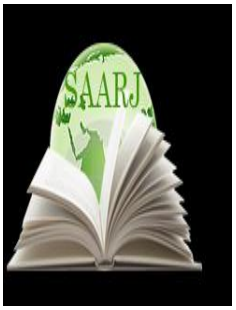
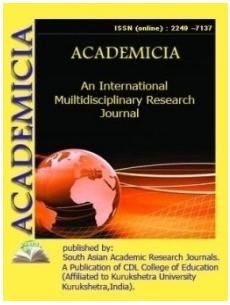


ACADEMICIA
An International Multidisciplinary Research Journal
(Double Blind Refereed & Peer Reviewed Journal)



DOI: 10.5958/2249-7137.2021.02181.9

AN OVERVIEW OF MACHINE LEARNING FROM THEORY TO ALGORITHMS

Mr Mrinal Paliwal*

*SBAS, Sanskriti University, Mathura,
Uttar Pradesh, India, INDIA
Email id: mrinalpaliwal.cse@sanskriti.edu.in

ABSTRACT

The present SMAC (Social, Mobile, Analytical, and Cloud) technological trend points to a future in which intelligent machines, networked processes, and big data are all combined. The enormous quantity of data produced by this virtual environment is hastening the adoption of machine learning solutions and techniques. Computers can mimic and modify human-like behaviour thanks to machine learning. Each interaction, each action done, becomes something the system may learn and utilize as experience for the next time using machine learning. This paper provides an introduction of a data analytics technique that allows computers to learn and do what people do naturally, namely, learn from experience. It covers the fundamentals of machine learning, including definitions, terminology, and applications that explain what, how, and why it works. Machine learning's technological roadmap is explored in order to better comprehend and validate its potential as a market and industrial practice. The main goal of this paper is to explain why machine learning is the way of the future.

KEYWORDS: *Algorithms, Ensemble Learning, Instant Learning, Machine Learning, Supervised Learning, Unsupervised Learning.*

1. INTRODUCTION

Learning is the process of learning new behaviours, beliefs, information, abilities, or preferences, or altering old ones. The philosophy of personal learning, or how humans learn, is defined by behaviourism, cognitivism, constructivism, experientialism, and social learning. Machines, in contrast to humans, rely on data instead of learning through experience[1]. Machine learning (ML) is a type of artificial intelligence that allows computers to think and learn on their own, at its most basic level. It's all about getting computers to change their activities in order to increase

their accuracy, with accuracy being defined as the number of times the chosen actions result in right behaviours[2].ML has been explicitly defined by researchers in the relevant literature. Arthur Samuel created the phrase in 1959, defining it as an area of research that allows computers to learn without having to be explicitly programmed. Tom Mitchell recently provided a "well-posed" term that has shown to be more effective in engineering set-up: If a computer program's performance on a task increases with experience, it is said to learn from experience with regard to that task and that performance metric.

