



The Significance of Machine Learning and its Applicability in the Research Field

Sultan bin Saad Saud Al-Harbi,

Computer teacher at Gulf Colleges

Submit Date: 2023-06-11 **Accept Date:** 2024-06-02

DOI: [10.21608/ijimct.2024.80261.1020](https://doi.org/10.21608/ijimct.2024.80261.1020)

Abstract

The last decade has seen a significant number of remarkable expansions in machine learning research. The field has achieved unprecedented popularity by developing new areas and increased momentum on the existing sites. Whereas the symbolic methods have been exceedingly active, the sub-symbolic approaches, especially the neural networks, have had tremendous growth. Unlike the efforts whose main concentrations were on the single learning strategies, the new trends have been to integrate various methods and for the development of cognitive learning architectures. Besides, there is an increased interest in the experimental comparisons of several methods and learning algorithms' theoretical analysis. The same data sets have often been shared amongst the researchers who have led to applying the same techniques in dealing with the same problems to understand the relative merits of various methods. There are new insights that have been brought into the complexity of the learning process through theoretical investigations. This research paper seeks to transmit the knowledge of machine learning, which has become very popular and the researchers' applications of the experience.

KEYWORDS: Algorithm, Machine Learning, Supervised learning, unsupervised learning,

Machine Learning

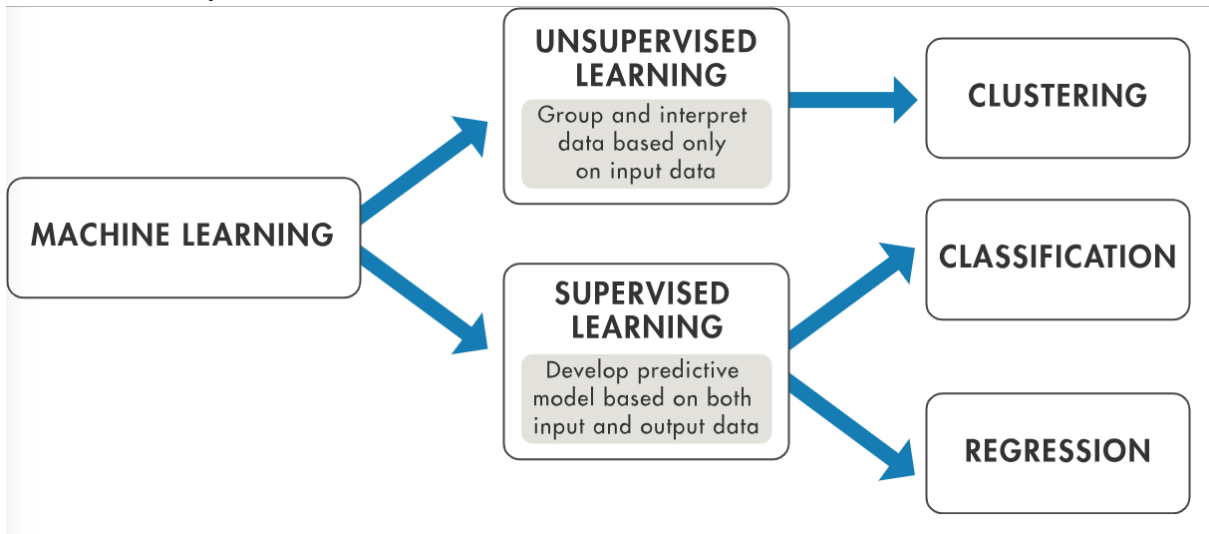
Introduction

Learning can be defined as a process of enhancement and change in several behaviors by exploring new information in time. When such learning is conducted using machines, it is termed as machine learning. The poor performance results that have been produced through statistical estimates have been in the estimation area for over one decade. Their inability to handle lack of reasoning capabilities, the spread of data points, cope with missing data points, and categorical data has led to an increase in the number of studies that use non-traditional means like machine learning technique. Machine learning can be defined as studying situational methods of improving performance through the mechanization of knowledge acquisitions from gained experience. Machine learning seeks to provide increasing levels of automation in understanding the engineering process. Machine learning applications can be applied in various fields such as biology, calculated finance, generation of energy, production, aviation, automotive, speech and handwriting recognition, image processing and computer vision, and natural language processing. Since human beings cannot analyze and interpret data, methods and algorithms in machine learning have been developed for this purpose. This research paper seeks to establish the significance of machine learning and its applicability in the research field through various techniques and implementation designs.

Literature Review

In recent years, researchers in the adversarial data retrieval community have been moving towards machine learning. Most modern machine learning research are devoted to improving the accuracy of prediction (Dey, 2016). However, there is less attention paid to the machine deployment and deep learning systems, time-series similarities, irregular temporal data analysis, and supervised/unsupervised techniques. Most of these deployments are found in the cloud with scalable and abundant resources, and a free choice of the computation platform (Mishra et al., 2018). Machine learning techniques are essential in enabling the systems to learn from experiences. It is a broad discipline which has been used by modern researchers to improve the accuracy of predictions. However, there is little attention paid for the deployment of deep and machine

learning systems, unsupervised and supervised techniques that are essential for healthcare data mining, and time series irregular temporal data analysis and similarity.



