DISCUSSION PUBLISHERS
Research Center · Open Access Edition

Machine Learning from Theory to Algorithms: An Overview

Jafar Ahmad Abed Alzubi, Anand Nayyar, Akshi Kumar

Discussion Research ID	DPR-2026-1150
Year	2018
Field	Computer Science
DOI	10.1088/1742-6596/1142/1/012012
Author ORCID	0000-0001-6724-1421
Licence	CC-BY
Original source	Journal of Physics Conference Series
Permanent link	https://discussionpublishers.com/research/606097fc-1d41-4e24-b65e-9915932ce021
Publisher	Discussion Publishers discussionpublishers.com

Abstract

The current SMAC (Social, Mobile, Analytic, Cloud) technology trend paves the way to a future in which intelligent machines, networked processes and big data are brought together. This virtual world has generated vast amount of data which is accelerating the adoption of machine learning solutions & practices. Machine Learning enables computers to imitate and adapt human-like behaviour. Using machine learning, each interaction, each action performed, becomes something the system can learn and use as experience for the next time. This work is an overview of this data analytics method which enables computers to learn and do what comes naturally to humans, i.e. learn from experience. It includes the preliminaries of machine learning, the definition, nomenclature and applications' describing it's what, how and why. The technology roadmap of machine learning is discussed to understand and verify its potential as a market & industry practice. The primary intent of this work is to give insight into why machine learning is the future.

Republished by Discussion Publishers under the original CC-BY licence. Copyright remains with the original authors; please cite the original work.

Cite as: Jafar Ahmad Abed Alzubi, Anand Nayyar, Akshi Kumar (2018). Machine Learning from Theory to Algorithms: An Ove. Discussion

Journal of Physics:
Conference Series

PURPOSE-LED PUBLISHING™

PAPER • **OPEN ACCESS**

Machine Learning from Theory to Algorithms: An Overview

To cite this article: Jafar Alzubi *et al* 2018 *J. Phys.: Conf. Ser.* **1142** 012012

View the article online for updates and enhancements.

You may also like

- Classification of Rock Mineral in Field X based on Spectral Data (SWIR & TIR) using Supervised Machine Learning Methods
S A Pane and F M H Sihombing
- 26th IAHR Symposium on Hydraulic Machinery and Systems
Yulin Wu, Zhengwei Wang, Shuhong Liu et al.
- Implementation of Naïve Bayes Classifier-based Machine Learning to Predict and Classify New Students at Matana University
Simon Prananta Barus

Enabling Electrochemistry Inside and Outside the Lab

PalmSens

Discover our diverse range of Potentiostats engineered to support your research wherever it takes you

- Intuitive software for Windows and Android
- Proven in over 4000 scientific publications
- Single-channel and multi-channel
- Benchtop, portable and wearable

Discover More



This content was downloaded from IP address 104.23.243.19 on 18/11/2026 at 20:08

Discussion Publishers edition DPR-2026-1150 https://discussionpublishers.com/research/606097fc-1d41-4e24-b65e-9915932ce021

IOP Conf. Series: Journal of Physics: Conf. Series 1142 (2018) 012012 doi:10.1088/1742-6596/1142/1/012012

Machine Learning from Theory to Algorithms: An Overview

Jafar Alzubi[1], Anand Nayyar[2], Akshi Kumar[3]

[1]School of Engineering, Al-Balqa Applied University, Jordan

[2]Graduate School, Duy Tan University, Da Nang, Viet Nam

[3]Department of CSE, Delhi Technological University, Delhi, India

j.zubi@bau.edu.jo, anandnayyar@duytan.edu.vn, akshikumar@dce.ac.in

Abstract. The current SMAC (Social, Mobile, Analytic, Cloud) technology trend paves the way to a future in which intelligent machines, networked processes and big data are brought together. This virtual world has generated vast amount of data which is accelerating the adoption of machine learning solutions& practices. Machine Learning enables computers to imitate and adapt human-like behaviour. Using machine learning, each interaction, each action performed, becomes something the system can learn and use as experience for the next time. This work is an overview of this data analytics method which enables computers to learn and do what comes naturally to humans, i.e. learn from experience. It includes the preliminaries of machine learning, the definition, nomenclature and applications' describing it's what, how and why. The technology roadmap of machine learning is discussed to understand and verify its potential as a market & industry practice. The primary intent of this work is to give insight into why machine learning is the future.

Keywords: Machine Learning, Supervised Learning, Unsupervised Learning, Algorithms, Instant Learning, Ensemble Learning

1. Introduction

Learning as a generic process is about acquiring new, or modifying existing, behaviors, values, knowledge, skills, or preferences. Behaviorism, Cognitivism, Constructivism, Experientialism and Social Learning define the theory of personal learning, i.e. how humans learn. Machines rely on data contrary to what comes naturally to humans: learning from experience. At the very fundamental level machine learning (ML) is a category of artificial intelligence that enables computers to think and learn on their own. It is all about making computers modify their actions in order to improve the actions to attain more accuracy, where accuracy is measured in terms of the number of times times the chosen actions results into correct ones.

Researchers have formally defined ML across pertinent literature. The term was coined by Arthur Samuel in 1959, who defined ML as a field of study that provides learning capability to computers without being explicitly programmed [1].More recently, Tom Mitchell gave a "well-posed" definition that has proven more useful to engineering set-up: "A computer program is said to learn from experience E with respect to some task T and some performance measure P, if its performance on T, as measured by P, improves with experience E [2]"

Content from this work may be used under the terms of the Creative Commons Attribution 3.0 licence. Any further distribution of this work must maintain attribution to the author(s) and the title of the work, journal citation and DOI.
Published under licence by IOP Publishing Ltd

Discussion Publishers edition DPR-2026-1150 https://discussionpublishers.com/research/606097fc-1d41-4e24-b65e-9915932ce021

Machine learning is a multi-disciplinary field having a wide-range of research domains reinforcing its existence. These are as shown in the following figure 1.