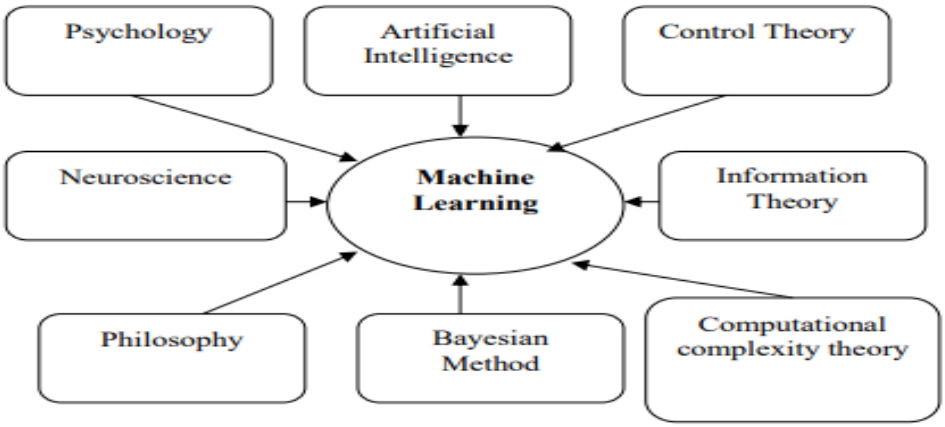


Fig. 1: The Multi-disciplinary Machine Learning (ML) fields[3]**.**

Machine learning is a multidisciplinary area with a diverse set of study fields supporting it. As illustrated in Fig. 1, there are the following. Computational Statistics, whose major goal is to make predictions using computers, is closely connected to the simulation of ML models. It's also linked to Mathematical Optimization, a branch of statistics that deals with models, applications, and frameworks. The enormous complexity of real-world issues makes them ideal candidates for machine learning. Machine learning can be used to design and program explicit algorithms with high performance output in a variety of areas of computing, such as email spam filtering, social network fraud detection, online stock trading, face & shape detection, medical diagnosis, traffic prediction, character recognition, and product recommendation, to name a few. Machine learning is used in real-world applications such as self-driving Google cars, Netflix recommending movies and shows a person might like, online recommendation engines such as Facebook friend suggestions, Amazon's "more items to consider" and "get yourself a little something," and credit card fraud detection. The major goal of this paper is to provide an overview of Machine Learning's evolution to the current day, as well as numerous machine learning methods, applications, and problems. This paper provides an overview of machine learning, including its history, origins, and evolution over the decades to the current day. The research discusses the challenges in data science that machine learning techniques may solve, as well as the progress of machine learning to the present day. It also includes a description of the general model of machine learning, as well as an explanation of the machine learning process. It also goes through the major machine learning paradigms and methods, as well as a quick rundown of the ML difficulties.

2. DISCUSSION

2.1. Data Science Problems and Machine Learning:

Machine learning is necessary to make computers smart enough to do tasks without human involvement, based on learning and continuously expanding experience to comprehend the issue complexity and requirement for flexibility[4].

2.1.1. Tasks Performed by Human Beings:

Human people execute a variety of activities on a daily basis, but the key issue is to complete the duties flawlessly and according to a well-defined schedule. Cooking, driving, and speech recognition are among examples.

2.1.2. Tasks Beyond Human Capabilities:

Another group of activities that machine learning can effectively do is the analysis of big and complicated data sets, such as remote sensing, weather forecasting, ecommerce, and web search. It becomes extremely difficult for humans to forecast relevant data when there is a vast volume of data.Machine learning has demonstrated its ability to tackle data science issues on its own. Data science, according to Hayashi and Chikio, is a concept that combines statistics, data analysis, machine learning, and related methodologies in order to comprehend and analyze real-world events using data [5]. Before beginning to solve an issue, it must be properly classified so that the most appropriate machine learning method may be used. As illustrated in Fig. 2, any data science challenge may be classified into one of the five categories.

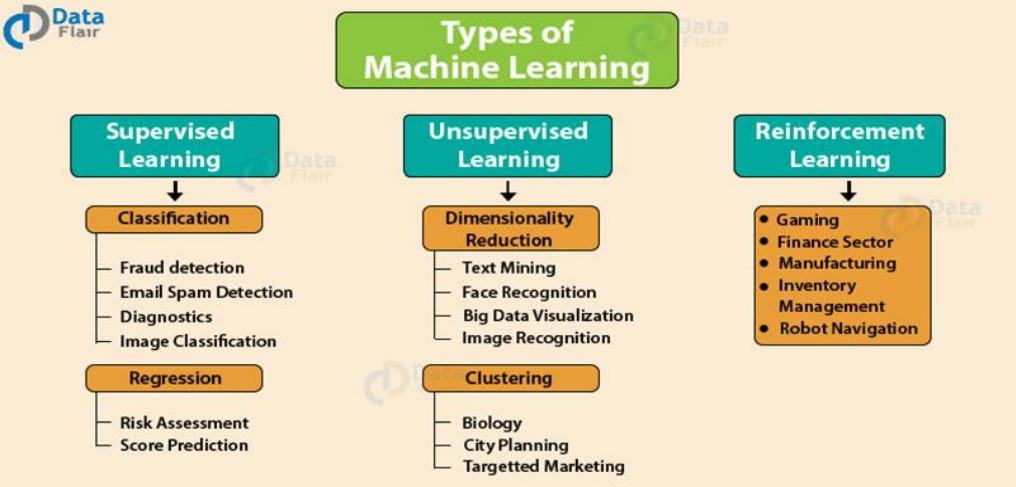


Fig. 2: Types of Machine Learning techniques along with the types of problems they are suitable for[6]**.**

As a result, depending on the kind of challenge, a machine learning technique can be used. The following are the various categories:

- *Classification Problem:*

A classification issue is one in which the output can only be one of a set number of output classes known a priori, such as Yes/No or True/False. The problem might be a binary or multi-class classification problem, depending on the number of output classes.

- *Anomaly Detection Problem:*

This category includes problems that examine a pattern and discover changes or abnormalities in the pattern. Credit card firms, for example, employ anomaly detection algorithms to detect deviations from their clients' typical transaction behaviour and issue warnings whenever there is an unexpected transaction. Finding outliers is one of these challenges.

- *Regression Problem:*

When dealing with situations that have both continuous and numeric output, regression methods are utilized. These are typically used for situations involving queries such as "how much" or "how many."

