Deep Science Publishing, 2025
https://doi.org/10.70593/978-93-7185-365-1_2

DeepScience
Open Access Books

Chapter 2: Understanding the core algorithms behind machine learning: Key learning algorithms explained

Priyambada Swain

Infosys Limited

1. Introduction to Machine Learning

Machine learning is a branch of artificial intelligence. It enables computers to automatically learn and improve from experience without being explicitly programmed [1]. Machine learning algorithms are usually classified into supervised, unsupervised, neural network-based, and reinforcement learning methods. Supervised learning methods build a mathematical model from a set of data that contains both the inputs and the desired outputs. The model is studied to predict the output values for given inputs [1-2]. These methods can also be used to find patterns (usually called classes) in data and assign new data points to one of the predefined classes. The goal of unsupervised learning is to find meaningful patterns in a set of data points that correspond closely to some intuitive notion of similarity. Neural network-based methods belong to the class of algorithms inspired by biological neural networks [2-4]. Neural networks replace a simulated neuron's simple threshold function with a real-valued differentiable function that gives the neuron the ability to output a real number. Reinforcement learning is an area of machine learning concerned with how an agent should take actions in an environment to maximize a reward signal.

2. Supervised Learning

Because of its ability to predict the outcome of never-seen-before data with various uncertainties, supervised learning is the most popular paradigm of machine learning. The objective in supervised learning is to achieve the best possible prediction performance on explanatory unlabelled testing data [5-6].

Linear regression is a classical algorithm for indicating the relationship between a scalar dependent variable y and one or more explanatory variables denoted together as a vector x in the form $y = w \cdot x + b$, where w and b are respectively the weight vector and the bias term. For a classification task, logistic regression employs a logistic transformation of the linear-regression formulation to represent the conditional probability of a certain class label y for an input x in a binary format. Provided with a set of explanatory variables, decision trees rely on thresholds to divide data into categories. Based on the concept of decision trees, random forests utilize several individual decision trees to solve a problem and first classify the input to the determined label by majority voting and then to the mean value in regression. The other useful supervised-learning algorithms include support vector machines (SVMs), artificial neural networks (ANNs), Gaussian processes (GPs), gradient boosting decision trees, and so on.

2.1. Linear Regression

Linear Regression is a supervised machine learning algorithm for predicting the value of a continuous target variable from multiple predictor variables. It belongs to the category of supervised learning, where the algorithm is trained on labelled data—in this case, a set of points on a graph—and used to estimate the relationships between the predictor variables and the target variable. The basic model assumes that the target variable y can be expressed as a linear combination of the predictor variables x_1, x_2, ..., x_p, weighted by coefficients β_1, β_2, ..., β_p and shifted by an intercept β_0:

$$y = \beta_0 + \beta_1 x_1 + ... + \beta\_p\, x\_p + \varepsilon$$

where ε represents the residual error, the difference between the predicted value and the true value.

