities.

Monte Carlo Tree Search (MCTS): Often employed in games like Go, MCTS is a decision-making technique that uses scenario-playing to determine the best options.

A general summary of the most prevalent kinds of machine learning algorithms is given by these categories. Each is more suited for certain jobs than others due to its advantages and optimal use cases.

categories. Each is more suited for certain jobs than others due to its advantages and optimal use cases.

Applications of Machine Learning

Self-driving cars and gaming are examples of machine learning difficulties. Here are a few well-known real-world uses for machine learning.(Alzubi et al., 2018)

Playing Checkers: Through experience gained from competing against itself, a computer program learns how to play checkers and adjusts its performance based on its ability to succeed at different classes of activities involving the game.

Speech Recognition: Machine learning methods are used in some capacity by the most advanced voice recognition systems available today. For instance, the SPHINX system uses voice signals to learn words and sounds unique to a speaker. Several neural network learning techniques for deciphering hidden Markov models perform quite well for autonomously modifying noise, speakers, dictionaries, and other elements.

Autonomous Vehicles: These days, machine learning models are used to operate autonomous vehicles, such as cars and drones. For instance, Tesla vehicles and Google driverless cars. Sensor-based applications may also be effectively controlled by machine learning methods.

Email Filtering (Spam Emails): Spam emails may be filtered using machine learning. The machine learning-based model will just commit to memory any email that the user has flagged as spam. The machine learning-based algorithm will search, compare, and rely on the prior spam emails when a new

email comes in the inbox. A new email will be pushed to the user's inbox if it matches any of them; if not, it will be flagged as spam.

Artificial Intelligence and Robotics: Machine learning is thought to be a better method of issue resolution. Learning may be enhanced using foundational knowledge and training data using machine learning models, propelling robots and artificial intelligence to new heights.

2 Literature Review

(Dey, 2016) This study has examined a number of machine learning methods. To mention a few uses, these algorithms are used in machine learning, predictive analytics, data mining, and other fields. The primary benefit of machine learning is that its algorithms can do tasks automatically after they have learnt how to handle data.

(Alzubi et al., 2018) The present SMAC (Social, Mobile, Analytic, Cloud) technology movement will eventually merge big data, networked processes, and intelligent machines. Adoption of machine learning methods and solutions is accelerating due to the massive amount of data generated by this virtual environment. Machine learning enables computers to imitate and adapt to human-like behaviour. Every interaction and activity is converted by machine learning into something that the system can use to learn and apply to other situations in the future. The data analytics approach that enables computers to learn & do human-like activities, including learning from experience, is introduced in this work. It explains the what, how, and why of machine learning by going over its idea, vocabulary, and applications. The technological roadmap is examined to comprehend and assess the commercial and industrial practice possibilities of machine learning. The fundamental goal of this research is to define machine learning and its potential future applications.

(Zaidi et al., 2018) Algorithms for machine learning have evolved remarkably, transforming a variety of industries and influencing the science of artificial intelligence. From the beginning to the present, the development of machine learning algorithms is mapped out in this review paper's thorough historical examination. This research attempts to provide a comprehensive view of the development of these algorithms by examining the turning points, innovations, and paradigm changes.

(Xu, 2019) In this study, machine learning methods in artificial intelligence are primarily introduced. It begins by outlining how these algorithms are categorised and the primary use cases for them. The document then gives the author's opinions and explains the underlying algorithms. Lastly, growing trend of algorithms for machine learning is projected.

(Pineda-Jaramillo, 2019) In recent decades, researchers have used a number of machine learning approaches to study a wide range of topics connected to transport planning.This review article begins with a summary of the many machine learning algorithms often used in transportation research. Artificial Neural Networks, Support Vector Machines, Decision Trees, and Cluster Analysis are some of these techniques. The many methods that scholars have used to depict the choice of travel mode are then compared with Multinomial Logit Model, most popular discrete choice model. Following the description

of machine learning algorithms, Random Forest (RF), a variant of Decision Tree algorithms, offers the best method for simulating travel mode choice.

