

# Machine Learning: A Brief Review for the Beginners

Haradhan Kumar Mohajan<sup>1</sup>

<sup>1</sup> Chairman and Associate Professor, Department of Mathematics, Premier University, Chittagong, Bangladesh

Correspondence: Haradhan Kumar Mohajan, Chairman and Associate Professor, Department of Mathematics, Premier University, Chittagong, Bangladesh.

doi:10.63593/IST.2788-7030.2026.03.004

## Abstract

Machine learning (ML) is a branch of artificial intelligence (AI) that focuses on developing models, studies statistical algorithm, teaches the systems to think and understand like humans by learning from the data, and performs tasks without explicit instructions. It is one of the most relevant technologies of the 21<sup>st</sup> century that provides systems the ability to automatically learn and improve from experience without being explicitly programmed. It opens an entirely new realm of what humans can do with computers and other machines. It describes the capacity of systems to learn from problem-specific training data to automate the process of analytical model building and solve associated tasks. It can enable an organization to autonomously learn and improve using neural networks and deep learning (DL), without being explicitly programmed, by feeding it large amounts of data. This paper tries to discuss elementary ideas of machine learning for the benefit of the new researchers in this field.

**Keywords:** machine learning, deep learning, artificial intelligence, cyber-attacks, cyber-security

## 1. Introduction

Machine learning (ML) is a branch of artificial intelligence (AI) that focuses on enabling computers and machines to imitate the way that humans learn to perform the tasks autonomously, and to improve their performance and accuracy through experience and exposure to more data (Alpaydin, 2020). It is for designing algorithms that allow a computer to learn. It aims at enabling machines to perform their jobs skillfully by using intelligent software. It is using in nearly every industry and business activity that helps the logistics industry optimize shipping and delivery routes, the retail industry personalize shopping experiences and manage inventory, manufacturers automate factories, and helps secure organizations everywhere (Fujii & Managi, 2018).

In the 1930s, American brilliant electrochemical expert Thomas Neil Ross (1909-2010) made the first attempt to develop a machine that simulated the behavior of a living creature in performance (Ross, 1938). The term machine learning was first coined in 1959 by American pioneer in the field of computer gaming and artificial intelligence Arthur Lee Samuel (1901-1990) (Samuel, 1959). In recent years ML has grown rapidly in the context of data analysis and computing that typically allows the applications to function in an intelligent manner. Actually, the ML usually provides systems with the ability to learn and enhance from experience automatically without being specifically programmed and is generally referred to as the most popular latest technologies in the fourth industrial revolution (Sarker, 2021). The ML is used in web search, drug design, spam filters, credit scoring, fraud detection, recommender systems, ad placement, stock trading, and many other applications (Domingos, 2012).

## 2. Literature Review

A literature review is an overview of previously published works on a particular topic. It provides the researchers general information of an existing knowledge of a particular topic (Bolderston, 2008). It is a comprehensive survey of scholarly sources on a specific topic that provides an overview of current knowledge, which

synthesizes, analyzes, and critically evaluates existing research to identify key themes, debates, and gaps in the literature (Galvan, 2015). A good literature review has a proper research question, a proper theoretical framework, and a chosen research methodology (Creswell, 2013). Arthur Lee Samuel has been investigated two ML procedures in some detail using the game of checkers, and has observed that a computer can be programmed so than a programmer. It can be done in a very short period of time when given only the rules of the game, a sense of direction, and a redundant and incomplete list of parameters (Samuel, 1959). Pedro Domingos has shown that ML is widely used in computer science and other fields. He has summarized some key lessons, such as pitfalls to avoid important issues to focus on, and answers to common questions (Domingos, 2012).

Tom Mitchell has described that ML as 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  $E$  (Mitchell, 1997). Niklas Kühl and his coworkers have wanted to clarify the relationship between ML and DL, and to specify the contribution of machine learning to artificial intelligence. They have presented a conceptual framework which clarifies the role of ML to build AI agents with the more terminological clarity and a starting point for interdisciplinary discussions and future research (Kühl et al., 2019). Iqbal H. Sarker has presented a comprehensive view on ML algorithms that can be applied to enhance the intelligence and the capabilities of an application through the explaining the principles of different ML techniques and their applicability in various real-world application domains, such as cyber security systems, smart cities, healthcare, e-commerce, agriculture, etc. (Sarker, 2021).