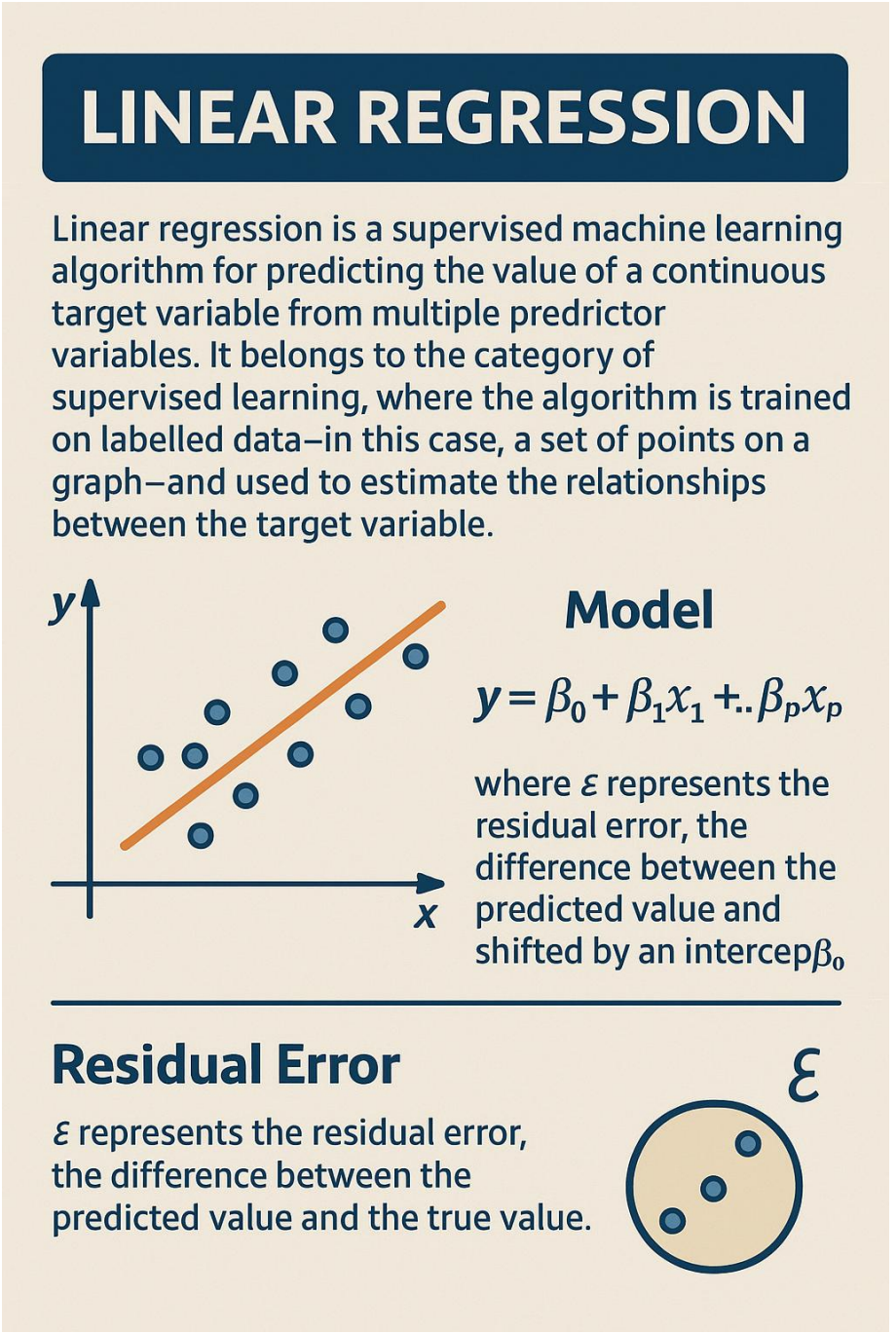

Fig 1 . Linear Regression

The coefficients in a linear regression model are of interest because they provide insight into the importance of each predictor variable in relation to the target variable. The goal in training is to find the set of coefficients β that minimizes

the residual error on the training data. This optimization is usually performed by minimizing the sum of the squares of the residual errors, known as the loss.

2.2. Logistic Regression

Although termed "logistic regression," this algorithm is largely employed for classification rather than regression. Logistic regression is designed for categorical dependent variables, typically dichotomous with two classes—though it can be adapted for multiple classes. The sigmoid function maps any real number input to the (0,1) interval, making it suitable for expressing a probability. Logistic regression can also be viewed as a single-layer neural network with a sigmoid activation. Cross-entropy, a measure reflecting the difference between predicted and actual labels, serves as the loss function, quantifying how observed labels diverge from the predictions.

The sigmoid function σ is defined by $\sigma(z) = 1/(1 + e^{-z})$. Logistic regression is well suited for the supervised classification of labeled data—where each data point is paired with a corresponding label—and can be trained with gradient descent. Unlike linear regression, which employs mean square error as a loss function, logistic regression uses cross-entropy, defined as $L = -\sum_i [o_i \log(p_i) + (1 - o_i) \log(1 - p_i)]$, where o_i represents the observed label and p_i the predicted probability.

2.3. Decision Trees

Decision trees constitute a methodology for supervised learning. Using a recursive partitioning procedure, they identify homogeneous sets of cases, whether classification categories or regression values, on the basis of a set of predictor variables [7,8]. Their decision tree analogy stems from the fact that they recursively split the predictors into two or more child nodes until the point that not enough data remain to continue. The algorithm then assigns a class to a child node (for classification) or a value (for regression).

These algorithms prove easy to understand, interpret, and implement with different types of variables for both classification and regression [9-12]. They are fast and involve only a small fraction of predictor variables at each split in the tree-building process. Although binary splits remain the most common, some algorithms search for splits into three or more child nodes. Besides that, they do not require any feature standardization [7,13-15]. Because of their relative simplicity, individual trees rarely provide the most accurate predictor. To tackle this shortcoming, ensemble models such as random forests and gradient boosting combine multiple trees into more accurate, robust predictors.

2.4. Random Forests

The Random Forest algorithm is an ensemble learning model based on Decision Trees. It operates by constructing multiple Decision Trees during training and combining their outputs—either through majority voting for classification tasks or by averaging for regression problems. Random Forests excel in handling both regression and classification problems, facilitating tasks from medicine to stock-market predictions, while also providing insights into feature importance [9,16-18].

The algorithm enhances the basic Decision Tree by adding a bagging technique. It generates subsets of the training data by sampling with replacement, ensuring each subset contains samples from all classes. The training subset for each tree is typically 70% of the data, with duplication ensured through sampling with replacement. Subsequently, a tree model is fitted on each constructed training subset, with the final output derived by merging predictions across all trees.

3. Unsupervised Learning

Algorithms belonging to the Unsupervised Learning family do not require training data with labeled outcomes. They are useful in circumstances when one wishes to discover structure and patterns in the data. Typically, these algorithms are applied to clustering and dimensionality reduction problems.

A clustering technique such as K-means creates disjoint groups of data points based on a distance measure. Each data point belongs to the cluster with the nearest mean in the feature space [2,19-20]. Hierarchical clustering organizes data points in a multi-level hierarchy, which can be represented by a dendrogram. The dendrogram can be cut to obtain the desired number of clusters. In the reducing dimensionality paradigm, principal component analysis (PCA) transforms the data variables into a new feature space defined by a set of orthogonal axes, along which the data variance is maximized.

3.1. K-Means Clustering

K-Means Clustering is perhaps the algorithm most simply and classically associated with unsupervised learning. It is a centroid-based clustering technique that attempts to discover clusters that minimize within-cluster distance and maximize between-cluster distance. Algorithmically, K-Means seeks to find k groups in the data, where k is several clusters predefined by the user.

