

# How to Read Research Papers

## → When learning a new topic!

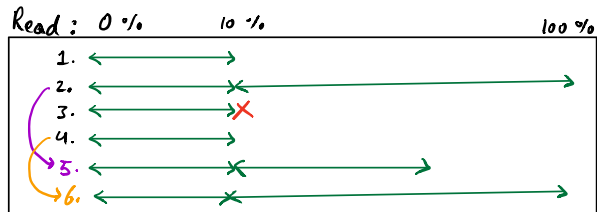
→ Compile a list of papers (+ medium / blog posts)

→ Skip around the list:

↳ Read 10% of all papers

↳ Prune the ones that are not interesting or covered are in other papers

↳ Collect more papers which interest you through references



To have a basic understanding of a topic: Read 5-20 papers

To have a very good understanding of a topic: Read 50-100 papers

## → When Reading 1 paper!

→ Take multiple passes of the paper

1. Read Title / Abstract / Figures

2. Read Intro + Conclusion + Figures + Skim rest (Skip or skim related work)

3. Read out but skip / skim the math.

4. Read whole thing but skip parts that don't make sense.

## → After reading a paper try to answer:

→ What did authors try to accomplish?

→ What were the key elements of the approach?

→ What can you use yourself from this paper?

→ What other references to follow?



To truly understand the math behind a paper, try to derive it from scratch after reading through the paper. (This will also help you in deriving your own proofs).

# A Brief Introduction into Machine Learning

**Gunnar Rätsch**

Friedrich Miescher Laboratory of the Max Planck Society,  
Spemannstraße 37, 72076 Tübingen, Germany

<http://www.tuebingen.mpg.de/~raetsch>

## 1 Introduction

The Machine Learning field evolved from the broad field of *Artificial Intelligence*, which aims to mimic intelligent abilities of humans by machines. In the field of *Machine Learning* one considers the important question of how to make machines able to “learn”. Learning in this context is understood as *inductive inference*, where one observes *examples* that represent incomplete information about some “statistical phenomenon”. In *unsupervised* learning one typically tries to uncover hidden regularities (e.g. clusters) or to detect anomalies in the data (for instance some unusual machine function or a network intrusion). In *supervised learning*, there is a *label* associated with each example. It is supposed to be the answer to a question about the example. If the label is discrete, then the task is called *classification problem* – otherwise, for real-valued labels we speak of a *regression problem*. Based on these examples (including the labels), one is particularly interested to *predict* the answer for other cases before they are explicitly observed. Hence, learning is not only a question of remembering but also of *generalization to unseen cases*.

## 2 Supervised Classification

An important task in Machine Learning is *classification*, also referred to as pattern recognition, where one attempts to build algorithms capable of automatically constructing methods for distinguishing between different exemplars, based on their differentiating patterns.

Watanabe [1985] described a pattern as “the opposite of chaos; it is an entity, vaguely defined, that could be given a name.” Examples of patterns are human faces, text documents, handwritten letters or digits, EEG signals, and the DNA sequences that may cause a certain disease. More formally, the goal of a (supervised) classification task is to find a functional mapping between the input data  $X$ , describing the input pattern, to a class label  $Y$  (e.g.  $-1$  or  $+1$ ), such that  $Y = f(X)$ . The construction of the mapping is based on so-called *training data* supplied to the classification algorithm. The aim is to accurately predict the correct label on unseen data.

A pattern (also: “example”) is described by its *features*. These are the characteristics of the examples for a given problem. For instance, in a face recognition task some features could be the color of the eyes or the distance between the eyes. Thus, the input

to a pattern recognition task can be viewed as a two-dimensional matrix, whose axes are the examples and the features.

Pattern classification tasks are often divided into several sub-tasks:

1. Data collection and representation.
2. Feature selection and/or feature reduction.
3. Classification.

Data collection and representation are mostly problem-specific. Therefore it is difficult to give general statements about this step of the process. In broad terms, one should try to find invariant features, that describe the differences in classes as best as possible.

Feature selection and feature reduction attempt to reduce the dimensionality (i.e. the number of features) for the remaining steps of the task. Finally, the classification phase of the process finds the actual mapping between patterns and labels (or targets). In many applications the second step is not essential or is implicitly performed in the third step.

### 3 Classification Algorithms

Although Machine Learning is a relatively young field of research, there exist more learning algorithms than I can mention in this introduction. I chose to describe six methods that I am frequently using when solving data analysis tasks (usually classification). The first four methods are traditional techniques that have been widely used in the past and work reasonably well when analyzing low dimensional data sets with not too few labeled training examples. In the second part I will briefly outline two methods (Support Vector Machines & Boosting) that have received a lot of attention in the Machine Learning community recently. They are able to solve high-dimensional problems with very few examples (e.g. fifty) quite accurately and also work efficiently when examples are abundant (for instance several hundred thousands of examples).

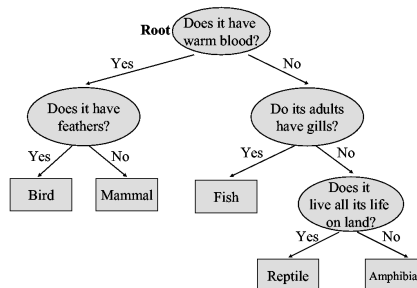
#### 3.1 Traditional Techniques

**k-Nearest Neighbor Classification** Arguably the simplest method is the *k-Nearest Neighbor* classifier [Cover and Hart, 1967]. Here the  $k$  points of the training data closest to the test point are found, and a label is given to the test point by a majority vote between the  $k$  points. This method is highly intuitive and attains – given its simplicity – remarkably low classification errors, but it is computationally expensive and requires a large memory to store the training data.

**Linear Discriminant Analysis** computes a hyperplane in the input space that minimizes the within-class variance and maximizes the between class distance [Fisher 1936]. It can be efficiently computed in the linear case even with large data sets. However, often a linear separation is not sufficient. Nonlinear extensions by using kernels exist [Mika et al., 2003], however, making it difficult to apply it to problems with large training sets.

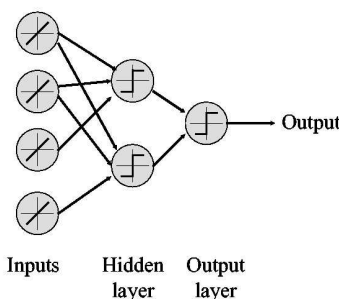
**Decision Trees** Another intuitive class of classification algorithms are *decision trees*. These algorithms solve the classification problem by repeatedly partitioning the in-

put space, so as to build a tree whose nodes are as pure as possible (that is, they contain points of a single class). Classification of a new test point is achieved by moving from top to bottom along the branches of the tree, starting from the root node, until a terminal node is reached. Decision trees are simple yet effective classification schemes for small datasets. The computational complexity scales unfavorably with the number of dimensions of the data. Large datasets tend to result in complicated trees, which in turn require a large memory for storage. The C4.5 implementation by [Quinlan \[1992\]](#) is frequently used and can be downloaded at <http://www.rulequest.com/Personal>.