(Mahesh, 2020) The scientific investigation of statistical models and methods used by computer systems to carry out certain tasks without explicit programming is known as machine learning (ML). Acquiring knowledge of algorithms in several applications that we use on a regular basis. One of the reasons online search engines like Google perform so effectively is because they have a learning algorithm that has already figured out how to rank web sites. To mention a few uses, these algorithms are used in predictive analytics, data mining, image processing, and other fields. The primary benefit of machine learning is that algorithms can do tasks automatically after they have learnt how to handle data. This article provides a concise overview and outlook on the wide range of uses for machine learning techniques.

(Lu & Zhou, 2021) Because there were no sophisticated technology or machines in the ancient past, people had to do all kinds of labour by hand. Numerous cutting-edge technologies had not yet been built, and science had not yet advanced. Therefore, the operation was entirely relied on the people, and people have realised that "the technology of the future is the wisdom of the moment." After its inception in the 1950s, artificial intelligence (AI) has seen substantial development. This essay offers a thorough analysis of AI's past, current, and future advances.

(Sarker, 2021) In the current Fourth Industrial Revolution, there is a wealth of data fromInternet of Things, social media, business, cybersecurity, mobile, and health. One has to have a firm grasp of artificial intelligence (AI), and particularly machine learning (ML), in order to analyse this data efficiently and create the associated intelligent and automated applications. Numerous algorithms, including reinforcement learning, supervised, unsupervised, as well as semi-supervised learning, are included in the field of machine learning. Furthermore, deep learning is capable of intelligently analysing vast volumes of data as a subset of a larger family of machine learning algorithms. This article offers a thorough analysis of several machine learning techniques that may be used to improve an application's intelligence and capabilities.

(Abdel-Jaber et al., 2022) Artificial neural networks are used in deep learning to identify patterns and draw conclusions from them. Deep learning is a kind of machine learning that uses artificial neural networks to model the human brain. It learns on its own in deep structures using machine learning techniques including supervised, semi-supervised, and unsupervised learning algorithms. It has become quite famous because to its amazing capacity to learn from vast volumes of data. Deep learning algorithms have been shown to be an excellent way to analyse large amounts of data.

(Woodman & Mangoni, 2023) A renewed interest in machine learning research, notably data-hungry deep learning algorithms, is being fuelled by the global expansion of access to health data. These days, improved computationally effective algorithms provide special chances to improve risk assessment, diagnostics, and individualised patient care strategies. These potential are especially pertinent to the care of elderly patients, a population that exhibits substantial interindividual variability within homeostatic capability, organ function, and responsiveness to therapy, as well as complicated multimorbidity

patterns. With major ramifications for almost all medical specialisations in the next age of digital medicine, clinical tools that use machine learning algorithms to decide the optimal course of therapy are progressively receiving the required regulatory body clearance and being implemented into healthcare.

(Negi & Bhavsar, 2023) In the beginning, computers were used to carry out intricate computations and store data along with instructions on how to manage it. Managing, modifying, and storing data while adhering to instructions was the primary emphasis of computer development. Additionally, algorithms were developed that could read vast volumes of data, apply categorisation and prediction rules that comprehended the data, and draw conclusions from it. Machine learning, a branch of artificial intelligence, got its start with this.

(Ezugwu et al., 2024) Although machine learning (ML) has revolutionised many disciplines, its future development depends on an understanding of its underlying research. Through a thorough bibliometric analysis research, this work analyses the state-of-the-art publications spanning twelve decades and provides an overview of the main classical machine learning algorithms. In order to find important insights, we used methods like citation and keyword analysis to examine a dataset of highly cited articles from prestigious ML conferences and publications. The report also identifies the most important articles and authors, shows how the ML community's collaboration networks are changing, and indicates popular research subjects and new areas of interest.

(Harrison, 2024) In the ever-changing world of modern technology, this paper explores the revolutionary path of machine learning (ML). The abstract offers a succinct synopsis of the main elements of the research. From early learning algorithms to the revival of interest in neural networks, the historical study reveals the beginnings of machine learning and highlights significant events that influenced its course. Extensive case studies highlight the many uses of machine learning, demonstrating its influence in sectors such as manufacturing, healthcare, and finance.