Christian Janiesch and his coauthors have summarized the fundamentals of ML and DL to generate a broader understanding of the methodical underpinning of current intelligent systems. They have provided a conceptual distinction between relevant terms and concepts, explain the process of automated analytical model building through ML and DL, and discuss the challenges that arise when implementing such intelligent systems in the field of electronic markets and networked business (Janiesch et al., 2021). Juan D Pineda-Jaramillo has provided a brief explanation of some ML algorithms commonly used for transportation research, such as artificial neural networks (ANN), decision trees (DT), support vector machines (SVM) and cluster analysis (CA). Later, the characterization of ML algorithms is discussed and random forest (RF), a variant of decision tree algorithms, is presented as the best methodology for modeling travel mode choice (Pineda-Jaramillo, 2019). Hamed Alqahtani and his coworkers have studied various popular ML classification algorithms, such as Bayesian network, naive Bayes classifier, decision tree, random decision forest, random tree, decision table, and artificial neural network to detect intrusions due to provide intelligent services in the domain of cyber-security (Alqahtani et al., 2020).

### **3. Research Methodology of the Study**

Research is a searching for knowledge and truth. It is a creative and systematic work undertaken to increase the stock of knowledge that involves the collection, organization, and analysis of evidence to increase understanding of a topic (Grover, 2015). The primary purposes of research are documentation, discovery, interpretation, the research and development (R&D) of methods, and systems for the advancement of human knowledge (Song et al., 2010). Methodology is the study of research methods that refers to the philosophical discussion of associated background assumptions. It is divided into quantitative and qualitative research areas. Quantitative research is the main methodology of the natural sciences that uses precise numerical measurements (Adams et al., 2007). On the other hand, qualitative research is more characteristic of the social sciences, such as surveys, interviews, focus groups, and the nominal group technique that aim more at an in-depth understanding of the meaning of the studied phenomena and less at universal and predictive laws (Berg, 2009). Also, many social scientists use mixed-methods research that combines quantitative and qualitative methodologies (Creswell, 2013, Mohajan, 2018b, 2020). A research methodology is a way of explaining how a researcher intends to carry out the research (Kara, 2012). It describes the techniques and procedures used to identify and analyze information regarding a specific research topic (Eyler, 2020). It provides a detailed plan that helps to keep researchers on track, making the process smooth, effective, and manageable (Mohajan, 2017; Groh, 2018).

### **4. Objective of the Study**

Machine learning (ML) is a subfield of artificial intelligence (AI) that replaces the need for developing computer programs manually and lets computers to create programs themselves from the data. It also relates broadly with many fields, such as statistics, mathematics, physics, theoretical computer science, etc. (Koza et al., 1996). It enables machines to make predictions, perform clustering, extract association rules, and make decisions from a given dataset. It studies computer algorithms for learning to do stuff, complete a task, make accurate predictions, and behave intelligently that are being done always based on some sort of data, such as examples, direct experience, and instruction. The goal of ML is to devise learning algorithms that does the learning automatically without human intervention. Hence, without codifying knowledge into computers, ML seeks to automatically learn meaningful relationships and patterns from examples and observations (Janiesch et al., 2021). Main objective of this article is to provide introductory ideas of ML. Other minor objectives of the study are as follows

(Mohajan, 2018a):

- 1) to highlight on overview and types of ML,
- 2) to focus on common algorithms of ML, and
- 3) to discuss application and importance of ML.

### 5. An Overview of ML

Machine learning and artificial intelligence, as well as the terms data mining, deep learning and statistical learning are related, often present in the same context and sometimes used interchangeably (Bousquet et al., 2011). The ML is defined as an application of AI where available information is used through algorithms for processing the statistical data. It involves concepts of automation, a high level of generalization to get a system, and requires human guidance (Mohri et al., 2012). In brief, ML is the process of turning data into programs. From a statistical point of view, ML can be regarded as an implementation of statistical learning. But, in computer science, it has the focus of designing efficient algorithms to solve problems with computational resources (Brink, 2017). The ML algorithms are organized into taxonomy, based on the desired outcome of the algorithm. Common algorithm types are supervised learning, unsupervised learning, semi-supervised learning, reinforcement learning, etc. (Anzai, 2012). The ML is used to define a group of methods or algorithms that allow computers to mechanize data driven model programming and build models by means of a methodical detection of patterns in statistically significant data (Bhavsar et al., 2017).