**Figure 1:** An example a decision tree (Figure taken from [Yom-Tov \[2004\]](#)).

**Neural Networks** are perhaps one of the most commonly used approaches to classification. Neural networks (suggested first by [Tur-ing \[1992\]](#)) are a computational model inspired by the connectivity of neurons in animate nervous systems. A further boost to their popularity came with the proof that they can approximate any function mapping via the Universal Approximation Theorem [\[Haykin, 1999\]](#). A simple scheme for a neural network is shown in Figure 2. Each circle denotes a computational element referred to as a *neuron*, which computes a weighted sum of its inputs, and possibly performs a nonlinear function on this sum. If certain classes of nonlinear functions are used, the function computed by the network can approximate any function (specifically a mapping from the training patterns to the training targets), provided enough neurons exist in the network and enough training examples are provided.



**Figure 2:** A schematic diagram of a neural network. Each circle in the hidden and output layer is a computational element known as a neuron. (Figure taken from [Yom-Tov \[2004\]](#))

### 3.2 Large Margin Algorithms

Machine learning rests upon the theoretical foundation of *Statistical Learning Theory* [e.g. [Vapnik, 1995](#)] which provides conditions and guarantees for good generalization of learning algorithms. Within the last decade, *large margin classification techniques* have emerged as a practical result of the theory of generalization. Roughly speaking, the margin is the distance of the example to the separation boundary and a large margin classifier generates decision boundaries with large margins to almost all training examples. The two most widely studied classes of large margin classifiers are *Support Vector Machines (SVMs)* [\[Boser et al., 1992, Cortes and Vapnik, 1995\]](#) and *Boosting* [\[Valiant, 1984, Schapire, 1992\]](#):

**Support Vector Machines** work by mapping the training data into a feature space by the aid of a so-called kernel function and then separating the data using a *large margin hyperplane* (cf. Algorithm 1). Intuitively, the kernel computes a similarity

between two given examples. Most commonly used kernel functions are *RBF kernels*  $k(\mathbf{x}, \mathbf{x}') = \exp\left(-\frac{\|\mathbf{x}-\mathbf{x}'\|^2}{\sigma^2}\right)$  and *polynomial kernels*  $k(\mathbf{x}, \mathbf{x}') = (\mathbf{x} \cdot \mathbf{x}')^d$ .

The SVM finds a large margin separation between the training examples and previously unseen examples will often be close to the training examples. Hence, the large margin then ensures that these examples are correctly classified as well, i.e., the decision rule *generalizes*. For so-called *positive definite* kernels, the optimization problem can be solved efficiently and SVMs have an interpretation as a hyperplane separation in a high dimensional feature space [Vapnik, 1995, Müller et al., 2001, Schölkopf and Smola, 2002]. Support Vector Machines have been used on million dimensional data sets and in other cases with more than a million examples [Mangasarian and Musicant, 2001]. Research papers and implementations can be downloaded from the kernel machines web-site <http://www.kernel-machines.org>.

---

**Algorithm 1** The Support Vector Machine with regularization parameter  $C$ .

---

Given labeled sequences  $(\mathbf{x}_1, y_1), \dots, (\mathbf{x}_m, y_m)$  ( $\mathbf{x} \in X$  and  $y \in \{-1, +1\}$ ) and a kernel  $k$ , the SVM computes a function

$$f(s) = \sum_{i=1}^m \alpha_i k(\mathbf{x}_i, \mathbf{x}) + b,$$

where the coefficients  $\alpha_i$  and  $b$  are found by solving the optimization problem

$$\begin{aligned} & \text{minimize} && \sum_{i,j=1}^m \alpha_i \alpha_j k(s_i, s_j) + C \sum_{i=1}^m \xi_i \\ & \text{subject to} && y_i f(\mathbf{x}_i) \geq 1 - \xi_i \end{aligned}$$

---

**Boosting** The basic idea of boosting and *ensemble learning algorithms* in general is to iteratively combine relatively simple *base hypotheses* – sometimes called *rules of thumb* – for the final prediction. One uses a so-called *base learner* that generates the base hypotheses. In boosting the base hypotheses are linearly combined. In the case of *two-class classification*, the final prediction is the weighted majority of the votes. The combination of these simple rules can boost the performance drastically. It has been shown that Boosting has strong ties to support vector machines and large margin classification [Rätsch, 2001, Meir and Rätsch, 2003]. Boosting techniques have been used on very high dimensional data sets and can quite easily deal with than hundred thousands of examples. Research papers and implementations can be downloaded from <http://www.boosting.org>.

## 4 Summary

Machine Learning research has been extremely active the last few years. The result is a large number of very accurate and efficient algorithms that are quite easy to use for a practitioner. It seems rewarding and almost mandatory for (computer) scientist and engineers to learn how and where Machine Learning can help to automate tasks

or provide predictions where humans have difficulties to comprehend large amounts of data. The long list of examples where Machine Learning techniques were successfully applied includes: Text classification and categorization [e.g. [Joachims, 2001](#)] (for instance spam filtering), network intrusion detection [e.g. [Laskov et al., 2004](#)], Bioinformatics (e.g. cancer tissue classification, gene finding; e.g. [Furey et al., 2000](#)), [Zien et al., 2000](#)), [Sonnenburg et al., 2002](#)), brain computer interfacing [e.g. [Blankertz et al., 2003](#)], monitoring of electric appliances [e.g. [Onoda et al., 2000](#)], optimization of hard disk caching strategies [e.g. [Gramacy et al., 2003](#)] and disk spin-down prediction [e.g. [Helmbold et al., 2000](#)]), drug discovery [e.g. [Warmuth et al., 2003](#)]), high-energy physics particle classification, recognition of hand writing, natural scene analysis etc.

Obviously, in this brief summary I have to be far from being complete. I did not mention regression algorithms (e.g. ridge regression, regression trees), unsupervised learning algorithms (such as clustering, principle component analysis), reinforcement learning, online learning algorithms or model-selection issues. Some of these techniques extend the applicability of Machine Learning algorithms drastically and would each require an introduction for them self. I would like to refer the interested reader to two introductory books [\[Mendelson and Smola, 2003\]](#), [\[von Luxburg et al., 2004\]](#) which are the result of the last annual Summer Schools on Machine Learning (cf. <http://mlss.cc>).

**Acknowledgments** I would like to thank Julia Lüning for encouraging me to write this paper. Furthermore, thanks to Sören Sonnenburg and Nora Toussaint for proof-reading the manuscript. For the preparation of this work I used and modified some parts of [\[Müller et al., 2001\]](#), [\[Rätsch, 2001\]](#), [\[Meir and Rätsch, 2003\]](#), [\[Yom-Tov, 2004\]](#). This work was partly funded by the PASCAL Network of Excellence project (EU #506778).

## References

- B.. Blankertz, G. Dornhege, C. Schäfer, R. Krepki, J. Kohlmorgen, K.-R. Müller, V. Kunzmann, F. Losch, and G. Curio. BCI bit rates and error detection for fast-pace motor commands based on single-trial EEG analysis. *IEEE Transactions on Neural Systems and Rehabilitation Engineering*, 11:127–131, 2003.
- B.E. Boser, I.M. Guyon, and V.N. Vapnik. A training algorithm for optimal margin classifiers. In D. Haussler, editor, *Proceedings of the 5th Annual ACM Workshop on Computational Learning Theory*, pages 144–152, 1992.
- C. Cortes and V.N. Vapnik. Support vector networks. *Machine Learning*, 20:273–297, 1995.
- T.M. Cover and P.E. Hart. Nearest neighbor pattern classifications. *IEEE transaction on information theory*, 13(1):21–27, 1967.
- R.A. Fisher. The use of multiple measurements in taxonomic problems. *Annals of Eugenics*, 7:179–188, 1936.
- T. Furey, N. Cristianini, N. Duffy, D. Bednarski, M. Schummer, and D. Haussler. Support vector machine classification and validation of cancer tissue samples using microarray expression data. *Bioinformatics*, 10:906–914, 2000.
- R.B. Gramacy, M.K. Warmuth, S.A. Brandt, and I. Ari. Adaptive caching by refetching. In *In Advances in Neural Information Processing Systems 15*, pages 1465–1472. MIT Press, 2003.