- *Clustering Problem:*

Clustering is classified as an unsupervised learning technique. These algorithms attempt to learn structures inside the data and create clusters based on data structure similarities[1]. After that, the various classes or clusters are labelled. When the algorithm is trained, it adds fresh, previously unseen data to one of the clusters.

- *Reinforcement Problem:*

When a choice must be made based on previous learning experiences, reinforcement algorithms are utilized. The machine agent learns its behaviour by interacting with the constantly changing environment in a trial and error manner. It allows agents to be programmed utilizing the idea of rewards and punishments without having to explain how the job should be completed. Reinforcement learning is used in a variety of applications, including game play and temperature management.

2.2. Development of Machine Learning:

TABLE 1: ILLUSTRATES DEVELOPMENT OF MACHINE LEARNING (ML)[3]**.**

Year	Development
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 Neighbours 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.
1985	NetTalk was invented by Terry Sejnowski, which learnt to pronounce English words in the same manner that children learn.
1990	The focus of Machine Learning shifted from Knowledge-driven to Data Driven. Machine Learning was implemented to analyse large chunks of data and derive conclusions from it.
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 is a Deep Neural Network to detect patterns in Videos and Images.
2014	Facebook invented the "DeepFace" algorithm based on Deep Neural Networks capable of 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. 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.
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

Artificial Intelligence (AI) and Machine Learning (ML) are not new terms. For more than 60 years, computer scientists, engineers, researchers, students, and industry experts have investigated, used, used, and re-invented them[7]. Algebra, statistics, and probability are the mathematical foundations of machine learning. Researchers such as Alan Turing, John McCarthy, Arthur Samuels, Alan Newell, and Frank Rosenblatt contributed to the serious development of Machine Learning and Artificial Intelligence in the 1950s and 1960s. On the Optimizing Checkers Program, Samuel proposed the first practical machine learning model. Rosenblatt invented the Perceptron, a widely used machine learning method based on biological neurons that lay the groundwork for Artificial Neural Networks. The illustrious, expansive, and practical growth of machine learning is depicted in Table 1.

2.3. The Generic Model of Machine Learning:

Computer learning is used to tackle a variety of issues that need the machine to learn. There are three characteristics of a learning problem:

- The process of gaining experience.
- Performance measure to be improved.
- Task classes.

2.4. Machine Learning Paradigms:

2.4.1. Supervised Learning:

Prediction based on historical data may benefit from supervised learning. For example, a recognition system that identifies whether an item is a galaxy, a quasar, or a star given a coloured picture of an object via a telescope, or given a person's e-commerce browsing history, product suggestion by e-commerce websites.

2.4.2. Unsupervised Learning:

The goal of unsupervised learning is to recognize previously undiscovered patterns in data in order to generate rules from them. In situations when the data categories are uncertain, this approach is suitable. The training data is not labelled in this case. Unsupervised learning is a statistic-based learning method that addresses the issue of uncovering latent structure in unlabelled data.

2.4.3. Reinforcement Learning:

Because the algorithm is simply given a response that indicates whether the output is accurate or not, reinforcement learning is considered an intermediate kind of learning[8]. To arrive at the right result, the algorithm must investigate and reject out different alternatives. Because the algorithm makes no recommendations or answers to the issue, it is referred to as learning with a critic.

2.4.4. Evolutionary Learning:

Biological creatures that adapt to their surroundings are the source of inspiration. The algorithm learns from the behaviour and adjusts to the inputs, excluding solutions that are improbable. It proposes the optimal solution to the issue based on the concept of fitness.

2.4.5. Semi-Supervised Learning:

These algorithms provide a method for combining the benefits of both supervised and unsupervised learning. In the preceding two kinds of output labels, either all of the observations are labelled or none are labelled. Some observations may be labelled, but the bulk of observations remain unlabelled owing to the high expense of labelling and a lack of competent human knowledge. Semi-supervised algorithms are ideally suited for model development in such circumstances. Semi-supervised learning may be used to classification, regression, and prediction issues.

2.4.6. Ensemble Learning:

It's a machine learning model in which a large number of students (individual models) are taught to answer a common issue[9]. Unlike other machine learning techniques, which try to learn by constructing a set of hypotheses from the training data and combining them to make a prediction model in order to reduce bias, variance, or improve predictions, ensemble learning acquires knowledge by constructing a set of hypotheses from the training data and incorporating them to make a statistical method in order to reduce bias, variability, or improve predictions.

2.5. Machine Learning Algorithms:

In this part, we'll look at some of the most prominent machine learning algorithms from the various paradigms discussed before. Despite the fact that each paradigm has a large number of algorithms that have been published in the literature, we only look at a few of them in this research[10]. A handful of these algorithms are briefly explained in Table 2. These algorithms have a broad range of applications, some of which are discussed below.Table 3 illustrates machine learning applications.