### 6. Types of ML

Machine learning is about designing algorithms that allow a computer to learn. There are three kinds of models used in ML: i) supervised learning, ii) unsupervised learning, and iii) reinforcement learning. I have also highlighted on the other two types, such as semi-supervised learning and transduction inference. Among these five types, supervised learning is considered as the most developed branch of ML (Alpaydin, 2020).

#### 6.1 Supervised Learning

Supervised learning is a subcategory of machine learning model (MLM) in which we teach or train the machine using data which is well labeled that means some data is already tagged with the correct answers. It uses labeled training data and a collection of training examples to infer a function. Therefore, we are given a labeled training dataset from which a machine learning algorithm can learn a model that can predict labels of unlabeled data points. It is carried out when certain goals are identified to be accomplished from a certain set of inputs. For example, given a corpus of spam and non-spam email, a supervised learning task would be to learn a model that predicts to which class, spam or non-spam, and new emails belong (Sarker, 2021).

It uses structured data to map a specific feature to a label, where the output is known accurately, and the model is trained on data of the known output. For example, names, dates, addresses, credit card numbers, stock information, geo-location, etc. are structured data. Therefore, it covers three main portions; labeled data, direct feedback, and predict outcome (Hirt et al., 2017). It indicates the presence of a supervisor or a teacher. It always aims to build a model by applying an algorithm on a set of known data points to gain insight on an unknown set of data. Therefore, it has a well-defined structure, conforms to a data model following a standard order that is highly organized and easily accessed, and used by an entity or a computer program. The supervised tasks are “classification” that separates the data, and “regression” that fits the data (Sarker et al., 2020).

#### Mathematical Notations

We define a function that use to approximate some unknown function,

$$y = f(x),$$

where  $x$  is a vector of input features associated with a training example and  $y$  is the outcome we want to predict. In classification, we define the hypothesis function as,

$$h: X \rightarrow Y$$

where  $X = R^m$  and  $Y = \{1,2,\dots,k\}$  with class labels  $k$ . In regression, the task is to learn a function,

$$h: R^m \rightarrow R .$$

Given a training set

$$D = \left\{ \langle \mathbf{x}^i, \mathbf{y}^i \rangle, i = 1, 2, \dots, m \right\}$$

where  $\mathbf{x} \in X$  and  $\mathbf{y} \in Y$ , and the training pairs  $\langle \mathbf{x}^i, \mathbf{y}^i \rangle$  are drawn from a joint distribution  $P(X, Y) = P(X)P(Y/X) = P(Y)P(X/Y)$ . Here  $\mathbf{x} \in X$  is a training example with  $m$  features, such as height, weight, age, etc. of a person, represented as a column vector,

$$\mathbf{x} = \begin{bmatrix} x_1 \\ x_2 \\ \cdot \\ \cdot \\ \cdot \\ x_m \end{bmatrix}.$$

Similarly,  $\mathbf{y} \in Y$  can be represented as a column vector,

$$\mathbf{y} = \begin{bmatrix} y_1 \\ y_2 \\ \cdot \\ \cdot \\ \cdot \\ y_n \end{bmatrix}.$$

Supervised learning can be classified into two categories of algorithms; i) classification: a classification problem is used when the output variable is a category, where the outputs are discrete labels, as in spam filtering, such as red or blue, disease or no disease, and ii) regression: a regression problem is used when the output variable is a real-valued, such as dollars or weight (Mohammed et al., 2016).

### 6.2 Unsupervised Learning

Unsupervised learning is a MLM that uses unstructured data to learn patterns without the need for human interference, where the correctness of the output is not known ahead of time. It mainly deals with the unlabeled data and tries to analyze and discover patterns within (Han et al., 2011). It is widely used for extracting generative features, identifying meaningful trends and structures, groupings in results, and exploratory purposes. The algorithm learns from the data without human input and categorizes it into groups based on attributes. It allows one to perform more complex processing tasks compared to supervised learning. Therefore, it covers three main portions; no labels, no feedback, and find hidden structure in data (Hirt et al., 2017).