The simulation of ML models is significantly related to Computational Statistics whose main aim is to focus on making predictions via computers. It is also co-related to Mathematical Optimization which relates models, applications and frameworks to the field of statistics.

Real world problems have high complexity which make them excellent candidates for application of ML. Machine learning can be applied to various areas of computing to design and program explicit algorithms with high performance output, for example, email spam filtering, fraud detection on social network, online stock trading, face & shape detection, medical diagnosis, traffic prediction, character recognition and product recommendation amongst others. The self-driving Google cars, Netflix showcasing the movies and shows a person might like, online recommendation engines—like friend suggestions on Facebook, "more items to consider" and "get yourself a little something" on Amazon, and credit card fraud detection, are all real-world examples of application of machine learning.

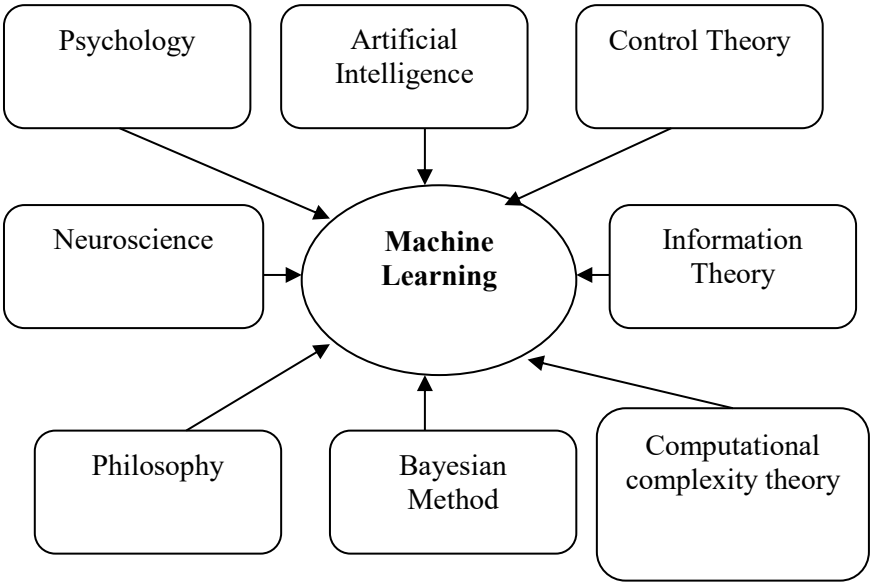


Fig.1. The Multi-disciplinary ML

The main objective of this work is to give an overview of development of Machine Learning to the present day, various machine learning algorithms, applications and challenges.

This paper presents an overview of machine learning, its origin and development through the decades to the present age. The study is organized as follows: Section 2 explains the problems in data science for which ML approaches offer a solution. Section 3 gives the evolution of ML to the present day. Section 4 describes the generic model of machine learning followed by the explanation of the machine learning process in section 5. Section 6 and 7 explain the various machine learning paradigms and algorithms respectively while section 8 presents a brief summary of the challenges faced by ML and its future scope as a discussion followed by conclusion of the study in section 9.

2. Data Science problems and Machine Learning

Machine learning is required to make the computers sophisticatedly perform the task without any intervention of human beings on the basis of learning and constantly increasing experience to understand the problem complexity and need for adaptability.

Discussion Publishers edition DPR-2026-1150 https://discussionpublishers.com/research/606097fc-1d41-4e24-b65e-9915932ce021

- *Tasks performed by Human Beings:* There are lots of tasks performed on day to day basis by human beings, but the main concern is that to perform the tasks perfectly and to perform under well-defined program. Examples: Cooking, Driving, Speech Recognition.
- *Tasks beyond Human Capabilities:* Another category of tasks includes which can be done by machine learning in effective manner is analysis of large and complex data sets like Remote Sensing, Weather forecasting, Ecommerce, Web search etc. With large amounts of data, it becomes really complex for human beings to predict meaningful data.

Machine learning has proven capabilities to inherently solve the problems of data science. Hayashi and Chikio [3] define data science as, "a concept to unify statistics, data analysis, machine learning and their related methods in order to understand and analyze actual phenomena" with data". Before taking to problem solving, the problem must be categorized suitably so that the most appropriate machine learning algorithm can be applied to it. Any problem in data science can be grouped in one of the following five categories as shown in Figure 2.

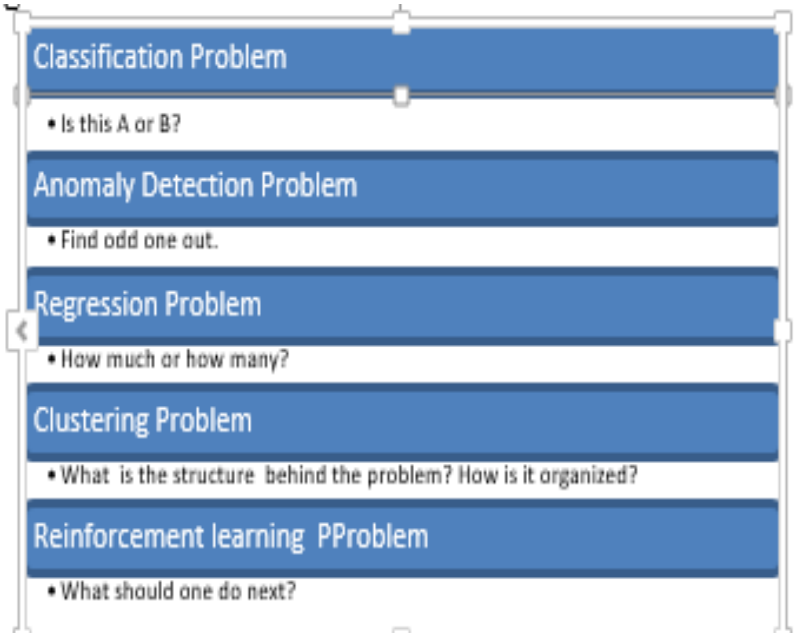


Fig 2. Types of Problems

Thus depending on the type of problem, an appropriate machine learning approach can be applied. The various categories are explained below:

- *Classification Problem*- A problem in which the output can be only one of a fixed number of output classes known apriori like Yes/No, True/False, is called a classification problem. Depending on the number of output classes, the problem can be a binary or multi-class classification problem.
- *Anomaly Detection Problem*- Problems which analyze a certain pattern and detects changes or anomalies in the pattern fall under this category. For example, credit card companies use anomaly detection algorithms to find deviation from the usual transaction behavior of their client and raises alerts whenever there is an unusual transaction. Such problems deal with finding out the outliers.

Discussion Publishers edition DPR-2026-1150 https://discussionpublishers.com/research/606097fc-1d41-4e24-b65e-9915932ce021

- *Regression Problem*- Regression algorithms are used to deal with problems with continuous and numeric output. These are usually used for problems that deal with questions like, 'how much' or 'how many'.
- *Clustering Problem*- Clustering falls under the category of unsupervised learning algorithms. These algorithms try to learn structures within the data and attempt to make clusters based on the similarity in the structure of data. The different classes or clusters are then labeled. The algorithm, when trained, puts new unseen data in one of the clusters.
- *Reinforcement Problem*-Reinforcement algorithms are used when a decision is to be made based on past experiences of learning. The machine agent learns the behaviour using trial and error sort of interaction with the continuously changing environment. It provides a way to program agents using the concept of rewards and penalties without specifying how the task is to be accomplished. Game playing programs and programs for temperature control are some popular examples using reinforcement learning.

3. Development of Machine Learning

The words, Artificial Intelligence and Machine Learning are not new. They have been researched, utilized, applied and re-invented by computer scientists, engineers, researchers, students and industry professionals for more than 60 years. The mathematical foundation of machine learning lies in algebra, statistics, and probability. Serious development of Machine Learning and Artificial Intelligence began in 1950's and 1960's with the contributions of researchers like Alan Turing, John McCarthy, Arthur Samuels, Alan Newell and Frank Rosenblatt. Samuel proposed the first working machine learning model on Optimizing Checkers Program. Rosenblatt created Perceptron, a popular machine learning algorithm based on biological neurons which laid the foundation of Artificial Neural Network [4, 5, 6]. The following table 1 depicts the illustrious, expansive and practical development of machine learning.

Table 1. Development of ML

1950	Alan Turning created "Turning Test" to check a machine's intelligence. In order to pass the Turning Test, the machine should be able to convince humans that there they are actually talking to a human and not a machine.
1952	Samuel created a highly capable learning algorithm than can play the game of Checkers with itself and get self-trained.
1956	Martin Minsky and John McCarty with Claude Shannon and Nathan Rochester organized a conference in Dartmouth in 1956 where actually Artificial Intelligence was born.
1958	Frank Rosenblatt created Perceptron, which laid the foundation stone for the development of Artificial Neural Network (ANN).
1967	The Nearest Neighbor Algorithm was proposed which could be used for "Pattern Recognition".
1979	Stanford University students developed "Stanford Cart", a sophisticated robot that could navigate around a room and avoid obstacles in its path.
1981	Explanation Based Learning (EBL) was proposed by Gerald Dejong, whereby, a computer can analyze the training data and create rules for discarding useless data [7]
1985	NetTalk was invented by Terry Sejnowski, [8] which learnt to pronounce English words in the same manner that children learn.
1990s	The focus of Machine Learning shifted from Knowledge-driven to Data Driven. Machine Learning was implemented to analyze large chunks of data and derive conclusions from it [9]
1997:	IBM invented the Deep Blue computer which was able to beat World Chess Champion Gary Kasparov.
2006	The term "Deep Learning" was coined by Geoffery Hinton which referred to a new architecture of neural networks that used multiple layers of neurons for learning.
2011	IBM's Watson, built to answer questions posed in a natural language, defeats a Human Competitor at Jeopardy Game.
2012	Jeff Dean from Google, developed GoogleBrain, which isa Deep Neural Network to detect patterns in Videos and Images.
2014	Facebook invented the "DeepFace" algorithm based on Deep Neural Networks capable of

Discussion Publishers edition DPR-2026-1150 https://discussionpublishers.com/research/606097fc-1d41-4e24-b65e-9915932ce021

	recognizing human faces in photos.
2015	Amazon proposed its own Machine Learning Platform. Microsoft created "Distributed Machine Learning Toolkit" for efficient distribution of machine learning problems to multiple computers to work parallel to find a solution [10,11]. Elon Musk and Sam Altman, created a non-profit organization- OoeaAI, with the objective of using Artificial Intelligence to serve human beings.
2016	Google proposed DeepMind which is regarded as the most complex Board Game. Google AlphaGo program becomes the first Computer Go program to beat a professional human player. It is based on the combination of machine learning and tree searching techniques [12].
2017	Google proposed Google Lens, Google Clicks, Google Home Mini and Google Nexus based phones which use Machine Learning and Deep Learning Algorithms. Nvidia proposed NVIDIA GPUs- The Engine of Deep Learning. Apple proposed Home Pod which is a Machine Learning Interactive device.

4. The Generic Model of ML

ML is used to solve various problems that require learning on the part of the machine. A learning problem has three features:

- Task classes (The task to be learnt)
- Performance measure to be improved
- The process of gaining experience

For example, in a game of checkers, the learning problem can be defined as:

- Task T: Playing the game
- Performance Measure P: number of games won against the opponent.
- Experience E: practicing via playing games against itself and consistently improving the performance.