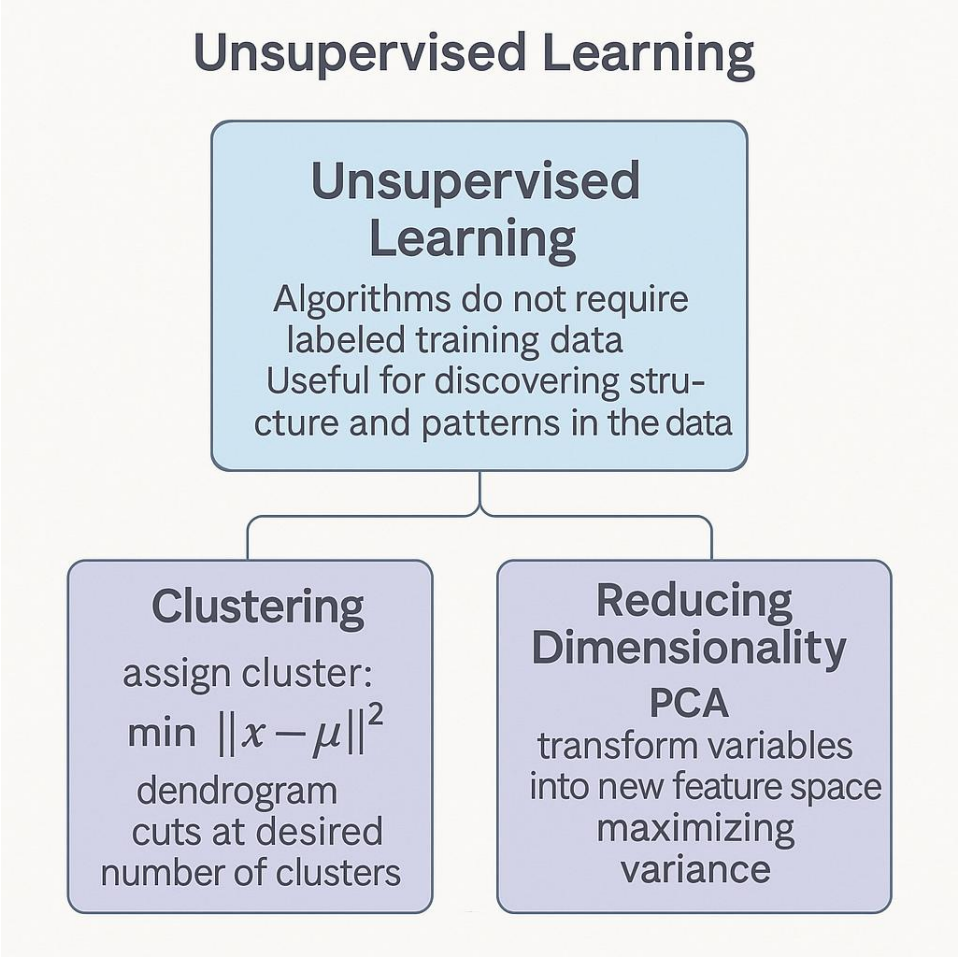

Fig 1. Unsupervised Learning

K-Means Clustering is relatively simple and straightforward to implement, but it is not without pitfalls. The number of clusters needs to be known in advance, making K-Means unsuitable for applications where explorative cluster discovery is the goal. It is also susceptible to the problem of local minima due to a suboptimal initialization. Several other issues plague K-Means; the algorithm performs badly on clusters with varying sizes, densities, and non-globular shapes.

3.2. Hierarchical Clustering

Hierarchical clustering produces a nested series of partitions, resembling the structure of a tree. A root cluster includes all objects, and the leaves each contain one object only. Each cluster that is not a leaf is a parent of other clusters [9,21-

23]. The parent is a superset of its clusters. Hierarchies are displayed in dendrograms.

There are two types. Agglomerative hierarchical clustering is a bottom-up approach. Every observation on its own is a cluster. Pairs of clusters are merged as one moves up the hierarchy. Divisive hierarchical clustering is a top-down approach. All observations start in one cluster, which is recursively split as one moves down the hierarchy. The latter procedure is not much used though, as it is Harder to compute. Nevertheless, it allows the use of arbitrary distances.

Indifferent to the linkage criterion, agglomerative hierarchical algorithms possess the locality property in the sense that decisions are irrevocable: Once two clusters have been merged, this cannot be undone in subsequent iterations. This causes the extrema linkage methods to be sensitive to anomalous points. Complete linkage tends to produce compact clusters. It reduces the effect of anomalous points, but it is still sensitive to noise. Single linkage can follow concavities or holes in the data [24-26]. However, it is also sensitive to anomalous points and tends to produce elongated clusters.

3.3. Principal Component Analysis (PCA)

A special case of Linear Discriminant Analysis has been Principal Component Analysis (PCA), which is, however, not a classification method but addresses dimensionality reduction. In many contexts, particularly those addressed by unsupervised learning techniques, data points are labelled. Principal Component Analysis also falls under this category and has several applications. The goal of PCA is to represent the data graphically and to reduce the dimensionality (the number of features) by removing redundant features. It achieves this by projecting the original n-dimensional dataset onto a k-dimensional space, where $k < n$, in a manner that the greatest variance by some projection of the data lies on the first coordinate (called the first principal component); the second greatest variance—as orthogonal to the first principle component as possible—on the second coordinate; the third greatest variance on the third coordinate; and so on. The resultant k-dimensional feature subspace will be made up of a set of k derived features, called the principal components [8,27-30].

PCA searches for the directions of maximum variability in high-dimensional data and projects the data onto a smaller-dimensional subspace while retaining most of the information. For example, in two-dimensional data, the first Principal Component (PC1) defines the first axis of maximum variance, and the second Principal Component (PC2) has the second largest variance under the constraint of being orthogonal to PC1. Figure 4 below shows the representation of the first

two components in PCA. This projected data can then be used in machine learning algorithms to achieve the benefits listed above. No information is lost when k = n, but by selecting k such that k < n, dimensionality reduction is completed by minimizing the loss of information.

4. Neural Networks

Deep learning techniques have come to dominate many applications of machine learning. Despite early exploration by researchers such as Hubert Dreyfus, who also highlighted significant limitations, the field continued to grow. The introduction of convolutional neural networks (CNNs) saw a resurgence of interest, and sequences of recurrent neural networks. More recently, the transformer architecture became the main deep learning model for natural-language processing [9,31-33].

Neural networks are deeply connected with two other important groups of machine-learning algorithms. Backpropagation is the fundamental algorithm for training neural networks. Reinforcement learning trains an agent within an environment to maximize its reward; when agents are represented by neural networks, the corresponding sub-field is called deep reinforcement learning.

4.1. Introduction to Neural Networks

Three main categories of artificial neural networks are currently in use, but these are bestconsidered subcategories of the broad family of machine-learning algorithms, since they lie at thecore of the more famous deep learning machine-learning methods. Although their architecturesand properties vary, all artificial neurons operate in a similar fashion: a neuron accepts multipleinputs and produces a single output, guided by a corresponding output function [34-36].