Unsupervised learning is good at descriptive modeling and pattern matching. The most common unsupervised learning algorithms used present are fuzzy means, k-means clustering, hierarchical clustering, and partial least squares. It is classified into two categories of algorithms: clustering and dimension reduction (Sarker et al., 2020). It is used by clustering algorithms to find patterns in data so that it can be grouped, and is also used for dimensionality reduction, such as compressing data onto a lower-dimensional subspace or manifold. By spotting distinctions between data points that humans have missed, computers can aid data scientists. The most common unsupervised learning tasks are clustering, density estimation, feature learning, dimensionality reduction, finding association rules, anomaly detection, etc. (Mohammed et al., 2016).

### 6.3 Reinforcement Learning

Reinforcement learning is a MLM that can be described as “learn by doing” through a series of trial and error experiments. It aims at using observations gathered from the interaction with the environment to take actions that would maximize the reward or minimize the risk (Mohammed et al., 2016). An agent, such as a robot or controller learns to perform a defined task through a feedback loop until its performance is within a desirable range. Instead of having correct or false label for each step, the learner must discover or learn a behavior that maximizes the reward for a series of actions. The agent receives positive reinforcement when it performs the task well and negative reinforcement when it performs poorly (Raschka & Mirjalili, 2017). In that sense, it is not a supervised setting and somewhat related to unsupervised learning; however, reinforcement learning really is its own category of machine learning. Therefore, it covers three main portions; decision process, reward system, and learn series of actions (Hirt et al., 2017). Typical applications of reinforcement learning involve playing games, such as chess, Atari video games; and some form of robots, such as drones, warehouse robots, and self-

driving cars (Raschka & Mirjalili, 2017).

#### 6.4 Semi-Supervised Learning

The semi-supervised learning can be described as a mix between supervised and unsupervised learning, as it operates on both labeled and unlabeled data, where the algorithm must figure out how to organize and structure the data to achieve a known result (Han et al., 2011). In semi-supervised learning tasks, some training examples contain outputs, but some do not. We then use the labeled training subset to label the unlabeled portion of the training set, which we then also utilize for model training. For instance, in the ML model it is told that the result is a pear, but only some training data is labeled as a pear. It is applied in machine translation, fraud detection, labeling data, and text classification. Actually, the labeled data could be rare in several contexts, and unlabeled data are numerous, and semi-supervised learning is useful to provide a better outcome for prediction than that produced using the labeled data alone from the model (Mohammed et al., 2016).

#### 6.5 Transduction Inference

Transduction inference is reasoning from observed, specific training cases to specific cases. It is similar to supervised learning, but does not explicitly construct a function. Instead, it tries to predict new outputs based on training inputs, training outputs, and new inputs. It was introduced in a computer science context by Russian statistician, researcher, and academician Vladimir Vapnik in the 1990s (Vapnik & Kotz, 2006). Transduction algorithms can be broadly divided into two categories: those that seek to assign discrete labels to unlabeled points, and those that seek to regress continuous labels for unlabeled points (de Finetti, 1970).

### 7. Common Algorithms of ML

There are numerous ML algorithms in use, and the most common learning algorithms used at present are linear classifiers (logical regression, naïve Bayes classifier, perceptron, and support vector machine), quadratic classifiers, k-means clustering, boosting, decision tree, random forest, neural networks, and Bayesian networks. Now we will describe neural networks, k-means clustering, decision trees, and random forests in briefly (Cao, 2017).

#### 7.1 Neural Networks

Neural network has a massive number of connected processing nodes that is inspired by the structure and functions of biological neural networks, and mimics how the human brain functions (Bishop, 2006). It consists of connected units or nodes called artificial neurons, which loosely model the neurons in the brain. It is actually performing a number of regression and classification tasks at once, although commonly each network performs only one (Bishop, 1995). *American neurophysiologist and cybernetician* Warren S. McCulloch (1898-1969) and *American logician* Walter Pitts (1923-1969) introduced the concept of artificial neural network (ANN), and it was designed to simulate the functions and structure of the nervous systems in living beings (McCulloch & Pitts, 1943). The ANNs are composed of a large number of neurons that are interconnected in parallel and work in unison to solve diverse problems (Bishop, 1995).