The generic model of machine learning consists of six components independent of the algorithm adopted. The following figure 3 depicts these primary components.

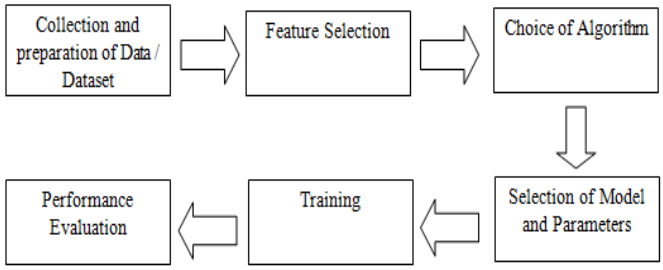


Fig 3. Components of a Generic ML model

Each component of the model has a specific task to accomplish as described next.

i. **Collection and Preparation of Data**: The primary task of in the machine learning process is to collect and prepare data in a format that can be given as input to the algorithm. A large amount may be available for any problem. Web data is usually unstructured and contains a lot of noise, i.e., irrelevant data as well as redundant data. Hence the data needs to be cleaned and pre-processed to a structured format.

ii. **Feature Selection**: The data obtained from the above step may contain numerous features, not all of which would be relevant to the learning process. These features need to be removed and a subset of the most important features needs to be obtained.

Discussion Publishers edition DPR-2026-1150 https://discussionpublishers.com/research/606097fc-1d41-4e24-b65e-9915932ce021

iii. **Choice of Algorithm**: Not all machine learning algorithms are meant for all problems. Certain algorithms are more suited to a particular class problem as explained in the previous section. Selecting the best machine learning algorithm for the problem at hand is imperative in getting the best possible results. The various ML algorithms are discussed in Section 6.
iv. **Selection of Models and Parameters**: Most of machine learning algorithms require some initial manual intervention for setting the most appropriate values of various parameters.
v. **Training**: After selecting the appropriate algorithm and suitable parameter values, the model needs to be trained using a part of the dataset as training data.
vi. **Performance Evaluation**: Before real-time implementation of the system, the model must be tested against unseen data to evaluate how much has been learnt using various performance parameters like accuracy, precision and recall.

5. Machine Learning Paradigms

Depending on how an algorithm is being trained and on the basis of availability of the output while training, machine learning paradigms can be classified into ten categories. These include: supervised learning, semi-supervised learning, unsupervised learning, reinforcement learning, evolutionary learning, ensemble learning, artificial neural network, Instance-based learning, dimensionality reduction algorithms and hybrid learning [4, 5, 6, 13,14,15,16 17]. Each of these paradigms is explained in the following sub-sections.

5.1. Supervised Learning

Under supervised learning, a set of examples or training modules are provided with the correct outputs and on the basis of these training sets, the algorithm learns to respond more accurately by comparing its output with those that are given as input. Supervised learning is also known as learning via examples or learning from exemplars. The following figure 4 explains the concept.

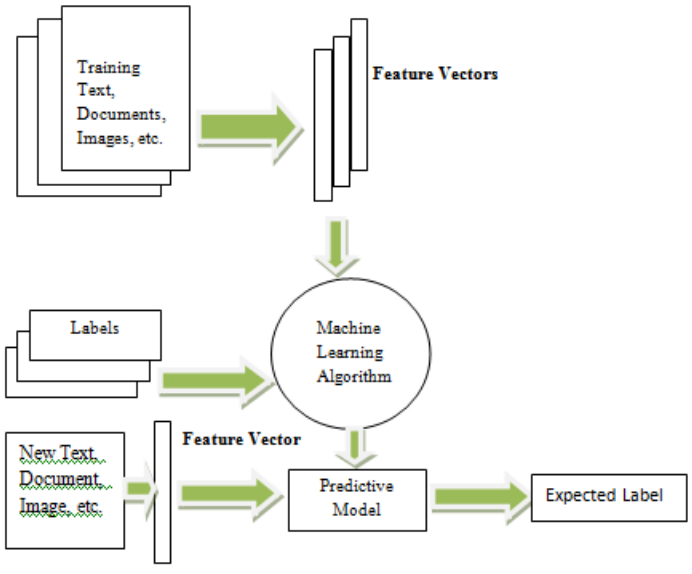


Fig. 4. Supervised Learning [4]

Supervised learning finds applications in prediction based on historical data. For example: to predict the Iris species given a set of its flower measurements or a recognition system that determines whether an object is a galaxy, a quasar or a star given a colored image of an object through a telescope, or given an e-commerce surfing history of a person, recommendation of the products by e-commerce websites

Discussion Publishers edition DPR-2026-1150 https://discussionpublishers.com/research/606097fc-1d41-4e24-b65e-9915932ce021

[4]. Supervised learning tasks can be further categorized as classification tasks and regression tasks. In case of classification, the output labels are discrete whereas they are continuous in case of regression.

5.2. Unsupervised Learning

The unsupervised learning approach is all about recognizing unidentified existing patterns from the data in order to derive rules from them. This technique is appropriate in a situation when the categories of data are unknown. Here, the training data is not labeled. Unsupervised learning is regarded as a statistic based approach for learning and thus refers to the problem of finding hidden structure in unlabeled data. Figure 5 explicates the concept.