- S. Haykin. *Neural Networks: A comprehensive foundation, 2nd Ed.* Prentice-Hall, 1999.
- D.P. Helmbold, D.D.E. Long, T.L. Sconyers, and B. Sherrod. Adaptive disk spin-down for mobile computers. *Mobile Networks and Applications*, 5(4):285–297, 2000.
- T. Joachims. *The Maximum-Margin Approach to Learning Text Classifiers*. PhD thesis, Computer Science Dept., University of Dortmund, 2001.
- P. Laskov, C. Schäfer, and I. Kotenko. Intrusion detection in unlabeled data with quarter-sphere support vector machines. In *Proc. DIMVA*, pages 71–82, 2004.
- O.L. Mangasarian and D.R. Musicant. Lagrangian support vector machines. *JMLR*, 1:161–177, March 2001.
- R. Meir and G. Rätsch. An introduction to boosting and leveraging. In S. Mendelson and A. Smola, editors, *Advanced Lectures on Machine Learning*, LNAI, pages 119–184. Springer, 2003.
- S. Mendelson and A. Smola, editors. *Advanced Lectures on Machine Learning*, volume 2600 of LNAI. Springer, 2003.
- S. Mika, G. Rätsch, J. Weston, B. Schölkopf, A.J. Smola, and K.-R. Müller. Constructing descriptive and discriminative nonlinear features: Rayleigh coefficients in kernel feature spaces. *IEEE Transactions on Pattern Analysis and Machine Intelligence*, 25(5):623–627, May 2003.
- K.-R. Müller, S. Mika, G. Rätsch, K. Tsuda, and B. Schölkopf. An introduction to kernel-based learning algorithms. *IEEE Transactions on Neural Networks*, 12(2):181–201, 2001.
- T. Onoda, G. Rätsch, and K.-R. Müller. A non-intrusive monitoring system for household electric appliances with inverters. In H. Bothe and R. Rojas, editors, *Proc. of NC'2000*, Berlin, 2000. ICSC Academic Press Canada/Switzerland.
- J.R. Quinlan. *C4.5: Programs for Machine Learning*. Morgan Kaufmann, 1992.
- G. Rätsch. *Robust Boosting via Convex Optimization*. PhD thesis, University of Potsdam, Neues Palais 10, 14469 Potsdam, Germany, October 2001.
- R.E. Schapire. *The Design and Analysis of Efficient Learning Algorithms*. PhD thesis, MIT Press, 1992.
- B. Schölkopf and A.J. Smola. *Learning with Kernels*. MIT Press, Cambridge, MA, 2002.
- S. Sonnenburg, G. Rätsch, A. Jagota, and K.-R. Müller. New methods for splice-site recognition. In JR. Dorronsoro, editor, *Proc. International conference on artificial Neural Networks – ICANN'02*, pages 329–336. LNCS 2415, Springer Berlin, 2002.
- A.M. Turing. Intelligent machinery. In D.C. Ince, editor, *Collected works of A.M. Turing: Mechanical Intelligence*, Amsterdam, The Netherlands, 1992. Elsevier Science Publishers.
- L.G. Valiant. A theory of the learnable. *Communications of the ACM*, 27(11):1134–1142, November 1984.
- V.N. Vapnik. *The nature of statistical learning theory*. Springer Verlag, New York, 1995.
- U. von Luxburg, O. Bousquet, and G. Rätsch, editors. *Advanced Lectures on Machine Learning*, volume 3176 of LNAI. Springer, 2004.
- M. K Warmuth, J. Liao, G. Rätsch, Mathieson. M., S. Putta, and C. Lemmem. Support Vector Machines for active learning in the drug discovery process. *Journal of Chemical Information Sciences*, 43(2):667–673, 2003.
- W. Watanabe. *Pattern recognition: Human and mechanical*. Wiley, 1985.
- E. Yom-Tov. An introduction to pattern classification. In U. von Luxburg, O. Bousquet, and G. Rätsch, editors, *Advanced Lectures on Machine Learning*, volume 3176 of LNAI, pages 1–23. Springer, 2004.
- A. Zien, G. Rätsch, S. Mika, B. Schölkopf, T. Lengauer, and K.-R. Müller. Engineering Support Vector Machine Kernels That Recognize Translation Initiation Sites. *Bioinformatics*, 16(9):799–807, September 2000.

# Machine Learning: An Overview

Matthew N. O. Sadiku<sup>1</sup>, Samuel A. Ajayi<sup>2</sup>, Janet O. Sadiku<sup>3</sup>

<sup>1</sup>Roy G. Perry College of Engineering, Prairie View A&M University, Prairie View, TX, USA

<sup>2</sup>Texas Southern University, Houston, TX, USA

<sup>3</sup>Juliana King University, Houston, TX, USA

## ABSTRACT

Machine learning (ML) is a branch of artificial intelligence that performs data analysis tasks without explicit instructions. It is about learning some properties of a data set and applying them to new data. It refers to the automated detection of meaningful patterns in a given data. It involves using the past experience to optimize the performance of a given algorithm. ML algorithms open up a realm where computers analyze and learn from data to make predictions and decisions. As machine learning continues to evolve, its uses are set to revolutionize how we engage with technology, becoming an essential part of everyday life. This paper provides an overview of machine learning.

**KEYWORDS:** machine learning, artificial intelligence, prediction, deep learning.

**How to cite this paper:** Matthew N. O. Sadiku | Samuel A. Ajayi | Janet O. Sadiku "Machine Learning: An Overview" Published in International Journal of Trend in Scientific Research and Development (ijtsrd), ISSN: 2456-6470, Volume-9 | Issue-5, October 2025, pp.1121-1131, URL: [www.ijtsrd.com/papers/ijtsrd97693.pdf](http://www.ijtsrd.com/papers/ijtsrd97693.pdf)



Copyright © 2025 by author (s) and International Journal of Trend in Scientific Research and Development Journal. This is an Open Access article distributed under the terms of the Creative Commons Attribution License (CC BY 4.0) (<http://creativecommons.org/licenses/by/4.0>)



## INTRODUCTION

Technology is becoming more and more intelligent as scientists, engineers, and programmers teach computers to learn independently. Robots and machines have been occupying mankind for several centuries. Machine learning is the concept that a computer program can learn and adapt to new data without human intervention. It is focused on teaching computers to learn from data and to improve with experience – instead of being explicitly programmed to do so. It is used in different sectors for various reasons. It can be applied in a variety of areas, such as in investing, advertising, marketing, ecommerce, banking, organizing news, fraud detection, and more [1].

Machine learning (ML) is the discipline that gives computers the ability to learn without being explicitly programmed. It is the science of facilitating computers to learn from past experience. It is a field within computer science and is closely related to artificial intelligence (AI) and computational statistics. It may also be regarded as a branch of artificial intelligence (AI) that trains algorithms on data to build models capable of doing things usually

reserved for humans, like sorting images, analyzing patterns, or predicting trends. Artificial intelligence is an umbrella term for different strategies and techniques used to make machines more human-like. ML is based on the idea that machines should be able to learn and adapt through experience.