In most of the cases, the network will have a single output variable, although in the case of many-state classification problems, this may correspond to a number of output units (Gurney, 1997). The neurons are aggregated into different layers and may perform different transformations on their inputs. Signals travel from the input layer to the output layer, possibly passing through multiple intermediate hidden layers (Dawson, 1998). Neural networks are used to extract complex patterns from the data, and perceive trends that are too complex to be observed by humans or other computer methods with their outstanding ability to derive meaning from data that is complex or inaccurate (Bishop, 1995). These are effective at identifying patterns and are crucial in applications, such as speech recognition, image creation, natural language translation, and image recognition (Bhadreshia, 1999). These are very powerful tools that have been used for numerous applications, such as medicine, transportation, optimization, and even quantum physics (Cantarella & de Luca, 2003).

#### 7.2 K-Means Clustering

Clustering problems are seen in many different applications, such as data compression and vector quantization, data mining and knowledge discovery, and pattern recognition and pattern classification (Fayyad et al., 1996). The k-means clustering is one of the simplest unsupervised learning algorithms that solve the well-known clustering problem through a certain number of clusters (assume k-clusters) fixed a priori. It does the three steps to convergence, determine the center coordinate, determine the distance of each object to the center, and group the object based on minimum distance (Bishop, 1995). The main idea is to define k centroids, one for each cluster. The next step is to take each point belonging to a given data set and associate it to the nearest centroid. When no point is pending, the first step is completed (Alsabti et al., 1998). Then re-calculate k new centroids, and if a loop has been generated, centroids move no more. The k-means is a simple algorithm that has been adapted to many problem domains. It aims at minimizing an objective function (Kanungo et al., 2004).

$$J = \sum_{j=1}^k \sum_{i=1}^n \|x_i^{(j)} - c_j\|^2$$

where  $\|x_i^{(j)} - c_j\|^2$  is a chosen distance measure between a data point  $x_i^{(j)}$  and the cluster center  $c_j$ , is an indicator of the distance of the  $n$  data points from their respective cluster centers.

### 7.3 Decision Trees

A decision tree (DT) is a flowchart-like structure and a non-parametric method that is oriented graphs formed by a finite number of nodes departing from the root nodes. It is a decision support recursive partitioning structure that uses a tree-like model of decisions and their possible consequences, including chance event outcomes, resource costs, and utility (von Winterfeldt & Ward, 1986). It simulates a real tree, which begins at a wide trunk, and as it rises is divided into narrower branches and the paths from root to leaf represent classification rules. These are powerful algorithms for classifying data, where a tree structure is used for modeling the different relationships between the features and potential output data. When a final decision is made, the DT ends with leaf nodes (Bhavsar et al., 2017).

Decision trees are simple to understand and interpret; and can determine worst, best, and expected values for different scenarios (Xu et al., 2023). These are useful for both categorizing data and regression that are the prediction of numerical values, which are simple to validate and audit. Functioning of DT is easy to understand and interpret, and needs little data preparation from the user to build an optimal DT (de Oña, 2016). Decision trees are commonly used in operations research, specifically in decision analysis, statistics, data mining, and machine learning to identify a strategy most likely to reach a goal. The measurements used to evaluate decision trees are accuracy, sensitivity, specificity, precision, miss rate, false discovery rate, and false omission rate (Kamiński et al., 2017).

### 7.4 Random Forests

Random forests (RFs) are tree-based algorithms that are associated with bagging. These are machine learning algorithms that *use many decision trees to make better predictions* (Dessi et al., 2013). These are ensemble learning method for classification, regression, and other tasks that work by creating a multitude of decision trees during training (Ye et al., 2008). In a random forest, the machine learning algorithm combines the output from various decision trees to predict a value or category (Denisko & Hoffman, 2018). The RFs combine both the different set of data called bootstrap aggregation and also numerous features selection to predict the outcome (Alqahtani et al., 2020). The first algorithm for random decision forests was created in 1995 by Chinese computer scientist Tin Kam Ho using the random subspace method (Ho, 1995). An extension of this algorithm was developed by Leo Breiman and Adele Cutler in 2001 (Breiman, 2001).