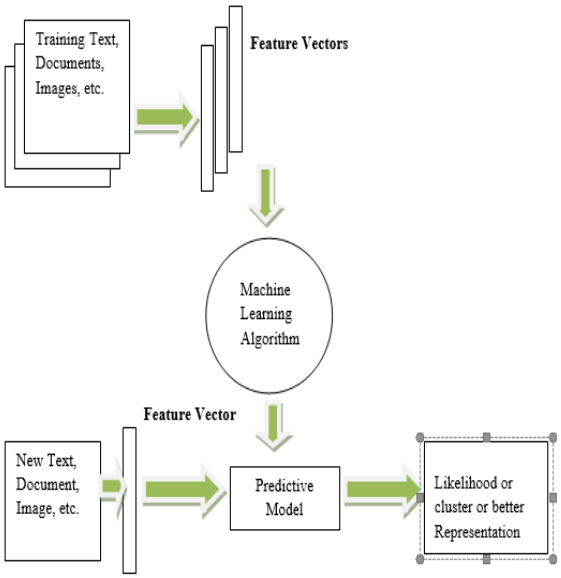


Fig 5. Unsupervised Learning

5.3. Reinforcement Learning

Reinforcement learning is regarded as an intermediate type of learning as the algorithm is only provided with a response that tells whether the output is correct or not. The algorithm has to explore and rule out various possibilities to get the correct output. It is regarded as learning with a Critic as the algorithm doesn't propose any sort of suggestions or solutions to the problem.

5.4. Evolutionary Learning

It is inspired by biological organisms who adapt to their environment. The algorithm understands the behavior and adapts to the inputs and rules out unlikely solutions. It is based on the idea of fitness to propose the best solution to the problem.

5.5. Semi-Supervised Learning

These algorithms provide a technique that harnesses the power of both - supervised learning and unsupervised learning. In the previous two types output labels areeither provided for all the observations or no labels are provided. There might be situations when some observations are provided with labels but majority of observations are unlabeled due to high cost of labeling and lack of skilled human expertise. In such situations, semi-supervised algorithms are best suited for model building. Semi supervised learning can be used with problems like classification, regression and prediction [4, 13,18].

Discussion Publishers edition DPR-2026-1150 https://discussionpublishers.com/research/606097fc-1d41-4e24-b65e-9915932ce021

It may further be categorized as Generative Models, Self-Training and Transductive SVM.

5.6. Ensemble Learning

It is a machine learning model in which numerous learners (individual models) are trained to solve a common problem. Unlike other machine learning techniques which learn a single hypothesis from the training data, ensemble learning tries to learn by constructing a set of hypotheses from the training data and by combining them to make a prediction model [15,19] in order to decrease bias(boosting), variance (bagging), or improve predictions(stacking). Ensemble learning can be further divided into two groups:

- Sequential ensemble approaches- These are the methods in which the base learners are constructed sequentially (AdaBoost). This method exploits the dependence between the base learners.
- Parallel ensemble approaches- In these, the base learners are independent of each other, so this relationship is exploited by constructing the base learners in parallel (e.g. Random Forest)
    - *Bagging:* It stands for bootstrap aggregation. It implements homogenous learners on sample populations and takes the mean of all predictions [4, 15,16]. For example, M different trees can be trained on dissimilar subsets of data and compute the ensemble as:

$$f(x) = 1/M \sum_{m=1}^{M} f_m(x)$$

    - *Boosting:* It is an iterative technique that adjusts the observation's weight on the basis of last classification. It tries to fit a sequence of weak learner models that performs a little better than just random predicting e.g. small decision trees [4,16]. AdaBoost stands for adaptive boosting and is the most widely used boosting algorithm.

5.7. Artificial Neural Network

Artificial neural networks (ANNs) are encouraged by the biological neural network. A neural network is an interconnection of neuron cells that help the electric impulses to propagate through the brain. The basic unit of learning in a neural network is a neuron, which is a nerve cell. A neuron consists of four parts, namely dendrites (receptor), soma (processor of electric signal), nucleus (core of the neuron) and axon (the transmitting end of the neuron). Analogical to a biological neural network, an ANN works on three layers: input layer, hidden layer and output layer. This type of network has weighted interconnections and learns by adjusting the weights of interconnections in order to perform parallel distributed processing. The Perceptron learning algorithm, Back-propagation algorithm, Hopfield Networks, Radial Basis Function Network (RBFN) are some popular algorithms.

Based on learning behavior, ANN can be further classified as:

- *Supervised Neural Network* – The inputs and the outputs are presented to the network as training data. The network is trained with this data by adjusting the weights to get accurate results. When it is fully trained, it is presented with unseen data to predict the output.
- *Unsupervised Neural Network*- In unsupervised neural network, the network is not provided with any output. It tries to find some structure or correlation among the input data and group those data together in a group or class. When new data is presented to it as input, it identifies its features and classifies it in one of the groups based on similarities.
- *Reinforcement Neural Network*- As humans interact with their environment and learn from mistakes, a reinforcement neural network also learns from its past decisions by way of penalties for wrong decisions and rewards for good decisions. The connection weights producing correct output are strengthened, while those producing incorrect responses are weakened.

5.8. Instance based learning

Discussion Publishers edition DPR-2026-1150 https://discussionpublishers.com/research/606097fc-1d41-4e24-b65e-9915932ce021

Unlike other machine learning methods where clear definition of the target function are provided from the training data, this lerning method does not describe any target function in the beginning. Rather it simply stores the training instance and generalizing is postponed until a new instance is classified. Hence it is also known as lazy learner. Such methods build up a database of training instances and whenever new data is presented as input it compares that data with other instances in the database using a similarity measure to find the nearest match and make the prediction [4]. The lazy learner estimates the target function differently and locally for every new instance to be classified instead of estimating it globally for the whole instance space hence it is faster to train but, takes time in making prediction [16]. K-Means, k-medians, hierarchical clustering and expectation maximization are some popular instance-based algorithms.