## WHAT IS MACHINE LEARNING?

Machine learning is a subfield of artificial intelligence that uses algorithms trained on data sets to create models capable of performing tasks that would otherwise only be possible for humans, such as categorizing images, analyzing data, or predicting price fluctuations. It uses algorithms (essentially lists of rules) trained on data sets to create self-learning models capable of predicting outcomes and classifying information without human intervention. It focuses on algorithms that can “learn” the patterns of training data and, subsequently, make accurate inferences about new data. This pattern recognition ability enables machine learning models to make decisions or predictions without explicit, hard-coded instructions. To ensure such algorithms work effectively, however, they must typically be refined

many times until they accumulate a comprehensive list of instructions that allow them to function correctly [2]. A symbol of machine learning is shown in Figure 1 [3].

Generally speaking, a learning problem considers a set of samples of data and then tries to predict properties of unknown data. ML builds heavily on statistics because when we train a machine to learn, we have to give it a statistically significant random sample as training data. Intelligent machines are increasing doing incredible things: Facebook recognizes faces in photos, Siri understands voices, and Google translates websites [4].

Machine learning techniques are transforming many fields including computer science, engineering, mathematics, physics, neuroscience, and cognitive science. We are surrounded by ML-based technologies: search engines learn how to bring us the best results, digital cameras learn to detect faces, credit card transactions are secured by a software that detects frauds, and cars are equipped with accident prevention systems that are built using ML algorithms [5]. In ML, data plays an indispensable role, and the learning algorithm is used to learn from the data. ML algorithms are now easy to use. One can download packages in Python. Programming languages used in ML include C++, Java, Python.

As its name indicates, machine learning works by creating computer-based statistical models that are refined for a given purpose by evaluating training data, rather than by the classical approach where programmers develop a static algorithm that attempts to solve a problem. Because the algorithm adjusts as it evaluates training data, the process of exposure and calculation around new data trains the algorithm to become better at what it does. Algorithms are the computational part of a machine learning project. Once trained, algorithms produce models with a statistical probability of answering a question or achieving a goal. Unlike in expert systems, the logic by which a machine learning model operates is not explicitly programmed; it is learned through experience. Machine learning has come to dominate the field of AI: it provides the backbone of most modern AI systems, from forecasting models to autonomous vehicles to large language models (LLMs) and other generative AI tools. Machine learning has become a household term in recent years as the concept moved from science fiction to a key driver of how businesses and organizations process information [6].

## HISTORICAL BACKGROUND

Actual efforts towards machine learning did not begin until the 1950s, at a time when computers were still in

their infancy and artificial intelligence could only be dreamed of. In 1950, Turing developed the Turing Test, a kind of game in which a computer pretends to be human. From then on, scientists began to trust their computers even more with more complex thinking tasks. The term “machine learning” was coined in 1959 by Arthur Samuel, an American pioneer in the field of computer gaming and artificial intelligence while working at IBM. More recently, in 1997, Tom M. Mitchell at Carnegie Mellon University provided a widely quoted, more formal definition of the algorithms studied in the machine learning field: “A computer program is said to learn from experience  $E$  with respect to some class of tasks  $T$  and performance measure  $P$  if its performance at tasks in  $T$ , as measured by  $P$ , improves with experience  $E$ ” [7]. Figure 2 show the evolution of machine learning [8].

## HOW MACHINE LEARNING WORKS

Machine learning uses algorithms—basically rule-based instructions—trained on past data to make predictions or classifications when faced with new input. Raw data is cleaned and transformed to train a machine learning model. The preprocessed data is used to train the machine learning algorithm. The algorithm tries to iteratively identify the mathematical correlation between the input and expected output from the training data.

Once ready, these trained algorithms become “machine learning models”—tools capable of tasks like recognizing images, forecasting housing prices, or even playing chess. The goal is to ensure the model can generalize beyond the training dataset. The central idea behind machine learning is an existing mathematical relationship between any input and output data combination. The machine learning model does not know this relationship in advance but can guess if sufficient examples of input-output data sets are given. In more advanced cases, multiple algorithms are stacked to form complex systems known as deep learning networks. Figure 3 shows how machine learning works [3].

## TYPES OF MACHINE LEARNING

Machine learning algorithms can be supervised or unsupervised. Basically, developers differentiate between supervised learning and unsupervised learning, with gradual intermediate stages. Besides these two main areas, there is also semi-supervised learning, reinforcement learning, and active learning. In addition, a distinction is made between shallow learning and deep learning. The relationship between some of these types of machine learning is shown in Figure 4 [9]. Different types of machine learning are explained below [10]:

- **Supervised Learning:** The program is “trained” on a pre-defined set of “training examples” from a “teacher,” which then facilitate its ability to reach an accurate conclusion when given new data. In this case, the data comes with additional attributes that we want to predict. A common case of supervised learning is to use historical data to predict statistically likely future events. Under supervised ML, we have regression ML and classification ML. Similarity learning is an area of supervised machine learning closely related to regression and classification, but the goal is to learn from examples using a similarity function that measures how similar or related two objects are.
- **Unsupervised Learning:** As their name suggests, unsupervised learning algorithms can be broadly understood as somewhat “optimizing themselves.” Unsupervised algorithms do not need to be trained with desired outcome data. The program is given a bunch of data and must find patterns and relationships therein. A typical goal of unsupervised learning may be as straightforward as discovering hidden patterns within a dataset. Without being told a “correct” answer, unsupervised learning methods can look at complex data and organize it in potentially meaningful ways. Unsupervised learning algorithms are used for more complex processing tasks than supervised learning systems.
- **Reinforcement Learning:** Reinforcement learning models are trained holistically through trial and error. Reinforcement learning is a method with reward values attached to the different steps that the algorithm must go through. So, the model’s goal is to accumulate as many reward points as possible and eventually reach an end goal. Reinforcement learning is an area of machine learning concerned with how software agents ought to take actions in an environment so as to maximise some notion of cumulative reward.
- **Classification Learning:** This involves seeking a yes-or-no prediction, such as “Does this product meet our quality standard.” We want to learn from already labeled data how to predict the class of unlabeled data. An example of classification problem would be the handwritten digit recognition example.
- **Regression Learning:** If the desired output consists of one or more continuous variables, then the task is called regression. In this case, the value being predicted falls somewhere on a spectrum. An example of a regression problem would be the

prediction of the length of a salmon as a function of its age and weight.

- **Deep Learning:** Deep learning (DL) is a specialized form of machine learning that uses artificial neural networks to mimic the human brain. It is a type of machine learning technique that is modeled on the human brain. It is an advanced technique for handling complex tasks like image and speech recognition. The way in which neural networks are trained can be described as deep learning. It is called deep because the network of neurons is arranged in several hierarchical levels. Deep learning laid the foundation for advances in generative artificial intelligence. Figure 5 shows the relationship between AI, ML, and DL: deep learning is a subset of machine learning, which is itself a subset of artificial intelligence [11].

## APPLICATIONS

Machine learning used to power applications like recommendation engines, fraud detection, and medical diagnostics by enabling systems to learn from experience and data, rather than being explicitly programmed for every task. There is a growing recognition that ML can play a key role in a wide range of critical applications such as data mining, natural language processing, image recognition, expert systems, simulation, agriculture, medicine, and manufacturing. Some of these application areas are shown in Figure 6 [12]. ML has been applied to each of these with great success. A sample of ML applications are provided next [10,13,14].