## 8. Applications of ML

The ML can be applied in almost every aspect of our life, such as web search (Bing, Google), email spam detection, game playing, weather prediction, sports predictions, stock predictions, identify credit card fraud, medical diagnoses, drug design, product recommendations, face detection and matching, ATMs, language translation, fraud detection, sentiment analysis, customer segmentation, self-driven vehicles (e.g., cars, drones), etc. (Warner & Hirschberg, 2012). It can be applied to filter spam emails. Various ML models and algorithms can be used to solve complex data science and analytics problems using Bayes' theorem (Balducci et al., 2018). In medical field, it is widely used to predict mortality in injured patients of Trauma & Injury Severity Score (TRISS) using logistic regression. It can contribute in areas as disparate as helping in the treatment of chronic diseases, fighting climate change, and anticipating cyber security threats (Fatima & Pasha, 2017). Using ML in practice requires that we make use of our own prior knowledge and experimentation to solve problems (Mohri et al., 2012).

## 9. Importance of ML

The ML helps to identify fraud, security threats, personalization and recommendations, automated customer service through the chatbots, transcription and translation, data analysis, etc. It opens an entirely new realm of what humans can do with computers and other machines (Russell & Norvig, 2015). It can control autonomous vehicles, drones, and airplanes, augmented and virtual reality, and robotics efficiently. Different ML techniques are used to meet the challenges of growing travel demands, safety concerns, energy consumption, emissions, and environmental degradation (Abduljabbar et al., 2019).

## 10. Conclusion

Machine learning is a branch of AI that aims at enabling machines to perform their jobs skillfully by using intelligent software. It is a key technology driver that encompasses the intelligent power to harness the

knowledge from the available data. It enables computers to imitate and adapt human-like behavior. In this study, I have briefly discussed various types of machine learning methods that can be used for making solutions to various real-world issues. In this article, a comprehensive review of ML process and algorithms are presented. I believe that my study on ML will be helpful to the new researchers of academia and industry professionals, and decision-makers.

## References

- Abduljabbar, R., et al. (2019). Applications of Artificial Intelligence in Transport: An Overview. *Sustainability*, 11(1), 189-190.
- Adams, J., et al. (2007). *Research Methods for Graduate Business and Social Science Students*. New Delhi: SAGE Publications.
- Alpaydin, E. (2020). *Introduction to Machine Learning*, (4<sup>th</sup> Ed.). MIT Press, Cambridge.
- Alqahtani, H. (2020). Cyber Intrusion Detection Using Machine Learning Classification Techniques. In N. Chaubey, S. Parikh, & K. Amin (Eds.), *Computing Science, Communication and Security: First International Conference*, pp. 121-131. COMS2 2020 Gujarat, India, March 26-27, 2020.
- Alsabti, K., et al. (1998). An Efficient k-Means Clustering Algorithm. In Proceedings of the First Workshop on High Performance Data Mining, Orlando, FL, March.
- Anzai, Y. (2012). *Pattern Recognition and Machine Learning*. Elsevier.
- Balducci, F., et al. (2018). Machine Learning Applications on Agricultural Datasets for Smart Farm Enhancement. *Machines*, 6(3), 38.
- Berg, B. L. (2009). *Qualitative Research Methods for the Social Sciences* (7<sup>th</sup> Ed.). Boston MA: Pearson Education Inc.
- Bhadeshia, H. K. D. H. (1999). Neural Networks in Materials Science. *ISIJ International*, 39(10), 966-979.
- Bhavsar, P., et al. (2017). Machine Learning in Transportation Data Analytics. In: Chowdhury, M., Apon, A. and Dey, K., (Eds.). *Data Analytics for Intelligent Transportation System*, pp. 283-307, Elsevier.
- Bishop, C. M. (1995). *Neural Networks for Pattern Recognition*. Oxford, England: Oxford University Press.
- Bishop, C. M. (2006). *Pattern Recognition and Machine Learning*. New York: Springer.
- Bolderston, A. (2008). Writing an Effective Literature Review. *Journal of Medical Imaging and Radiation Sciences*, 39(2), 86-92.
- Bousquet, O., et al. (2011). Advanced Lectures on Machine Learning: ML Summer Schools 2003, Canberra, Australia, February 2-14, 2003, Tübingen, Germany, August 4-16, 2003, Revised Lectures, vol. 3176. Springer.
- Breiman, L. (2001). Random Forests. *Machine Learning*, 45(1), 5-32.
- Brink, J. A. (2017). Big Data Management, Access, and Protection. *Journal of the American College of Radiology*, 14(5), 579-580.
- Cantarella, G., & de Luca, S. (2003). Modeling Transportation Mode Choice through Artificial Neural Networks. Fourth International Symposium on Uncertainty Modeling and Analysis (ISUMA), College Park, USA.
- Cao, L. (2017). Data Science: A Comprehensive Overview. *ACM Computing Surveys*, 50(3), 43.
- Creswell, J. W. (2013). *Research Design: Qualitative, Quantitative, and Mixed Method Approaches* (4<sup>th</sup> Ed.). Thousand Oaks, California: SAGE Publications.
- Dawson, C. W. (1998). An Artificial Neural Network Approach to Rainfall-Runoff Modelling. *Hydrological Sciences Journal*, 43(1), 47-66.
- de Finetti, B. (1970). *Theory of Probability: A Critical Introductory Treatment*. New York: John Wiley.
- Denisko, D., & Hoffman, M. M. (2018). Classification and Interaction in Random Forests. *Proceedings of the National Academy of Sciences*, 115(8), 1690-1692.
- de Oña, J., et al. (2016). Transit Service Quality Analysis Using Cluster Analysis and Decision Trees: A Step Forward to Personalized Marketing in Public Transportation. *Transportation*, 43(5), 725-747.
- Dessi, N., et al. (2013). *Enhancing Random Forests Performance in Microarray Data Classification*. Conference on Artificial Intelligence in Medicine in Europe.
- Domingos, P. (2012). A Few Useful Things to Know About Machine Learning. *Communications of the ACM*, 55(10), 78-87.