Supervised learning algorithms generate predictive models for labeled training data; hence, theycan locate patterns within the training examples but not in the properties of the examples themselves [3,37-39]. Similarly, unsupervised learning methods find patterns within the properties of the training setrather than within the data examples. In the reinforcement-learning approach, the learning agentcan adopt various heuristic policies toward solving the problem. In turn, the agent developsstrategies that optimize the summed reward for sequential decision-making, within an environmentthat can be fully or partially observable. A summary of choices in reinforcement learning isgiven, before introducing a second main category of machine-learning algorithms—the artificialneural networks—as a precursor to the section on Q-learning. For a detailed study, see.

4.2. Deep Learning Techniques

Convolutional Neural Networks Convolutional Neural Networks (CNNs) are a powerful type of neural networks widely and successfully used in image processing and computer vision. Mathematically, CNNs perform an approximation of a function f : , where the output is a multidimensional classification or regression. CNNs are designed to take advantage of 2D structure of the input data and use fewer parameters compared to fully connected networks.

The major modification that differentiates CNNs from fully connected networks is the neural architecture. Tensor inputs are connected with filters that slide with intervals of 1 over the image followed by nonlinear activation layers. Every activated unit generates a single value for a specific region of the image and therefore produces an activation map where the activations represent detected features from the image. Stacking several layers of these feature detectors allows the model to discover the structure and representation of the image automatically.

In addition to convolutional and nonlinear layers, there are also pooling and fully connected layers. Pooling layers lower the dimensionality of feature maps produced by detectors, while also keeping the most important information. In this way, pooling layers help to reduce memory and computational cost. Fully connected layers can still be included in CNNs and are located as the last layers of the models [36,40-42]. The softmax function at the last layer creates a score distribution over the output classes.

Recurrent Neural Networks Recurrent Neural Networks (RNNs) are a powerful neural architecture widely used in Natural Language Processing (NLP) and speech recognition. The recurrent structure is what makes an RNN different from a feed-forward network. The output depends not only on the input data at a given step but also on the previous step's output. The information is passed from the past to the future in a directed cycle (hence the temporal nature of RNNs). This directed cycle allows for the strong modeling of sequential data.

4.3. Convolutional Neural Networks (CNNs)

Convolutional Neural Networks (CNNs) are a particular type of neural network distinguished by the incorporation of one or more convolutional layers. These layers are inhospitable to inputs of arbitrary size. Thus, although fully connected networks can operate only on fixed-size inputs, Convolutional Neural Networks can be applied to inputs of diverse sizes, provided they contain at least one convolutional layer [40,43-44].

Convolutional layers were introduced to recognize features at spatial locations in the input image. This property is called spatial invariance—the ability to detect

the same feature, irrespective of its location. Spatial invariance arises from two attributes of convolutional layers: sparse connectivity and parameter sharing. Sparse connectivity significantly reduces the number of parameters in convolutional networks, enabling the construction and training of deep architectures without overfitting. Parameter sharing implies a translation-equivariant property or shifting property—shifting the input results in a corresponding shift in the feature maps.

4.4. Recurrent Neural Networks (RNNs)

Recurrent Neural Networks (RNNs) are a class of artificial neural networks in which connections between nodes form a directed graph along a temporal sequence. This structure allows RNNs to exhibit temporal dynamic behavior. Unlike feedforward neural networks, RNNs can use their internal state (memory) to process sequences of inputs, making them applicable to tasks such as unsegmented connected handwriting recognition or speech recognition.

In a standard RNN architecture, one or more hidden layers communicate among themselves through directed connections aiming to remember temporal sequences. The RNN may also have directed connections from the last hidden layer back to the input layer, allowing richness in the state model; however, hidden layers are usually connected back only to themselves, forming a self-loop that propagates information through time [3,45-48].

4.5. Transformers

The name Transformer derives from the architecture introduced in the seminal 2017 article by Vaswani et al. Transformer model captures similar context as a Bi-directional RNN can, but it can be trained more transparently and efficiently. The model is designed to transform one sequence into another through an attention mechanism that relates different positions of the input and output sequences, without relying on sequence-aligned RNNs or convolution.

A key point of a model based on self-attention is that all convolutional operations over the sequence used to compensate for the lack of sequence-alignment and caching in Transformer are eliminated, resulting in reduced training times. The Transformer architecture is divided into an encoder and a decoder, like the sequence-to-sequence model. Vanishing gradients in deep neural networks can, in principle, make them very hard to train [5,19,49-50]. Residual connections enable the model to maintain both the same distribution across all layers and allow lower layer representations to flow through the network. Batch Normalization uses the information learned from the dataset, helping the model to converge faster in deep neural networks.

4.6. Backpropagation

In the late 1980s, several networks shaped the development of deep learning techniques. Convolutional Neural Networks (CNNs) became the only deep learning method specializing in two-dimensional data, such as pictures and video. Recurrent Networks (RNNs) had loops in their structure, making them suitable for utilizing earlier information to influence later outputs in sequence data like speech or sentences. By 2017, Transformers evolved based on sequence-to-sequence models.

The Backpropagation algorithm, a supervised learning method suitable for training feedforward networks, existed prior to the development of deep learning. It is referred to as the fundamental algorithm for deep learning.

Reinforcement Learning

Deep reinforcement learning utilizes neural networks and backpropagation within the reinforcement learning framework [29,51-53].

Backpropagation

The algorithm, developed independently in the early 1960s and 1970s under the names "symmetric backpropagation algorithm" and "backpropagation of errors," targets feedforward neural networks using supervised learning. Indeed, it propagates backward the errors or differences between expected labels and outputs.