1. **Healthcare:** Using ML algorithms is a fast-growing trend in the healthcare industry. The proliferation of wearable sensors and devices has generated significant healthcare data. Wearable devices and sensors can be used to assess a patient's health in real time. The medical experts can analyze data to identify trends or red flags. ML algorithms can process more information and spot more patterns than their human counterparts. They have immense potential to enhance diagnostic and intervention research in smoking, depression, asthma, and chronic obstructive pulmonary disease (COPD) [6]. They can also be used to predict diabetes complications. ML has been applied to pharmacology for improved therapy. Figure 7 is a representation of machine learning in healthcare [15].
2. **Financial Services:** ML models have been successful in finance applications.

Financial machine learning projects improve risk analytics and regulation. Businesses in the

financial industry use ML technology for two key purposes: to prevent fraud and identify investment opportunities, or help investors know when to trade. Banks use it for credit-card fraud detection. ML is getting better and better at spotting potential cases of fraud. It can also be used to predict bankruptcy and credit risk management. A financial organization could train a machine learning system to classify fraudulent and genuine transactions. The system identifies patterns in known data to accurately guess or predict whether a new transaction is genuine. Financial services use machine learning and quantitative tools to make predictions about their prospects and customers. With this information, banks can answer questions like who is likely to default on a loan, which customers pose high or low risks, and which customers are the most lucrative to target resources.

3. *Retail:* For retailers, machine learning can help operations, sales, and more. Retail can use machine learning to improve customer service, stock management, upselling, and cross-channel marketing. For example, Amazon Fulfillment (AFT) cut infrastructure costs by 40 percent using a machine learning model to identify misplaced inventory. This helps them deliver on Amazon's promise that an item will be readily available to customers and arrive on time.
4. *Media and Entertainment:* Entertainment companies turn to machine learning to better understand their target audiences and deliver immersive, personalized, and on-demand content. For example, Disney uses machine learning to archive its media library. Machine learning tools automatically tag, describe, and sort media content, enabling Disney writers and animators to quickly search for and familiarize themselves with Disney characters.
5. *Transportation:* The transportation industry relies on making routes more efficient and predicting potential problems. The data analysis and modeling aspects of ML are important tools for public transportation and other transportation organizations.
6. *Image Recognition:* This is one of the most common uses of ML. There are also speech recognition, facial detection/recognition, object recognition, and character recognition. In speech recognition (also known as computer speech recognition), a software application recognizes spoken words.
7. *Manufacturing:* This is an area where the application of machine learning is very fruitful. The manufacturing industry today is facing an increasing volume of data which compromise a variety of different formats, semantics, and quality. ML techniques have been successfully utilized in various process optimization, monitoring and control applications in manufacturing. Machine learning can support predictive maintenance, quality control, and innovative research in the manufacturing sector. For example, manufacturing giant 3M uses machine learning to innovate sandpaper.
8. *Marketing:* Machine learning already has important functions for marketing. At the moment, however, it is primarily large companies that use the functions internally, such as Google.
9. *Science:* The intelligent processing of big data is an enormous relief for empirically working scientists. Machine learning also helps in medicine: even today, some doctors use artificial intelligence for diagnosis and treatment. Machine learning is also used for forecasting illnesses such as diabetes and heart attacks.
10. *Robots:* Robots are everywhere, especially in factories. They help, for example, with mass production if they are programmed to carry out consistent work steps. If self-learning systems are used in robotics, these machines should also be able to master new tasks. These developments are, of course, also very interesting for other areas: from space travel to the home, robots with artificial intelligence will be used in numerous areas.
11. *Autonomous Vehicles:* One of machine learning's big flagship products is the self-driving or autonomous car. Self-driving cars, a wonder of the 21st century, rely on deep learning models, as a specialized form of machine learning, to process sensor data, recognize road conditions, and make real-time driving decisions. Vehicles can maneuver themselves through actual traffic without causing accidents. This can only be achieved through machine learning, since it is not possible to program all situations that could possibly occur. For this reason, it is imperative for cars to be able to navigate themselves using intelligent machines.
12. *Personal Assistants:* Computer systems that are able to constantly learn also play an important role in your home. This is how simple homes become smart homes. Personal assistants such as Google Home and Amazon Echo also use

machine learning technologies to help them better understand their users. But many people now carry their assistants with them at all times: With Siri, Cortana, or Google Assistant, users can use voice control to send commands and questions to their smartphones.

13. *Gaming*: Since research began on artificial intelligence, scientists have been greatly motivated by how well machines are able to play games. Self-learning systems pitted themselves against humans in games of chess, checkers, and Go from China. Computer game developers also use machine learning to make their games more interesting. Game designers can use machine learning to create the most balanced gameplay possible.

Other potential applications include simulation, face recognition, object recognition, speech recognition, natural language processing (NLP), optical character recognition, affective computing, Internet fraud, medical diagnosis, IoT, stock market prediction, economics, automotive, defense and security, government, insurance, utilities, oil and gas, advertisement, email spam detection, drug recognition, education, social science, linguistics, management, computer vision, materials science, smart city, and remote sensing, and communication networks such wireless sensor networks and mobile ad-hoc networks (MANETs).

## BENEFITS

Machine learning technology can process large quantities of historical data, identify patterns, and predict new relationships between previously unknown data. It leverages data and algorithms to enable AI systems to learn and improve in a manner similar to humans, progressively enhancing their accuracy over time. A good machine learning model is crucial to ML applications. Other benefits of machine learning include the following [8,16]:

- *Automation*: Machine learning makes it easier to automate operations. Machine learning automates and optimizes the process of data collection, classification, and analysis. Machine learning algorithms can filter, sort, and classify data without human intervention. They can summarize reports, scan documents, transcribe audio, and tag content—tasks that are tedious and time-consuming for humans to perform. Since modern systems can independently adapt to new conditions with the help of machine learning, complex automations processes are also possible.
- *Limitations*: Although machine learning has been transformative in some fields, machine-learning

programs often fail to deliver expected results. Reasons for this are numerous: lack of (suitable) data, lack of access to the data, data bias, privacy problems, badly chosen tasks and algorithms, wrong tools and people, lack of resources, and evaluation problems. For example, attempts to use machine learning in healthcare with the IBM Watson system failed to deliver even after years of time and billions of dollars invested.

- *Enhanced Decision-making*: Machine learning systems can process and analyze massive data volumes quickly and accurately. Organizations can make data-driven decisions at runtime and respond more effectively to changing conditions. They can optimize operations and mitigate risks with confidence.
- *Personalization*: Machine learning has opened a new door for customer experiences through personalization. Machine learning transforms customer experiences through personalization. For example, retailers recommend products to customers based on previous purchases, browsing history, and search patterns. The personalized approach increases customer retention and brand loyalty.
- *Continuous Improvement*: A distinctive advantage of machine learning is its ability to improve as it processes more data. Machine learning systems adapt and learn from new data. They adjust and enhance their performance to remain practical and relevant over time.
- *Generative AI*: Generative AI tools like ChatGPT, Google Gemini, and Microsoft Copilot are becoming increasingly prevalent in professional settings. Technically, generative AI relies on sophisticated machine learning methods. However, its practical use differs from traditional machine learning. While conventional models are typically designed to carry out specific, repetitive tasks, generative AI focuses on producing novel and adaptive outputs in response to user input.
- *Predictive Analytics*: Businesses are using machine learning to look into the future. Predictive analytics uses historical data to forecast trends, helping companies make informed decisions. Whether it is predicting customer demand or identifying potential risks, machine learning helps businesses stay ahead of the curve. While descriptive analytics focuses on analyzing historical data and deriving inferences from them, predictive analytics focuses on predicting and understanding what could happen in the future.