Jordan & Mitchell (2015) proposed a feature transformation technique that uses a proposed Gaussian distance function to achieve dimensionality reduction to represent the original input dataset in the new transformation space. There is a further proposal of new computation expressions used to determine a threshold and equivalent deviation in the Gaussian area. In a research conducted by Goodfellow, Bengio, & Courville (2016) to determine the new framework that can be used to detect epileptic seizures from epileptic patients and EEG signals that should be recorded from the regular patients. The research proposed an approach designed for the classification of the abdominal call from the normal one automatically.

In education, Machine learning techniques are used in intelligent instructors to acquire specific knowledge about the learners' learning of new teaching approaches. There is no particular recipe that is used to ensure there is a successful implementation of the machine learning algorithms. This means that the techniques cannot be fully covered (Singh, Thakur, & Sharma, 2016). To understand how machine learning algorithms work, it is essential to implement one as a training exercise to reach the objectives. It is also possible for one to understand and internalize the algorithm and the coupled mathematics descriptions since they will be about the matrices and vectors as computational

intuitions and arrays for the transformations of the variable structures (Mishra et al., 2018). To implement machine learning algorithms, they should work through a wide range of micro-decisions that often lack formal algorithm descriptions. If they manage to parameterize and learn such decisions, they are likely to find themselves at an advanced or intermediate level of operating the machine learning process.

There are various procedures and patterns that one must follow to enhance their abilities to implement the machine learning algorithm from scratch. The more one makes the implementations, the more they are likely to learn faster, and they are likely to become efficient and customized in working with these algorithms. Five necessary steps are used during the implementation process (Dey, 2016). They include language, type of algorithm, problem, research, and testing. If one seeks to make the process accelerate along the path, they need to port the algorithms from the first language to another. There are many open-source implementations that one can re-implement and code review to other languages. Besides, it is also possible to open-source one's code when developing and after the development, adding the instructions and substantial comments on the uses and the process of building it. After doing this, one can market their skills, or inspire the beginners or meet fellow machine learning programmers who have interests in performing code reviews and audit for them (Goodfellow, Bengio, & Courville, 2016). Furthermore, when one decides to implement the algorithms by hand, they are more likely to develop valuable skills. Some of these skills are essential in mastering algorithms while others are used in the production system development and other for the classical research in the domain.

When working with machine learning, the traditional functions of the design; communicating with the stakeholders and crafting the product vision, apply, but machine learning is also likely to develop- other factors to the table. Besides, there are different design techniques, and they are also implemented differently in machine learning development (Mishra et al., 2018). At the base, the whole design is all about receiving the data by collecting a lot of it, making sure it is clean, ensuring you understand, and that you ultimately develop software on top of it. The process entails data collection, cleaning the data and visualizing, creating algorithms and models, and evaluating the algorithms and models.

Methodology

This research used the survey method for data collection. The researcher used virtual questionnaires which had a mixture of both open and close-ended questions. The targeted participants were senior employees in the department of computer and information technology for various organizations. The employees were asked to participate in their will as the researcher was strict on the ethical part. The respondents were also asked to answer only the questions they were comfortable with to see that they provide the most accurate information for the research.