5. Reinforcement Learning

Reinforcement learning deals with training agents to make sequences of decisions in pursuit of a goal. An agent interacts with an environment, making observations through a set of precepts. Based on its current observations and experience, an agent will perform an action that will influence the state of the environment and lead to a state transition. An agent’s goal is to maximize the expected return, or, more informally, the amount of reward it experiences (Sutton and Barto 2018).

Markov decision processes (MDPs) provide a mathematical framework for modelling decision making in situations where outcomes are partly random and partly under the control of a decision maker. In reinforcement learning, the decision maker interacts with its environment through actions and receives feedback from the environment in terms of rewards. MDP algorithms find the

best actions to maximize the expected long-term cumulative rewards. The agent uses MDPs to model a sequential decision-making process. Q-learning is a popular algorithm for model-free reinforcement learning [54-56]. It computes an optimal action-selection policy to reach the maximum future discounted rewards by learning the action representation value Q from the previous experience. Policy gradient methods take a policy as a parameterized distribution and perform gradient ascent on the expected return. These methods have been widely applied to many learning tasks in spoken language understanding, computer vision, and natural language processing.

5.1. Markov Decision Processes

Reinforcement Learning is a paradigm in machine learning in which an artificial agent is trained by interacting with the provided environment, to accomplish a given objective. The agent needs to learn a mapping from states to actions which maximizes the cumulative reward. The central question is how to learn a good mapping to maximize the long-term cumulative reward, when the model does not know what the model of the environment is.

The Markov Decision Process is a formal mathematical concept that is used for modelling decision making situations when the outcomes are partly under the control of the agent and partly stochastic. It provides a foundation for many results in the domain of reinforcement learning. A Markov Decision Process is defined as a tuple (S, A, P, R), where S is a set of states, A is a set of actions, P is the set of transition probabilities and R is the reward function. A policy is a mapping from states to actions, and the objective is to find the policy which maximizes the expected sum of discounted rewards obtained in the long run, i.e., the optimal policy.

5.2. Q-Learning

Reinforcement learning is a machine-learning paradigm in which agents learn to make sequences of decisions. An agent takes actions in an environment that ultimately maximizes a specified reward signal (feedback) and tries to predict actions and them

5.3. Policy Gradients

Policy gradient methods differ from Q-learning techniques in the way that they explicitly define and optimize the policy. Instead of learning a value function, policy gradient approaches adjust the parameters of a policy represented by $\pi_\theta(a \mid s)$. The objective is to find parameters θ that maximize the expected cumulative reward $J(\theta)$ by applying some gradient update rule.

With the policy gradient theorem, a mathematically tractable gradient estimator for J(θ) can be derived and used together with gradient ascent methods. The aim of policy gradient methods is to tune the parameters θ of a policy $\pi_\theta(a \mid s)$ that directly maps states to actions, optimizing a performance measure J(θ). This performance measure is defined as the expected sum of discounted rewards over an episode: $J(\theta) = E_{\tau \sim p_\theta(\tau)} [r(\tau)] = \sum_{t=0}^{T} E_{s_t, a_t \sim p_\theta(s_t, a_t)} [\gamma^t r(s_t, a_t)]$, where τ represents a trajectory. Applying the logarithmic derivative trick allows the gradient of J(θ) with respect to θ to be expressed as an expectation over trajectories involving the gradient of the log-policy multiplied by the returns. Although unbiased, this estimate can suffer from high variance; a baseline function $b(s_t)$ is typically subtracted from the returns to reduce variance without introducing bias.

6. Comparison of Learning Methods

Machine Learning Algorithms differ from one another due to different strengths, weaknesses, and appropriate applications. The following sections summarize these differences.

Supervised Learning Algorithms predict or classify future or previously-unseen examples, based on labelled examples. Unsupervised Learning Algorithms find clusters and latent factors among their examples. Neural Networks and Reinforcement-Learning Algorithms interact with the environment in a sequential manner. They choose an action, observe the reward for it, and update their parameters accordingly. Comparing the quality of algorithms across these different methods amounts to comparing apples and oranges.

7. Applications of Machine Learning Algorithms

The diverse algorithmic landscape of machine learning enables its application across a multitude of domains. Perhaps the broadest classification scheme distinguishes algorithms by learning style, resulting in four categories: Supervised, Unsupervised, Neural Networks, and Reinforcement. Subcategories within the third and fourth often leverage Deep Learning techniques [5,6].

Supervised learning algorithms use knowledge of the category or numerical values of input data to build discriminant functions. The results then classify unknown input. Supervised techniques seek predictability in labelled data as a

basis for classification and prediction. Unsupervised methods strive to discover hidden structures that may exist in unlabelled data, concentrating on data description and data abstraction. The dominant functions implemented by unsupervised algorithms include clustering and dimensionality reduction.