- **Predictive Maintenance:** By predictive maintenance, machine learning models can analyze vehicle performance data to detect potential mechanical failures before they occur. ML uses patterns learned from training data to make predictions about new, unseen data. By analyzing data obtained from the different machines, models can be generated that are capable of predicting when a failure will occur. This serves to improve processes and prevent failures before machines break down. In 2006, the media-services provider Netflix held the first "Netflix Prize" competition to find a program to better predict user preferences and improve the accuracy of its existing algorithm by at least 10%. When maintenance and repair data is collected manually, it is almost impossible to predict potential problems, let alone automate processes to predict and prevent them.
- **Data Compression:** Data compression aims to reduce the size of data files, enhancing storage efficiency and speeding up data transmission. There is a close connection between machine learning and compression. An optimal compressor can be used for prediction. An alternative view can show compression algorithms implicitly map strings into implicit feature space vectors, and compression-based similarity measures compute similarity within these feature spaces.
- **Data Mining:** Machine learning and data mining often employ the same methods and overlap significantly. While machine learning focuses on prediction, based on *known* properties learned from the training data, data mining focuses on the discovery of (previously) *unknown* properties in the data. Data mining uses many machine learning methods, but with different goals; on the other hand, machine learning also employs data mining methods as "unsupervised learning" to improve learner accuracy.

## CHALLENGES

A major challenge of increasing importance is the question what ML technique and model to choose. Most ML algorithms are time-intensive and energy-inefficient since they require several days or months to train a good model. The algorithms can be accelerated with graphic processing units (GPUs) and field-programmable gate arrays (FPGAs). Although the complexity of ML has restricted its use, the supply of competent ML designers has yet to catch up to this demand. Other challenges in machine learning implementation include the following [16,17]:

- **Data Quality:** The adage "garbage in, garbage out" applies to machine learning—the quality of

data is critical, during both the training phase and in production. Typically, machine learning models require a high quantity of reliable data to perform accurate predictions. High-quality data can lead to more accurate results delivered in a timely, efficient manner; low-quality data can create inaccuracies and distortion in resultant models. A machine learning model's performance depends on the data quality used for training. Issues such as missing values, inconsistent data entries, and noise can significantly degrade model accuracy. Lack of a sufficiently large dataset can prevent the model from learning effectively. Typically, machine learning models require a high quantity of reliable data to perform accurate predictions.

- **Data Security:** Despite its many benefits, machine learning can introduce a range of security issues. The data used in ML analysis may contain sensitive or proprietary information not meant for public consumption. Data preparation steps can both expose and address security vulnerabilities, particularly when data sets go through export or import processes between systems. Machine learning can enhance data security by detecting and responding to cybersecurity threats in real-time. ML algorithms can identify unusual patterns that may indicate a data breach, ensuring patient data remains protected.
- **Data Privacy:** One of the biggest concerns in machine learning is data privacy. For machine learning models to be effective, they need vast amounts of data—often personal or sensitive information. This raises important questions about how that data is collected, stored, and used. Ensuring sensitive data is not disclosed is an ongoing effort. Data anonymization is an emerging practice, but it may not always be available or sufficient. Information needs to be protected and used consistent with legal requirements. For retailers, machine learning can help operations, sales, and more.
- **Overfitting and Underfitting:** Overfitting occurs when a machine learning model learns the details and noise in the training data to the extent that it negatively impacts the model's performance on new data. On the other hand, underfitting happens when a model cannot learn the underlying pattern of the data, resulting in poor performance on both the training and testing data. Balancing the model's complexity and its ability to generalize is a critical challenge. Even among ML experts, selecting a good model can be a challenging task.

- *Bias*: Different machine learning approaches can suffer from different data biases. Algorithmic bias is a potential result of data not being fully prepared for training. The data may be imbalanced in many real-world applications. This imbalance can bias the training process, causing the model to perform well on the majority class while failing to predict the minority class accurately. For example, if historical data prioritizes a certain demographic, machine learning algorithms used in human resource applications may continue to prioritize those demographics.
- *Model Explainability*: Another challenge is the interpretation of the results. As machine learning models, particularly deep learning models, become more complex, their decisions become less interpretable. Developing methods to make models more interpretable without sacrificing performance is an important challenge.
- *Scalability*: Machine learning models, especially those that involve large datasets or complex algorithms like deep learning, require significant computational resources. Training these models can be time-consuming and costly. Optimizing algorithms to reduce computational demands involves challenges in algorithm design.

## CONCLUSION

Machine learning is a type of artificial intelligence (AI) where algorithms learn from data to identify patterns and make predictions or decisions with minimal human intervention.

It is the process of teaching computers to automatically recognize patterns of interest in data. It is an incredibly powerful tool that is used in a wide range of compelling application domains because it is applicable to many real-life problems. Machine learning is already transforming much of our world for the better. The journey of machine learning is just beginning, and the future holds incredible promise.

Machine learning is set to be a pillar of our future civilization. It requires a strong foundation in mathematics, statistics, coding, and data technologies. Those wishing to advance in machine learning should consider completing a master's degree in artificial intelligence or data science. They can use online courses to learn at their own pace and master specific skills. Additional information on machine learning is available in the books in [13,18-31].

## REFERENCES

- [1] "Understanding machine learning: Uses, example," April 2025,

<https://www.investopedia.com/terms/m/machine-learning.asp>

- [2] "What is machine learning? Definition, types, and examples," August 2025, <https://www.coursera.org/articles/what-is-machine-learning>
- [3] N. Nguyen, "Understanding the core concept of machine learning," June 2025, <https://shiftdasia.com/community/understanding-the-core-concept-of-machine-learning/>
- [4] "Machine learning" *Wikipedia*, the free encyclopedia [https://en.wikipedia.org/wiki/Machine\\_learning](https://en.wikipedia.org/wiki/Machine_learning)
- [5] S. Shalev-Shwartz and S. Ben-David, *Understanding Machine Learning: From Theory to Algorithms*. New York, NY: Cambridge University Press, 2014.
- [6] M. N. O. Sadiku, S. M. Musa, and O. S. Musa, "Machine learning," *International Research Journal of Advanced Engineering and Science*, vol. 2, no. 4, 2017, pp. 79-81.
- [7] S. Mullainathan and J. Spiess, "Machine learning: An applied econometric approach," *Journal of Economic Perspectives*, vol. 31, no. 2, Spring 2017, pp. 87-106.
- [8] "Machine learning in industry," <https://atriainnovation.com/en/blog/machine-learning-in-industry/>
- [9] D. Shewan, "10 Companies using machine learning in cool ways," November 2023, <https://www.wordstream.com/blog/ws/2017/07/28/machine-learning-applications>
- [10] T. Wuest et al., "Machine learning in manufacturing: Advantages, challenges, and applications," *Production & Manufacturing Research*, vol. 4, no. 1, 2016, pp. 23-45.
- [11] "What is machine learning? The ultimate beginner's guide," June 2021, <https://www.v7labs.com/blog/machine-learning-guide>
- [12] "What is machine learning? Key concepts and real-world uses," February 2025, <https://ischool.syracuse.edu/what-is-machine-learning/>
- [13] E. Alpaydm, *Introduction to Machine Learning*. Cambridge, MA: The MIT Press, 3rd edition, 2014.
- [14] <https://www.ionos.com/digitalguide/online-marketing/web-analytics/what-is-machine-learning-how-machines-learn-to-think/>