5.9. Dimensionality reduction algorithms

During the past few decades, intelligent machine learning models have been adopted in numerous complex and data intensive applications like climatology, biology, astronomy, medical, economy and finance. However, existing ML based systems are not sufficiently efficient and extensible enough to deal with massive and voluminous data. High dimensionality of data has proved to be a curse for data processing. Another challenge is sparsity of data. Global optimum is costly to find for such data. A dimensionality reduction algorithm helps in reducing the computational cost by reducing the number of dimensions of the data. It does so by reducing the redundant and irrelevant data and cleaning the data so as to improve the accuracy of results. Dimensionality reduction works in an unsupervised manner to search and exploit the implicit structure in the data [4, 5]. There are many algorithms for dimensionality reduction that can be adapted with classification and regression algorithms like Multidimensional scaling (MDS), Principal component analysis (PCA), Linear Discriminant Analysis (LDA), Principal component regression (PCR), and Linear Discriminant Analysis (LDA).

5.10. Hybrid Learning

Though ensemble learning appeared as a relief to researchers dealing with the common problems of computational complexity, over fitting and sticking to local minima in classification algorithms, researchers have found problems with ensemble learning. Complicated ensemble of multiple classifiers makes it difficult to implement and difficult to analyze the results. Instead of improving accuracy of the model, ensembles may tend to increase error at the level of individual base learner. Ensembles may result in poor accuracy as a result of selection of poor classifiers in combination. Recent approach to deal with such problems is hybridization i.e. creating ensemble of heterogeneous models. In this, more than one method is combined for example, combining clustering and decision tree or clustering and association mining etc.

Out of all the above-mentioned learning paradigms, the supervised learning is by far the most popular with researchers and practitioners.

6. Machine Learning Algorithms

In this section, we focus on some popular machine learning algorithms from the different paradigms [4, 5, 6, 16] explained in the preceding section. Although the number of algorithms falling within each paradigm are numerous and reported across pertinent literature, in this study we consider only few of these. The following table 2 briefly explains few of these algorithms.

These algorithms have a wide domain of practical applications, some of which are described in the next section.

Table 2. ML Algorithms

Discussion Publishers edition DPR-2026-1150 https://discussionpublishers.com/research/606097fc-1d41-4e24-b65e-9915932ce021

Paradigm	Algorithm	Description
Supervised Learning	Decision Tree	Decision Tree is a technique for approximating discrete valued target function which represents the learnt function in the form of a decision tree [10]. A decision tree classifies instances by sorting them from root to some leaf nodes on the basis of feature values. Each node represents some decision (test condition) on attribute of the instance whereas every branch represents a possible value for that feature. Classification of an instance starts at the root node called the decision node. Based on the value of node, the tree traverse down along the edge which corresponds to the value of the output of feature test. This process continues in the sub-tree headed by the new node at the end of the previous edge. Finally, the leaf node signifies the classification categories or the final decision. While using a decision tree, focus is on how to decide which attribute is the best classifier at each node level. Statistical measure like information gain, Gini index, Chi-square and entropy are calculated for each node to calculate the worth of that node [10]. Several algorithms are used to implement decision trees. The most popular ones are: Classification and Regression Tree (CART), Iterative Dichotomiser 3 (ID3), Automatic Interaction Detection (CHAID), Chi-Squared C4.5 and C5.0 and M5
	Naïve Bayes	Naïve Bayes classifies using Bayes' Theorem of probability. Bayes' theorem calculates the posterior probability of an event (A) given some prior probability of event B represented by P(A/B) as follows: $$\mathrm{P(A/B)} = \frac{P(B/A)P(A)}{P(B)}$$ Where, • A and B are events. • P(A) and P(B) are the probabilities of observing A and B independent of each other. • P(A/B) is the conditional probability, i.e. Probability of observing A, given B is true. • P(B/A) is the probability of observing B, given A is True. Naïve Bayes' classifiers fall under the category of simple probabilistic classifiers based on the concept of Bayes' Theorem having strong independence assumptions among the features.

Discussion Publishers edition DPR-2026-1150 https://discussionpublishers.com/research/606097fc-1d41-4e24-b65e-9915932ce021

		It is particularly suited when the dimensionality of the inputs is high [4,5].
	Support Vector Machines	SVMs can be used for classification as well as regression problems. It is a supervised learning algorithm. It works on the concept of margin calculation. In this algorithm, each data item is plotted as a point in n-dimensional space (where n is the number of features we have in our dataset). The value of each feature is the value of the corresponding coordinate. It classifies the data into different classes by finding a line (hyper plane) which separates the training datasets into classes. It works by maximizing the distances between the nearest data point (in both classes) and the hyper plane that we can call as margin.
	Regression Analysis	Regression analysis is a predictive modelling technique which investigates the relationship between a dependent (target) and independent variable(s) (predictor). It is an important tool for analysing and modelling of data. In this method, we try to fit the line/curve to the data points so as to minimize the differences between distances of data points from the curve or line. There are various kinds of regression analysis like linear, logistic and polynomial.
Unsupervised Learning	K-Means Clustering	K-means is a popular unsupervised machine learning algorithm for cluster analysis. Its goal is to partition ‘n’ observations into ‘k’ clusters in which each observation belongs to the cluster having the nearest mean, serving as a prototype of the cluster. The mean of the observations in a particular cluster defines the center of the cluster.
Instance based Learning	K-nearest Neighbours	It is a non-parametric method used for classification and regression. Given N training vectors, KNN algorithm identifies the k-nearest neighbours of an unknown feature vector whose class is to be identified.
Ensemble Learning	Random Forest	It is an ensemble learning method used in classification and regression. It uses bagging approach to create a bunch of decision trees with random subset of data. The output of all decision trees in the random forest is combined to make the final decision trees. There are two stages in Random Forest Algorithm, one is to create random forest, and the other is to make a prediction from the random forest classifier created in the first stage.