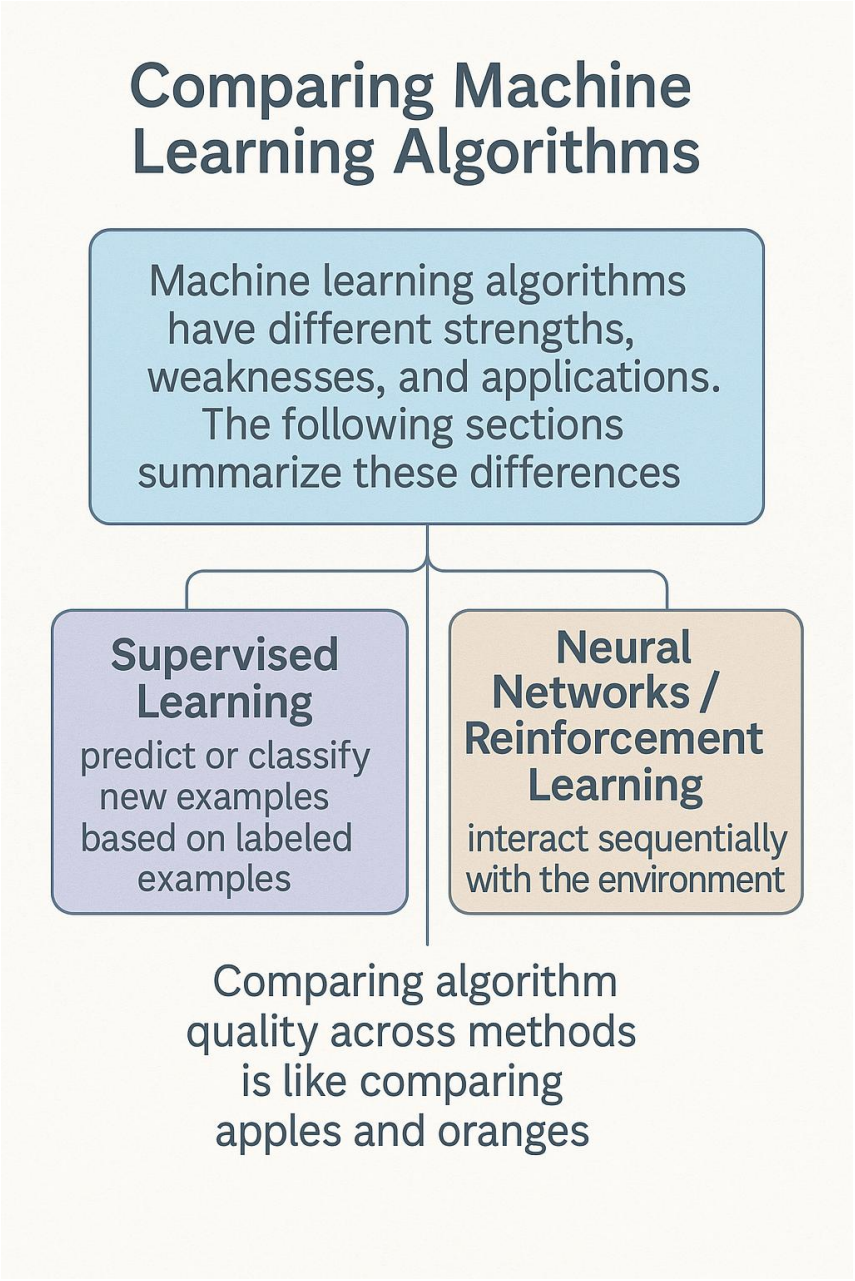

Fig 1. Machine Learning Algorithms

8. Challenges in Machine Learning

Multiple challenges complicate the use of machine learning. The selected model and methodology heavily influence its success and precise outcomes. The field has evolved in response to specific needs, with varying solutions developed by researchers and practitioners. As machine learning applications become more widespread, new users with limited expertise may find it difficult to keep pace with developments across all categories and domains.

Other obstacles are inherent to the applications themselves, such as the availability of appropriately formatted data with sufficient characteristics and quality to build accurate and robust models. Strong changes and trends periodically impact learning algorithms, including the need for precise interpretability, noise reduction, robust and scalable models, and the integration of prior knowledge for lifelong, transfer, and semi-supervised learning.

9. Future Trends in Machine Learning

Despite the significant research efforts in the past decades, machine learning is far from being solved. In fact, future trends have been already identified within the broader field of data analysis and computer-science problems. Together with the gradual elimination of formal methodologies that aim at fitting probability distributions—not only the normal ones—but also nonparametric types, more and more learning methods are emerging for overcoming architecture problems; for analysing many data; for developing generalizations and abstraction capabilities; for increasing the ability to find sensitive patterns among the features during the training phase; finally, for making the best use of prior knowledge incorporated in the training phase [7,9]. Furthermore, hybrid solutions have been also proposed for exploiting the best components of different learning methods: It is likely that many other hybrid combinations will be developed in the future for facing still unexplored problems.

Hybrid algorithms are often used for solving language modelling problems in speech recognition. According to this principle, language models should be learned by hybrid probabilistic automata–probabilistic neural networks combined with a very-long-memory mechanism for storing the recent context of the conversation, which is necessary for calculating the exact transition probabilities. Other interesting directions have been proposed for actively training different neural networks for different problems in different time. This

type of architecture is particularly suited to multiple speech recognition problems, where different neural networks are devoted to different speakers, whose recognition difficulty varies in time.

10. Conclusion

Machine learning algorithms have root in computer science, mathematics, and statistics. The first algorithm discussed appears in Computer Methods for Mathematical Computations by Forsythe, Malcolm, and Moler, published in 1977. Many other algorithms were developed by researchers in computer science and statistics over the following decades. Since the adoption of the name “deep learning” in 2006, the field has been dominated by study of neural networks.

The body of work is organized by the four primary machine learning categories. Supervised Learning defines the most popular models presenting Linear Regression, Logistic Regression, Decision Tree, and Random Forest. Unsupervised Learning presents the models used for clustering and dimension reduction through K-Means, Hierarchical Clustering, and Principal Component Analysis. Deep Learning Techniques summarizes the development of Convolutional Neural Networks, Recurrent Neural Networks, and Transformers. Backpropagation, the fundamental algorithm used to train neural networks, is also explained. Reinforcement Learning then defines the agent-environment interaction from a Markov Decision Process, presenting Q-Learning and Policy Gradients as the most popular methods for deciding behaviour. Each category − Supervised, Unsupervised, Deep Learning, and Reinforcement − are then compared.

References

[1] Tapeh AT, Naser MZ. Artificial intelligence, machine learning, and deep learning in structural engineering: a scientometrics review of trends and best practices. Archives of Computational Methods in Engineering. 2023 Jan;30(1):115-59.

[2] Panda SP. Artificial Intelligence Across Borders: Transforming Industries Through Intelligent Innovation. Deep Science Publishing; 2025 Jun 6.

[3] Theodosiou AA, Read RC. Artificial intelligence, machine learning and deep learning: Potential resources for the infection clinician. Journal of Infection. 2023 Oct 1;87(4):287-94.