- [15] “Benefits of machine learning in healthcare,” <https://www.foreseemed.com/blog/machine-learning-in-healthcare>
- [16] “What is machine learning?” [https://aws.amazon.com/what-is/machine-learning/?trk=80384a35-d261-4dc4-ae07-058697ea891c&sc\\_channel=ps&ef\\_id=EAIAIQobChMIgbifjFWhkAMVbUH\\_AR09QBjwEAA YASAAEgJp9vD\\_BwE:G:s&s\\_kwid=AL!4422!3!751258939107!e!!g!!machine%20learning!22542047959!178682823585&gad\\_campaignid=22542047959&gbraid=0AAAAADjHtp96vt\\_Q7N2DALMz6zQTWDGDN&gclid=EAIAIQobChMIgbifjFWhkAMVbUH\\_AR09QBjwEAA YASAAEgJp9vD\\_BwE](https://aws.amazon.com/what-is/machine-learning/?trk=80384a35-d261-4dc4-ae07-058697ea891c&sc_channel=ps&ef_id=EAIAIQobChMIgbifjFWhkAMVbUH_AR09QBjwEAA YASAAEgJp9vD_BwE:G:s&s_kwid=AL!4422!3!751258939107!e!!g!!machine%20learning!22542047959!178682823585&gad_campaignid=22542047959&gbraid=0AAAAADjHtp96vt_Q7N2DALMz6zQTWDGDN&gclid=EAIAIQobChMIgbifjFWhkAMVbUH_AR09QBjwEAA YASAAEgJp9vD_BwE)
- [17] M. Chen, “What is machine learning?” November 2024, <https://www.oracle.com/artificial-intelligence/machine-learning/what-is-machine-learning/>
- [18] G. Rebala, A. Ravi, and S. Churiwala, *An Introduction to Machine Learning*. Springer, 2019.
- [19] Y. Kodratoff, *An Introduction to Machine Learning*. Morgan Kaufman Publ., 1988.
- [20] B. Boehmke and B. M. Greenwell, *Hands-On Machine Learning with R*. Boca Raton, FL: CRC Press, 2019.
- [21] T. Danka, *Mathematics of Machine Learning: Master Linear Algebra, Calculus, and Probability for Machine Learning*. Packt Publishing, 2025.
- [22] J. James, *The Ultimate Guide to Machine Learning With Python Programming for Beginners: Mastering Python Fundamentals for Machine Learning With Practical Coding ... Chapter To Bring Yourself to the Next Level*. Independently Published, 2025.
- [23] O. Theobald, *Machine Learning for Absolute Beginners: A Plain English Introduction*. Independently Published, third edition, 2021.
- [24] C. Huyen, *Designing Machine Learning Systems: An Iterative Process for Production-Ready Applications*. O'Reilly Media, 2022.
- [25] Y. H. Liu, *Python Machine Learning by Example: Unlock Machine Learning Best Practices With Real-World Use Cases*. Packt Publishing, 4<sup>th</sup> edition, 2024.
- [26] P. Singh, *Fundamentals and Methods of Machine and Deep Learning Algorithms, Tools, and Applications*. Wiley, 2022.
- [27] S. Hack, *Machine Learning for Beginners A Math Guide to Mastering Deep Learning and Business Application. Understand How Artificial Intelligence, Data Science, and Neural Networks Work Through Real Examples*. Chopra International Consulting Limited, 2021
- [28] Ajayi, A.S., Kim, S. & Yun, R. Study of developing a condensation heat transfer coefficient and pressure drop model for whole reduced pressure ranges. *Int. J. Air-Cond. Ref.*32, 15 (2024).
- [29] Ajayi, A., Badmus, O., Iheuwa, G., Ehizojie, L., & Segun, S. (2025). Comparative Analysis of GitOps Tools and Frameworks. *Path of Science*, 11(5), 2017-2030.
- [30] Ajayi, A. S., Abimbola, J., Segun, S. E., Ajayi, E. I., & Bodude, A. D. (2026). Harnessing artificial intelligence for methane emissions control in industrial natural gas engines: Optimizing exhaust after treatment to advance U.S. clean energy goals—A review. *European Journal of Sustainable Development Research*, 10(1), em0345.
- [31] Oyetunji, O. R., Ajayi, A. Samuel, Amuda, B. A., & Morawo, I. I. (2023). Comparative study of mechanical properties of 3D printing materials (polylactic acid and acrylonitrile butadiene styrene) via simulations using COMSOL Multiphysics. *Advances in Multidisciplinary and Scientific Research Journal*, 9(3), 21–34.

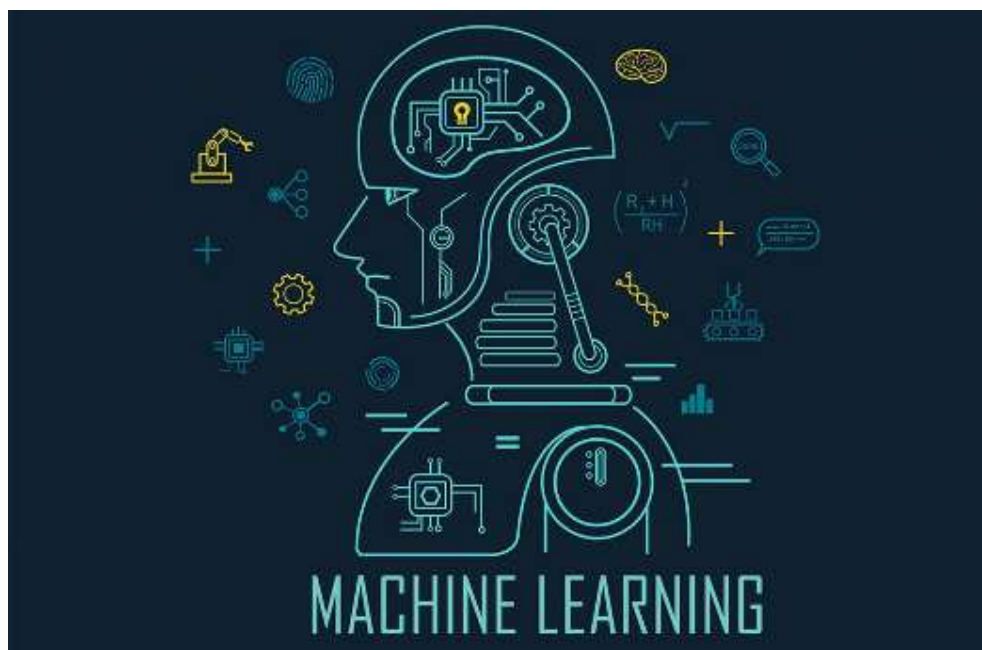


Figure 1 A symbol of machine learning [3].