TABLE 2: ILLUSTRATES MACHINE LEARNING (ML) ALGORITHMS[3]**.**

Paradigm	Algorithm	Description
Supervised Learning	Decision Tree	The learned function is represented in the form of a decision tree in the Decision Tree method for approximating discrete valued target functions. On the basis of feature values, a decision tree classifies instances by sorting them from root to leaf nodes. Every branch indicates a potential value for that feature, whereas each node represents a choice on a property of the instance. The decision node, which is the root node, is where an instance's classification begins. The tree traverses down along the edge that corresponds to the value of the result of feature test based on the value of node. In the sub-tree led by the new node at the end of the preceding edge, this process continues. Finally, the categorization categories or the final choice are represented by the leaf node. When utilizing a decision tree, the emphasis is on determining which characteristic at each node level is the best classifier. For each node, statistical measures like as information gain, Gini index, Chi-square, and entropy are computed to

		determine its value. Decision trees are implemented using a variety of algorithms.
	Naïve Bayes	Bayes' theorem of probability is used by Nave Bayes to classify. The Bayes theorem determines the posterior probability of an event (A) given a prior probability of event B denoted by P(A/B).
	Support Vector Machines	SVMs are useful for both classification and regression issues. It's a learning algorithm that's supervised. It is based on the idea of calculating margins. Each data item is plotted as a point in n-dimensional space in this method (where n is the number of features we have in our dataset). Each feature's value is the associated coordinate's value. It divides the data into classes by identifying a line that divides the training datasets into categories. It operates by maximizing the margin between the closest data point and the hyper plane.
	Regression Analysis	Regression analysis is a kind of predictive modelling that looks at the connection between a dependent and an independent variable. It is a crucial tool for data analysis and modelling. We attempt to fit the line or curve to the data points in this technique to minimize the disparities in data point distances from the curve or line. Regression analysis is divided into three types: linear, logistic, and polynomial.
Unsupervised Learning	K-Means Clustering	For cluster analysis, K-means is a common unsupervised machine learning method. Its aim is to divide 'n' observations into 'k' clusters, with each observation belonging to the cluster with the closest mean, which serves as the cluster's prototype. The cluster's centre is determined by the mean of the observations in that cluster.
Instance based Learning	K-nearest Neighbours	It is a non-parametric classification and regression technique. The KNN method finds the k-nearest neighbours of an unknown feature vector whose class has to be determined given N training vectors.
Ensemble Learning	Random Forest	It's a classification and regression technique that uses ensemble learning. It creates a lot of decision trees using a random selection of data using a bagging technique. To create the final decision trees, the output of all decision trees in the random forest is merged.

Dimensionality Reduction	Principal Component Algorithm	It is mainly used to reduce the number of dimensions in a data collection. It aids in the reduction of the number of characteristics in the data collection as well as the number of independent variables. It converts correlated data into a collection of linearly uncorrelated variables called principle components via orthogonal transformation.

2.6. Applications of Machine Learning:

TABLE 3: ILLUSTRATES MACHINE LEARNING APPLICATIONS.

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.

3. 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.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.

REFERENCES

1. A. Dey, "Machine Learning Algorithms: A Review," *Int. J. Comput. Sci. Inf. Technol.*, 2016.

2. S. Ding, H. Zhao, Y. Zhang, X. Xu, and R. Nie, "Extreme learning machine: algorithm, theory and applications," *Artif. Intell. Rev.*, 2015, doi: 10.1007/s10462-013-9405-z.

3. J. Alzubi, A. Nayyar, and A. Kumar, "Machine Learning from Theory to Algorithms: An Overview," 2018, doi: 10.1088/1742-6596/1142/1/012012.

4. K. Gibson, "The Value of Data Science and Machine Learning.," *Risk Manag.*, 2018.

5. C. Hayashi, "What is Data Science ? Fundamental Concepts and a Heuristic Example," 1998.

6. "Machine Learning." https://data-flair.training/blogs/types-of-machine-learning-algorithms/ (accessed Sep. 07, 2018).

7. Gilpress, "Visually Linking AI, Machine Learning, Deep Learning, Big Data and Data Science," *What's the Big Data?* 2016.

8. J. García and F. Fernández, "A comprehensive survey on safe reinforcement learning," *Journal of Machine Learning Research*. 2015.

9. V. Smolyakov, "Ensemble Learning to Improve Machine Learning Results," *Stats Bots*, 2017.

10. J. Deepika, T. Senthil, C. Rajan, and A. Surendar, "Machine learning algorithms: A background artifact," *Int. J. Eng. Technol.*, 2018, doi: 10.14419/ijet.v7i1.1.9214.