Discussion Publishers edition DPR-2026-1150 https://discussionpublishers.com/research/606097fc-1d41-4e24-b65e-9915932ce021

Dimensionality Reduction	Principal Component Algorithm	It is primarily used for reducing dimensionality of data set. It helps in reducing the number of features of the data set or the number of independent variables in the data set. It uses orthogonal transformation to convert correlated variables into a set of linearly uncorrelated variables called principal components.

7. Applications of Machine Learning

Machine learning problems range from game playing to self-driven vehicles. Table 3 enlists some popular real-life applications of ML.

Table 3. Applications of ML

Application	Description
Playing Checkers Game	A computer program learns to play checkers game, improvises its performance as determined by its ability to win at various class of tasks involving the game, through experience obtained by playing games against itself.
Speech Recognition	The most sophisticated speech recognition systems these days deploy machine learning algorithms in some forms. Example: SPHINX system [20] learns speaker-specific sounds and words from speech signals. Various Neural Network learning methodologies for interpreting hidden Markov Models are highly effective for automatically customizing speakers, dictionary, noise etc.
Autonomous Vehicles	Machine learning models are these days being applied to drive autonomous vehicles like Cars, Drones etc. Example: Google Driver Less Cars, Tesla Cars. Machine learning techniques are also highly effective in controlling sensor-based applications.
Filtering Emails (Spam Emails)	Machine learning can be applied to filter spam emails. The machine learning based model will simply memorize all the emails classified as spam emails by user. When new email arrives in inbox, the machine learning based model will search, compare and based on the previous spam emails. If new email matches any one of them, it will be marked as spam; else it will be moved to user's inbox.
Robotics and Artificial Intelligence	Machine learning is regarded as improved approach to problem solving. Using base knowledge and training data with machine learning models, learning can be improved which will take robotics and AI to next generation levels.
Web and Social Media	• Naïve Bayes classifiers have been successfully applied in the field of text mining, may it be spam filtering or classifying the web page, an email or any document. • Facebook uses Naïve Bayes' to analyze status update expressing positive and negative emotions. • Document Categorization: Google uses Naïve Bayes algorithm for document categorization. • K-means clustering is used by search engines like Google, Yahoo to cluster web pages by similarity. • Apriori is used by websites such as Amazon or Flipkart to recommend which items are purchased together frequently. • Another common application of Apriori is the Google auto-complete. When a person types a word, Google search engine looks for other associated words that go together with the word earlier typed word. • Sentiment analysis on social networking sites is a typical text classification problem solved using application of variety of ML algorithms [21, 22, 23,24]
Medical Field	TRISS: Trauma & Injury Severity Score, which is widely used to predict mortality in injured patients, was originally developed by Boyd et al. using logistic regression. Many other medical scales used to assess severity of a patient have been developed using logistic regression.
Bayesian Methods	Bayes' Theorem is one of the most popular methods for calculating conditional probabilities given the set of hypothesis. It can be used to solve complex data science and analytics problems by integrating with various Machine Learning models and algorithms. Some examples real world problems that can be solved using Bayesian methods are: • Suggestions on Netflix • Auto-correction

Discussion Publishers edition DPR-2026-1150 https://discussionpublishers.com/research/606097fc-1d41-4e24-b65e-9915932ce021

	• Credit card fraud detection • Google page ranking • Facial recognition • Climate modeling forecast • Stock trading systems

8. Analysis and Discussion

Given the wide range of applicability of ML, it faces a number of challenges. Some of them are as follows:

- The machine learning algorithms require large volumes of data to be accurate and efficient which is still not available to researchers. Tech giants like Facebook and Google have had access to such enormous data which is why they are leading in the field of Artificial Intelligence. It becomes even more difficult to get this data in the fields like banking and healthcare where sparse digital data is available making it tough to make accurate predictions.
- Spam Detection: Given email in an inbox, the intelligent systems developed so far are still not able to correctly detect the spam mail. It ends up in sending spam in inbox and non-spam mails to spam directory.
- Machine learning algorithms have not yet been successful in identifying the objects and images. This field is still an open research field for machine learning. Though we have mentioned a few challenges in machine learning, there are many more fields which still challenging for deep learning algorithms i.e. speech understanding, credit card fraud detection, face detection, digit recognition given a zip code, and product recommendation etc.

Machine learning algorithms are in continuous development and will definitely become more widespread in the years to come. They are useful in many different applications; there is significant brainpower and funding behind pushing the boundaries towards innovation [25]. Some open areas of application include:

- Deep learning, e.g. for predicting stock market trends, designing circuits, identifying illnesses, designing voice-controlled devices, and much more (save special attention for Generative Adversarial Neural Networks)
- Data mining and big data analytics, e.g. for predicting business market trends
- Natural language processing, e.g. in search engines
- Hardware accelerators for new AI architectures, e.g. from AMD and Intel
- Simulation environments for evaluation and testing, e.g. for self-driving cars and virtual reality
- Healthcare machine learning (medical imaging, working with clinical data, making sense of genomic data of huge populations)
- HCI (Human Computer Interaction), keep advancing better interfaces and usability between different devices with the rise of cloud computing and IoT

Given the current rate of advancement in the field, it has a bright future. Applications of machine learning have the potential to expand dramatically in the near future.

9. Machine Learning in Future

In near future machine learning is expected to be a part of almost every software application. There are some of the future predictions of machine learning applications: As machine learning helps computers understand the context and semantics of sentences using Natural Language Processing, so we do not have to wait long for a time when computers will learn to talk like humans. In the near future we can expect machine learning tools and techniques to connect to the internet and continuously retain on the most relevant information. This will help algorithms in constant retaining of algorithms and there will be no need to train the systems time and again. Personalization could be enhanced and recommendations could be improved leading to more beneficial and successful experiences.