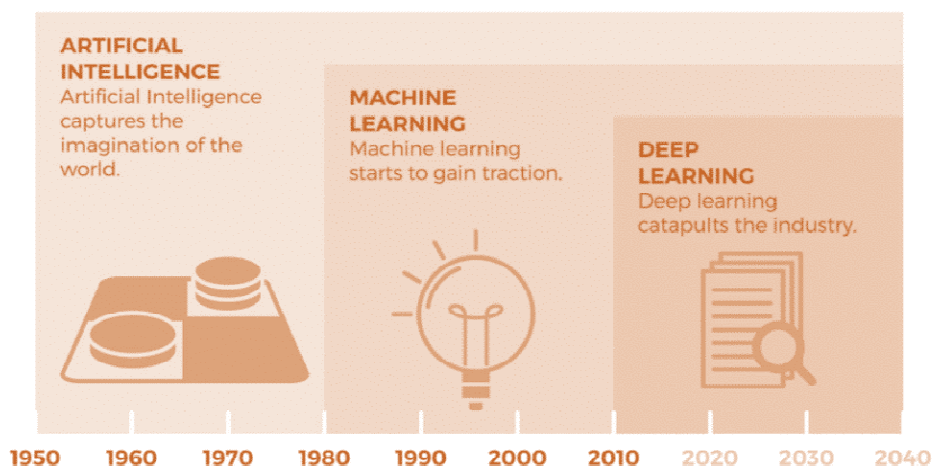


Figure 2 Evolution of machine learning [8].

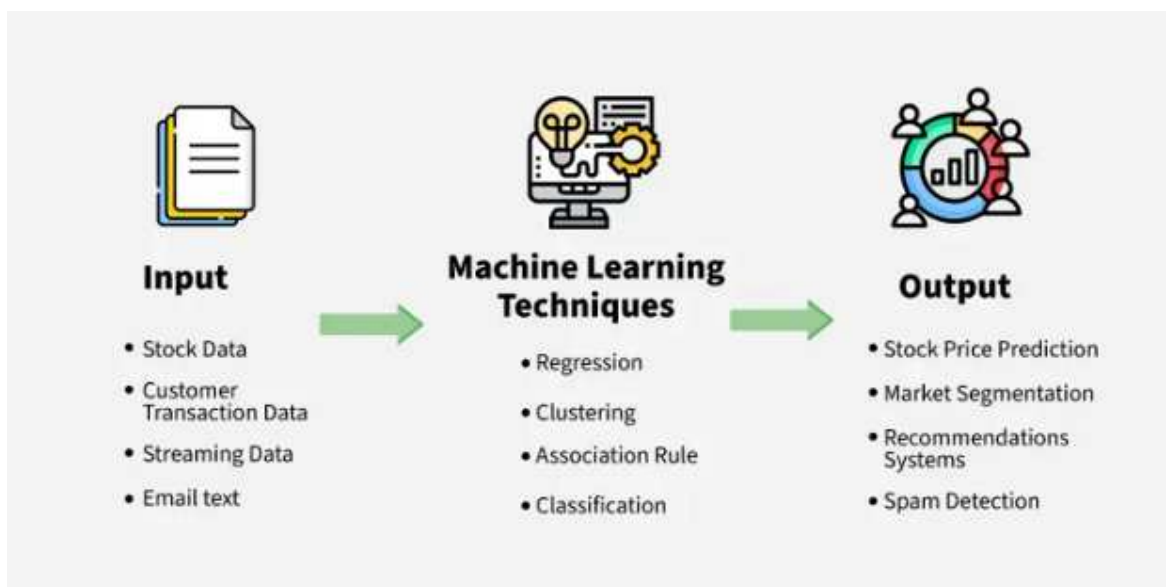
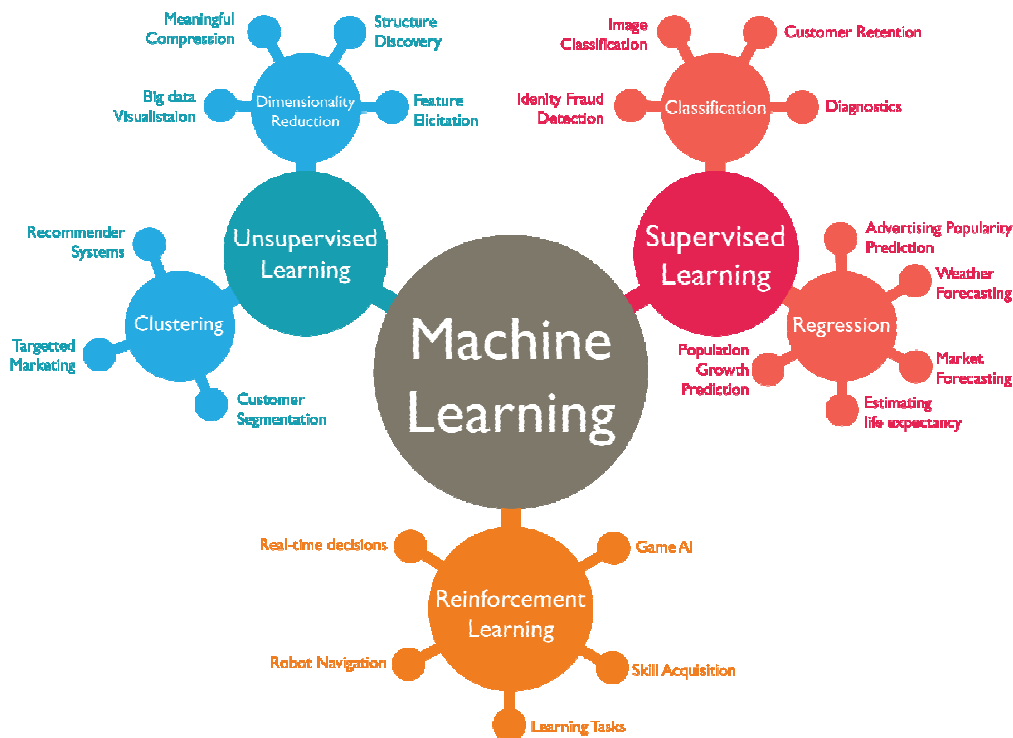
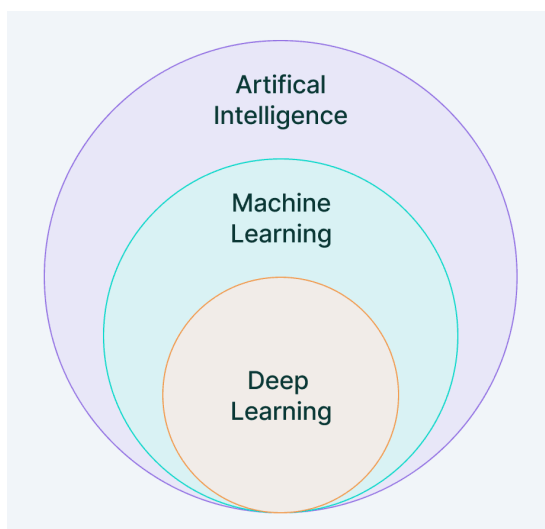


Figure 3 How machine learning works [3].



**Figure 4 The relationship between some types of machine learning [9].**



**Figure 5 The relationship between AI, ML, and DL [10].**

### **Machine Learning Applications: Industries**



**Figure 6 Some applications of machine learning [12].**



**Figure 7 A representation of machine learning in healthcare [15].**