[4] Kersting K. Machine learning and artificial intelligence: two fellow travelers on the quest for intelligent behavior in machines. Frontiers in big Data. 2018 Nov 19;1:6.

[5] Panda SP, Muppala M, Koneti SB. The Contribution of AI in Climate Modeling and Sustainable Decision-Making. Available at SSRN 5283619. 2025 Jun 1.
[6] Goldenberg SL, Nir G, Salcudean SE. A new era: artificial intelligence and machine learning in prostate cancer. Nature Reviews Urology. 2019 Jul;16(7):391-403.
[7] Shivadekar S. Artificial Intelligence for Cognitive Systems: Deep Learning, Neuro-symbolic Integration, and Human-Centric Intelligence. Deep Science Publishing; 2025 Jun 30.
[8] Sultan AS, Elgharib MA, Tavares T, Jessri M, Basile JR. The use of artificial intelligence, machine learning and deep learning in oncologic histopathology. Journal of Oral Pathology & Medicine. 2020 Oct;49(9):849-56.
[9] Guo K, Yang Z, Yu CH, Buehler MJ. Artificial intelligence and machine learning in design of mechanical materials. Materials Horizons. 2021;8(4):1153-72.
[10] Panda SP. Securing 5G Critical Interfaces: A Zero Trust Approach for Next-Generation Network Resilience. In2025 12th International Conference on Information Technology (ICIT) 2025 May 27 (pp. 141-146). IEEE.
[11] Rane J, Chaudhari RA, Rane NL. Data Privacy and Information Security in Deep Learning Applications: Risk Assessment and Patient Safety Protocols for Big Data Analytics. Ethical Considerations and Bias Detection in Artificial Intelligence/Machine Learning Applications. 2025 Jul 10:54.
[12] Mohapatra PS. Artificial Intelligence-Driven Test Case Generation in Software Development. Intelligent Assurance: Artificial Intelligence-Powered Software Testing in the Modern Development Lifecycle. 2025 Jul 27:38.
[13] Mich L. Artificial intelligence and machine learning. InHandbook of e-Tourism 2020 Sep 1 (pp. 1-21). Cham: Springer International Publishing.
[14] Mohapatra PS. Artificial Intelligence-Powered Software Testing: Challenges, Ethics, and Future Directions. Intelligent Assurance: Artificial Intelligence-Powered Software Testing in the Modern Development Lifecycle. 2025 Jul 27:163.
[15] Suura SR. Integrating Artificial Intelligence, Machine Learning, and Big Data with Genetic Testing and Genomic Medicine to Enable Earlier, Personalized Health Interventions. Deep Science Publishing; 2025 Apr 13.
[16] Panda SP. Relational, NoSQL, and Artificial Intelligence-Integrated Database Architectures: Foundations, Cloud Platforms, and Regulatory-Compliant Systems. Deep Science Publishing; 2025 Jun 22.
[17] Zhu S, Gilbert M, Chetty I, Siddiqui F. The 2021 landscape of FDA-approved artificial intelligence/machine learning-enabled medical devices: an analysis of the characteristics and intended use. International journal of medical informatics. 2022 Sep 1;165:104828.
[18] Mukhamediev RI, Popova Y, Kuchin Y, Zaitseva E, Kalimoldayev A, Symagulov A, Levashenko V, Abdoldina F, Gopejenko V, Yakunin K, Muhamedijeva E. Review of artificial intelligence and machine learning technologies: classification, restrictions, opportunities and challenges. Mathematics. 2022 Jul 22;10(15):2552.
[19] Gerke S, Babic B, Evgeniou T, Cohen IG. The need for a system view to regulate artificial intelligence/machine learning-based software as medical device. NPJ digital medicine. 2020 Apr 7;3(1):53.
[20] Paramesha M, Rane N, Rane J. Big data analytics, artificial intelligence, machine learning, internet of things, and blockchain for enhanced business intelligence. Artificial Intelligence,

Machine Learning, Internet of Things, and Blockchain for Enhanced Business Intelligence (June 6, 2024). 2024 Jun 6.
[21] Arel I, Rose DC, Karnowski TP. Deep machine learning-a new frontier in artificial intelligence research [research frontier]. IEEE computational intelligence magazine. 2010 Oct 18;5(4):13-8.
[22] Ongsulee P. Artificial intelligence, machine learning and deep learning. In2017 15th international conference on ICT and knowledge engineering (ICT&KE) 2017 Nov 22 (pp. 1-6). IEEE.
[23] Panch T, Szolovits P, Atun R. Artificial intelligence, machine learning and health systems. Journal of global health. 2018 Oct 21;8(2):020303.
[24] Das S, Dey A, Pal A, Roy N. Applications of artificial intelligence in machine learning: review and prospect. International Journal of Computer Applications. 2015 Jan 1;115(9).
[25] Balyen L, Peto T. Promising artificial intelligence-machine learning-deep learning algorithms in ophthalmology. The Asia-Pacific Journal of Ophthalmology. 2019 May 1;8(3):264-72.
[26] Michalski RS, Carbonell JG, Mitchell TM, editors. Machine learning: An artificial intelligence approach. Springer Science & Business Media; 2013 Apr 17.
[27] Soori M, Arezoo B, Dastres R. Artificial intelligence, machine learning and deep learning in advanced robotics, a review. Cognitive Robotics. 2023 Jan 1;3:54-70.
[28] Ghahramani Z. Probabilistic machine learning and artificial intelligence. Nature. 2015 May 28;521(7553):452-9.
[29] Jakhar D, Kaur I. Artificial intelligence, machine learning and deep learning: definitions and differences. Clinical and experimental dermatology. 2020 Jan 1;45(1):131-2.
[30] Siau K, Wang W. Building trust in artificial intelligence, machine learning, and robotics. Cutter business technology journal. 2018;31(2):47.
[31] Panda SP. The Evolution and Defense Against Social Engineering and Phishing Attacks. International Journal of Science and Research (IJSR). 2025 Jan 1.
[32] Kühl N, Schemmer M, Goutier M, Satzger G. Artificial intelligence and machine learning. Electronic Markets. 2022 Dec;32(4):2235-44.
[33] Karthikeyan A, Priyakumar UD. Artificial intelligence: machine learning for chemical sciences. Journal of Chemical Sciences. 2022 Mar;134(1):2.
[34] Helm JM, Swiergosz AM, Haeberle HS, Karnuta JM, Schaffer JL, Krebs VE, Spitzer AI, Ramkumar PN. Machine learning and artificial intelligence: definitions, applications, and future directions. Current reviews in musculoskeletal medicine. 2020 Feb;13(1):69-76.
[35] Shivadekar S, Halem M, Yeah Y, Vibhute S. Edge AI cosmos blockchain distributed network for precise ablh detection. Multimedia tools and applications. 2024 Aug;83(27):69083-109.
[36] Helm JM, Swiergosz AM, Haeberle HS, Karnuta JM, Schaffer JL, Krebs VE, Spitzer AI, Ramkumar PN. Machine learning and artificial intelligence: definitions, applications, and future directions. Current reviews in musculoskeletal medicine. 2020 Feb;13(1):69-76.
[37] Panda SP. Augmented and Virtual Reality in Intelligent Systems. Available at SSRN. 2021 Apr 16.
[38] Bhat M, Rabindranath M, Chara BS, Simonetto DA. Artificial intelligence, machine learning, and deep learning in liver transplantation. Journal of hepatology. 2023 Jun 1;78(6):1216-33.