- Eyler, A. A. (2020). *Research Methods for Public Health*. New York: Springer Publishing Company.
- Fatima, M., & Pasha, M. (2017). Survey of Machine Learning Algorithms for Disease Diagnostic. *Journal of Intelligent Learning Systems and Applications*, 9(1), 1-16.
- Fayyad, U. M., et al. (1996). *Advances in Knowledge Discovery and Data Mining*. AAAI/MIT Press.
- Fujii, H., & Managi, S. (2018). Trends and Priority Shifts in Artificial Intelligence Technology Invention: A Global Patent Analysis. *Economic Analysis and Policy*, 58, 60-69.
- Galvan, J. L. (2015). *Writing Literature Reviews: A Guide for Students of the Social and Behavioral Sciences* (6<sup>th</sup> Ed.). Pyczak Publishing.
- Groh, A. (2018). *Research Methods in Indigenous Contexts*. New York: Springer.
- Grover, V. (2015). Research Approach: An Overview. *Golden Research Thoughts*, 4(5), 1-8.
- Gurney, K. (1997). *An Introduction to Neural Networks*. UCL Press.
- Han, J., et al. (2011). *Data Mining: Concepts and Techniques*. Amsterdam: Elsevier.
- Hirt, R. et al. (2017). An End-to-End Process Model for Supervised Machine Learning Classification: From Problem to Deployment in Information Systems. In Maedche, A., vom Brocke, J., Hevner, A. (Eds.) in *Proceedings of the Design Science Research in Information Systems and Technology (DESRIST) 2017 Research-in-Progress*, pp. 55-63, Karlsruhe, Germany.
- Ho, T. K. (1995). Random Decision Forests. Proceedings of the 3<sup>rd</sup> International Conference on Document Analysis and Recognition, Montreal, QC, 14–16 August 1995. pp. 278-282.
- Janiesch, C., et al. (2021). Machine Learning and Deep Learning. *Electronic markets*, 31(3), 685-695.
- Kamiński, B., et al. (2017). A Framework for Sensitivity Analysis of Decision Trees. *Central European Journal of Operations Research*, 26(1), 135-159.
- Kanungo, T., et al. (2004). A Local Search Approximation Algorithm for k-Means Clustering. *Computational Geometry*, 28(2-3), 89-112.
- Kara, H. (2012). *Research and Evaluation for Busy Practitioners: A Time-Saving Guide*. Bristol: The Policy Press.
- Koza, J. R., et al. (1996). Automated Design of Both the Topology and Sizing of Analog Electrical Circuits Using Genetic Programming. In Gero, J. S., and Sudweeks, F. (Eds.), *Artificial Intelligence in Design '96*. Springer, Dordrecht.
- Kühl, N., et al. (2019). *Machine Learning in Artificial Intelligence: Towards a Common Understanding*. Proceedings of the 52<sup>nd</sup> Hawaii International Conference on System Sciences, pp. 5236-5245, Grand Wailea, Maui, Hawaii.
- McCulloch, W., & Pitts, W. (1943). A Logical Calculus of the Ideas Immanent in Nervous Activity. *The Bulletin of Mathematical Biophysics*, 5(4), 115-133.
- Mitchell, T. M. (1997). *Machine Learning*. McGraw-Hill Higher Education, New York.
- Mohajan, H. K. (2017). Two Criteria for Good Measurements in Research: Validity and Reliability. *Annals of Spuru Haret University Economic Series*, 17(3), 58-82.
- Mohajan, H. K. (2018a). Aspects of Mathematical Economics, Social Choice and Game Theory. PhD Dissertation. University of Chittagong, Chittagong, Bangladesh.
- Mohajan, H. K. (2018b). Qualitative Research Methodology in Social Sciences and Related Subjects. *Journal of Economic Development, Environment and People*, 2(1), 19-46.
- Mohajan, H. K. (2020). Quantitative Research: A Successful Investigation in Natural and Social Sciences. *Journal of Economic Development, Environment and People*, 9(4), 50-79.
- Mohammed, M., et al. (2016). *Machine Learning: Algorithms and Applications*. CRC Press, Taylor & Francis Group, New York.
- Mohri, M., et al. (2012). *Foundations of Machine Learning*. MIT press.
- Pineda-Jaramillo, J. D. (2019). A Review of Machine Learning (ML) Algorithms Used for Modeling Travel Mode Choice. *DYNA*, 86(211), 32-41.
- Raschka, S., & Mirjalili, V. (2017). *Python Machine Learning - Second Edition: Machine Learning and Deep Learning with Python, Scikit-Learn, and TensorFlow*. Packt Publishing, Birmingham, Mumbai.
- Ross, T. (1938). The Synthesis of Intelligence: Its Implications. *Psychological Review*, 45(2), 185-189.