Discussion Publishers edition DPR-2026-1150 https://discussionpublishers.com/research/606097fc-1d41-4e24-b65e-9915932ce021

10. Conclusion

Digitalization and the Internet revolution have led to a mounting volume of structured and unstructured data which needs to be utilized for analytics. Machine learning as a key technology driver encompasses the intelligent power to harness the knowledge from the available data. Moreover, the adoption of machine learning solutions for complex real-life problems by both researchers & practitioners has made this field a dynamic area of research with an active participation across industries & countries. In this paper, a comprehensive review of Machine Learning process and algorithms is presented. The purpose is clearly to understand the role, advantage and scope of Machine learning as a technology-based solution.

References

[1] Samuel AL. Some studies in machine learning using the game of checkers. IBM Journal of research and development. 1959 Jul;3(3):210-29.
[2] Mitchell, T. (1997). Machine Learning. McGraw Hill. p. 2. ISBN 978-0-07-042807-2.
[3] Hayashi, Chikio (1998-01-01). "What is Data Science? Fundamental Concepts and a Heuristic Example". Data Science, Classification, and Related Methods. Studies in Classification, Data Analysis, and Knowledge Organization. Springer Japan. pp. 40–51. doi:10.1007/978-4-431-65950-1_3. ISBN 9784431702085
[4] Sandhya N. dhage, Charanjeet Kaur Raina, "A review on Machine Learning Techniques", March 16 Volume 4 Issue 3 , International Journal on Recent and Innovation Trends in Computing and Communication (IJRITCC), ISSN: 2321-8169, PP: 395 – 399
[5] AyonDey , "Machine Learning Algorithms: A Review", (IJCSIT) International Journal of Computer Science
[6] A report by Royal Society, April 2017, "Machine learning: the power and promise of computers that learn by example ", ISBN: 978-1-78252-259-1.and Information Technologies, Vol. 7 (3) , 2016, 1174-1179
[7] Minton S, Zweben M. Learning, Planning, and Scheduling: An Overview. InMachine Learning Methods for Planning 1993 (pp. 1-29).
[8] Sejnowski T. Net talk: A parallel network that learns to read aloud. Complex Systems. 1987;1:145-68.
[9] Han J, Cai Y, Cercone N. Data-driven discovery of quantitative rules in relational databases. IEEE Transactions on Knowledge & Data Engineering. 1993 Feb 1(1):29-40.
[10] Shokri R, Stronati M, Song C, Shmatikov V. Membership inference attacks against machine learning models. InSecurity and Privacy (SP), 2017 IEEE Symposium on 2017 May 22 (pp. 3-18). IEEE.
[11] Li M, Andersen DG, Park JW, Smola AJ, Ahmed A, Josifovski V, Long J, Shekita EJ, Su BY. Scaling Distributed Machine Learning with the Parameter Server. InOSDI 2014 Oct 6 (Vol. 14, pp. 583-598).
[12] Chen JX. The evolution of computing: AlphaGo. Computing in Science & Engineering. 2016 Jul;18(4):4-7.
[13] Al-Jarrah OY, Yoo PD, Muhaidat S, Karagiannidis GK, Taha K. Efficient machine learning for big data: A review. Big Data Res. 2015; 2(3):87–93.
[14] A.Khan,B.Baharudin,Lan Hong Lee, " A Review of Machine Learning Algorithms for Text Documents Classification", Journal Of Advances in Information Technology, Vol. 1 , No. 1, Feb.2010.
[15] Zhou ZH. (2009) Ensemble Learning. In: Li S.Z., Jain A. (eds) Encyclopedia of Biometrics. Springer, Boston, MA
[16] Tom M. Mitchell , "Machine Learning ", WCB McGraw-Hill, 2013, ISBN-13:978-1-25-909695-2, ISBN-10: 1-25-909695-5.

Discussion Publishers edition DPR-2026-1150 https://discussionpublishers.com/research/606097fc-1d41-4e24-b65e-9915932ce021

[17] Tang W, Zhang T, Sazonov E. The application of machine learning in monitoring physical activity with shoe sensors. Cyber-Physical Systems: Integrated Computing and Engineering Design. 2013 Sep 26:187.
[18] Zhu X. Semi-supervised learning literature survey. Computer Science, University of Wisconsin-Madison. 2006;2(3):4.
[19] Witten IH, Frank E, Hall MA, Pal CJ. Data Mining: Practical machine learning tools and techniques. Morgan Kaufmann; 2016 Oct 1.
[20] Lee KF, Hon HW, Hwang MY, Mahajan S, Reddy R "The SPHINX speech recognition system", International Conference on Acoustics, Speech, and Signal Processing (ICASSP '89), Vol. 1, pp. 445 - 448, May, 1989
[21] KumarA., Sharma A., "Systematic Literature Review on Opinion Mining of Big Data for Government Intelligence." Webology, Vol.14, No. 2, December 2017.
[22] Kumar A., Jaiswal A., "Empirical Study of Twitter and Tumblr for Sentiment Analysis using Soft Computing Techniques" International Conference on Soft Computing and Applications (ICSCA 2017), World Congress on Engineering and Computer Science 2017, vol. 1, ISBN: 978-988-14047-5-6, pp. 1-5, Oct 2017.
[23] Kumar A., Abraham A. "Opinion Mining to Assist User Acceptance Testing for Open-Beta Versions", Journal of Information Assurance and Security, Vol.12, Issue 4, pp. 146-153, 2017, ISSN:1554-1010.
[24] Pazzani MJ, Billsus D. Content-based recommendation systems. InThe adaptive web 2007 (pp. 325-341). Springer, Berlin, Heidelberg.
[25] Bhatia MPS, Kumar A., "Information Retrieval & Machine Learning: Supporting Technologies for Web Mining Research & Practice", Webology, Vol. 5, No. 2.

Discussion Publishers edition DPR-2026-1150 https://discussionpublishers.com/research/606097fc-1d41-4e24-b65e-9915932ce021