[39] Tyagi AK, Chahal P. Artificial intelligence and machine learning algorithms. InChallenges and applications for implementing machine learning in computer vision 2020 (pp. 188-219). IGI Global Scientific Publishing.
[40] Shivadekar S, Kataria DB, Hundekar S, Wanjale K, Balpande VP, Suryawanshi R. Deep learning based image classification of lungs radiography for detecting covid-19 using a deep cnn and resnet 50. International Journal of Intelligent Systems and Applications in Engineering. 2023;11:241-50.
[41] Kühl N, Goutier M, Hirt R, Satzger G. Machine learning in artificial intelligence: Towards a common understanding. arXiv preprint arXiv:2004.04686. 2020 Mar 27.
[42] Rane J, Chaudhari RA, Rane NL. Data Analysis and Information Processing Frameworks for Ethical Artificial Intelligence Implementation: Machine-Learning Algorithm Validation in Clinical Research Settings. Ethical Considerations and Bias Detection in Artificial Intelligence/Machine Learning Applications. 2025 Jul 10:192.
[43] Neapolitan RE, Jiang X. Artificial intelligence: With an introduction to machine learning. CRC press; 2018 Mar 12.
[44] Rane NL, Paramesha M, Choudhary SP, Rane J. Artificial intelligence, machine learning, and deep learning for advanced business strategies: a review. Partners Universal International Innovation Journal. 2024 Jun 25;2(3):147-71.
[45] Rubinger L, Gazendam A, Ekhtiari S, Bhandari M. Machine learning and artificial intelligence in research and healthcare. Injury. 2023 May 1;54:S69-73.
[46] Gupta R, Srivastava D, Sahu M, Tiwari S, Ambasta RK, Kumar P. Artificial intelligence to deep learning: machine intelligence approach for drug discovery. Molecular diversity. 2021 Aug;25(3):1315-60.
[47] Entezari A, Aslani A, Zahedi R, Noorollahi Y. Artificial intelligence and machine learning in energy systems: A bibliographic perspective. Energy Strategy Reviews. 2023 Jan 1;45:101017.
[48] Mohapatra PS. Artificial Intelligence and Machine Learning for Test Engineers: Concepts in Software Quality Assurance. Intelligent Assurance: Artificial Intelligence-Powered Software Testing in the Modern Development Lifecycle. 2025 Jul 27:17.
[49] Bini SA. Artificial intelligence, machine learning, deep learning, and cognitive computing: what do these terms mean and how will they impact health care?. The Journal of arthroplasty. 2018 Aug 1;33(8):2358-61.
[50] Nuka ST. Next-Frontier Medical Devices and Embedded Systems: Harnessing Biomedical Engineering, Artificial Intelligence, and Cloud-Powered Big Data Analytics for Smarter Healthcare Solutions. Deep Science Publishing; 2025 Jun 6.
[51] Hyder Z, Siau K, Nah F. Artificial intelligence, machine learning, and autonomous technologies in mining industry. InResearch Anthology on Cross-Disciplinary Designs and Applications of Automation 2022 (pp. 478-492). IGI Global Scientific Publishing.
[52] Maguluri KK. Machine learning algorithms in personalized treatment planning. How Artificial Intelligence is Transforming Healthcare IT: Applications in Diagnostics, Treatment Planning, and Patient Monitoring. 2025 Jan 10:33.
[53] Mijwil MM, Aggarwal K, Sonia S, Al-Mistarehi AH, Alomari S, Gök M, Zein Alaabdin AM, Abdulrhman SH. Has the Future Started? The Current Growth of ArtificialIntelligence, Machine Learning, and Deep Learning. Iraqi Journal for Computer Science and Mathematics. 2022;3(1):13.

[54] Panda S. Observability in DevOps: Integrating AWS X-Ray, CloudWatch, and Open Telemetry. International Journal of Computer Application. 2025 Jan 1.
[55] Dunjko V, Briegel HJ. Machine learning & artificial intelligence in the quantum domain: a review of recent progress. Reports on Progress in Physics. 2018 Jun 19;81(7):074001.
[56] Cioffi R, Travaglioni M, Piscitelli G, Petrillo A, De Felice F. Artificial intelligence and machine learning applications in smart production: Progress, trends, and directions. Sustainability. 2020 Jan 8;12(2):492.