- Russell, S. J., & Norvig, P. (2015). *Artificial Intelligence: A Modern Approach*, (3<sup>rd</sup> Ed.), Pearson Publisher.
- Samuel, A. L. (1959). Some Studies in Machine Learning Using the Game of Checkers. *IBM Journal of Research and Development*, 3(3), 535-554.
- Sarker, I. H. (2021). Machine Learning: Algorithms, Real-World Applications and Research Directions. *SN Computer Science*, 2, 160.
- Sarker, I. H., et al. (2020). Cybersecurity Data Science: An Overview from Machine Learning Perspective. *Journal of Big Data*, 7(1), 1-29.
- Song, F., et al. (2010). Dissemination and Publication of Research Findings: An Updated Review of Related Biases. *Health Technology Assessment*, 14(8), 1-193.
- Vapnik, V., & Kotz, S. (2006). Estimation of Dependences Based on Empirical Data. *Journal of the Royal Statistical Society Series D*, 41(3), Publisher: Springer.
- von Winterfeldt, D., & Ward, E. (1986). *Decision Analysis and Behavioral Research*. Cambridge University Press, Cambridge.
- Warner, W., & Hirschberg, J. (2012). Detecting Hate Speech on the World Wide Web. *Proceedings of the Second Workshop on Language in Social Media*, pp. 19-26, Montréal, Canada.
- Xu, N., et al. (2023). Predicting and Assessing Wildfire Evacuation Decision-Making Using Machine Learning: Findings from the 2019 Kincade Fire. *Fire Technology*, 59(2), 793-825.
- Ye, Y., et al. (2008). Feature Weighting Random Forest for Detection of Hidden Web Search Interfaces. *Journal of Computational Linguistics and Chinese Language Processing*, 13, 387-404.

### Copyrights

Copyright for this article is retained by the author(s), with first publication rights granted to the journal.

This is an open-access article distributed under the terms and conditions of the Creative Commons Attribution license (<http://creativecommons.org/licenses/by/4.0/>).