(Sharma & Arora, 2024) This study provides a thorough examination of the development and diverse effects of artificial intelligence (AI) on modern civilisation. The presentation starts with the philosophical foundations and historical turning points before exploring the first wave of AI, which was characterised by rule-based systems, and the difficult AI Winter of the 1980s and 1990s. The groundwork for the revolutionary era of artificial intelligence is laid by the renewed interest in the late 1990s, especially with the introduction of machine training and neural networks.

3 Conclusion

In conclusion, machine learning (ML) has become an essential tool for efficiently handling and interpreting large datasets, especially when traditional methods fail. As the demand for ML continues to rise across various industries, it plays a vital role in extracting valuable insights from complex data. This review paper examined the evolution of machine learning algorithms, covering key types such as supervised learning (including classification and regression), unsupervised learning, and reinforcement learning. Additionally, it explored the diverse applications of ML, highlighting its growing importance in various sectors.

References

[1]. Abdel-Jaber, H., Devassy, D., Al Salam, A., Hidaytallah, L., & El-Amir, M. (2022). A Review of Deep Learning Algorithms and Their Applications in Healthcare. Algorithms, 15(2). https://doi.org/10.3390/a15020071

[2]. Alzubi, J., Nayyar, A., & Kumar, A. (2018). Machine Learning from Theory to Algorithms: An Overview. Journal of Physics: Conference Series, 1142(1). https://doi.org/10.1088/1742-6596/1142/1/012012

[3]. Dey, A. (2016). Machine Learning Algorithms: A Review. International Journal of Computer Science and Information Technologies, 7(3), 1174–1179. www.ijcsit.com

[4]. Ezugwu, A. E., Ho, Y.-S., Egwuche, O. S., Ekundayo, O. S., Van Der Merwe, A., Saha, A. K., & Pal, J. (2024). Classical Machine Learning: Seventy Years of Algorithmic Learning Evolution. Data Intelligence. https://doi.org/10.3724/2096-7004.di.2024.0051

[5]. Harrison, H. (2024). Navigating the Frontier : Machine Learning ' s Evolution in Modern Technology. Department of Computer Science, University of Manchester Abstract:, 1–15.

[6]. Lu, Y., & Zhou, Y. (2021). A review on the economics of artificial intelligence. Journal of Economic Surveys, 35(4), 1045–1072. https://doi.org/10.1111/joes.12422

[7]. Mahesh, B. (2020). Machine Learning Algorithms - A Review. International Journal of Science and Research (IJSR), 9(1), 381–386. https://doi.org/10.21275/art20203995

[8]. Negi, A., & Bhavsar, H. (2023). Understanding The Evolution Of Machine Learning Algorithms. Data Science and Intelligent Computing Techniques, January 2023, 255–265. https://doi.org/10.56155/978-81-955020-2-8-22

[9]. Pineda-Jaramillo, J. D. (2019). A review of machine learning (ML) algorithms used for modeling travel mode choice•. DYNA (Colombia), 86(211), 32–41. https://doi.org/10.15446/dyna.v86n211.79743

[10]. Sarker, I. H. (2021). Machine Learning: Algorithms, Real-World Applications and Research Directions. SN Computer Science, 2(3), 1–21. https://doi.org/10.1007/s42979-021-00592-x

[11]. Sharma, T., & Arora, R. (2024). The Evolution of Artificial Intelligence - A Comprehensive Review. International Journal of Science, Engineering and Technology.

[12]. Woodman, R. J., & Mangoni, A. A. (2023). A comprehensive review of machine learning algorithms and their application in geriatric medicine: present and future. In Aging Clinical and Experimental Research (Vol. 35, Issue 11). Springer International Publishing. https://doi.org/10.1007/s40520-023-02552-2

[13]. Xu, P. (2019). Review on Studies of Machine Learning Algorithms. Journal of Physics: Conference Series, 1187(5). https://doi.org/10.1088/1742-6596/1187/5/052103

[14]. Zaidi, N., Singh, B., & Yadav, S. (2018). The evolution of machine learning algorithms: A comprehensive historical review. International Journal of Applied Research, 4(9), 49–55. https://doi.org/10.22271/allresearch.2018.v4.i9a.11451