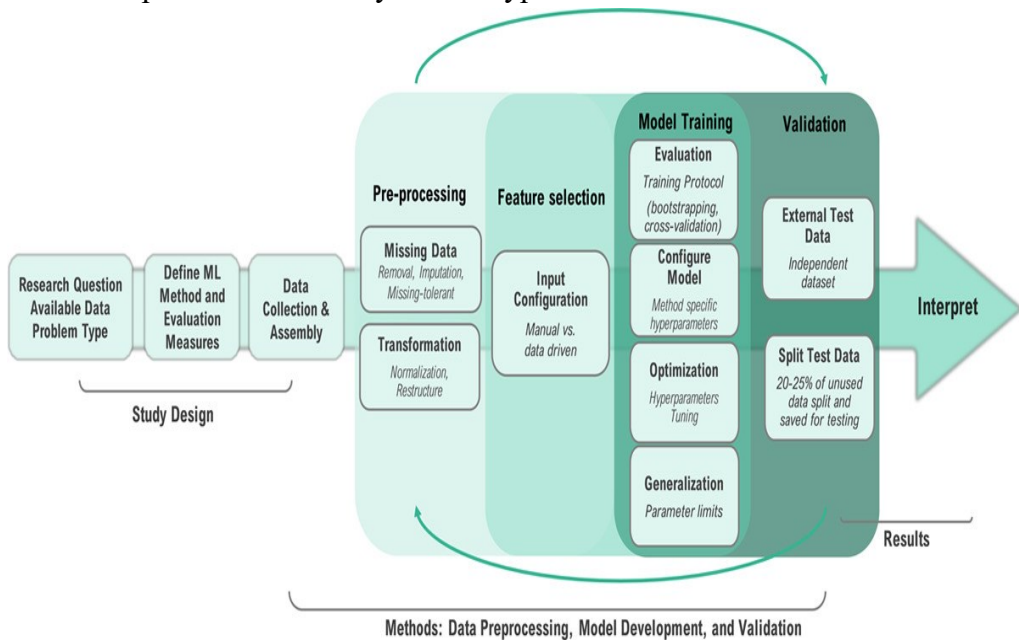
Findings

The research found that various techniques are applied by many organizations that are essential for their productivity. Various respondents had different designs and implementation techniques. Most of the respondents indicated that they used the modern methods of supervised machine learning. They said that the supervised machine learning is dependable as it turns data into real, actionable insights. It helps their organizations to utilize the data to understand and prevent the unwanted outcomes and at the same time, boost the desired results for the targeted variables (Singh, Thakur, & Sharma, 2016). The respondents who were using the supervised method of machine learning cited that the unsupervised one is rarely dependable. It is hard for one to get the precise information on the data sorting and the output as the user data is only labelled and is unknown. Besides, they also said that it is less accurate as the input data is not known. However, those who were pro unsupervised method indicated that it has the ability to handle problems that are impossible as a result of the bias and limited capacity. They said that the supervised practice is challenging as the classification of big data is stressing.

Discussion

Whereas the traditional statistical methods might suffice and need to be used, the machine learning methods must be used to offer marked advantages over the statistical methods. In various cases, this will be apparent in the ECG training, acoustic waveform, and interpretation of image which can hardly be interpreted through traditional regression means without engineering, or some feature regression (Dey, 2016). For clinical research, the machine learning rationale might be less obvious, such as improving substandard predictive

models or exploring complex data in a hypothesis-free manner. The respondents must understand the significance of the techniques selected before responding to the research questions (Mohammed, Khan, & Bashier, 2016). Machine learning method election depends on the research question's general objective, the data available, and the problem type (Caixinha & Nunes, 2017). For instance, with the supervised problems such as the prediction of clinical results and treatment response, which contain the variables with mixed data types, and method in machine learning such as regression tree and classification might be a superior option due to its ability to identify high imensional, non-linear relationship between a variety of data types.



The data regarding supervised machine learning indicates that it is essential to describe the prediction labels for applications such as detecting disease cases or clinical outcomes (Mohammed, Khan, & Bashier, 2016). The features of the candidates need to be identified from the data source. It is vital to address the limitations of the data used and the defined phenotypes for use as candidate features and outcomes (Jordan & Mitchell, 2015). On the other hand, unsupervised machine learning is often used to identify previously unknown structures in data or patterns. The way the data is interpreted depends on the data type and domain. Since there are no pre-specified hypotheses, the intuition behind the intended use and analysis of the findings must be described.

Conclusion

The complexity and size of the data are expanding rapidly, which calls for the use of machine learning in the analysis process. To help in the broader understanding and reproducibility and use of machine learning within the research community, the machine learning methods and analysis results must be presented in a relatable, standardized and concise manner to help facilitate the assessment and peer review through general readership. This study indicates that most of the essential concepts for reporting conventional statistics are also critical in machine learning with a few significant additions. Access to the analysis code and software details, ideally in a functional, publicly available workplace is a vital companion to a manuscript to promote reproducibility and ensure transparency. The research indicated that these principles are essential for the study to evaluate the valuable analyses appropriately as understanding machine learning becomes more widespread.

References

- Caixinha, M., & Nunes, S. (2017). Machine learning techniques in clinical vision sciences. *Current eye research*, 42(1), 1-15.
- Dey, A. (2016). Machine learning algorithms: a review. *International Journal of Computer Science and Information Technologies*, 7(3), 1174-1179.
- Goodfellow, I., Bengio, Y., & Courville, A. (2016). Machine learning basics. *Deep learning*, 1, 98-164.
- Jordan, M. I., & Mitchell, T. M. (2015). Machine learning: Trends, perspectives, and prospects. *Science*, 349(6245), 255-260.
- Mishra, P., Varadharajan, V., Tupakula, U., & Pilli, E. S. (2018). A detailed investigation and analysis of using machine learning techniques for intrusion detection. *IEEE Communications Surveys & Tutorials*, 21(1), 686-728.
- Mohammed, M., Khan, M. B., & Bashier, E. B. M. (2016). *Machine learning: algorithms and applications*. Crc Press.
- Singh, A., Thakur, N., & Sharma, A. (2016). A review of supervised machine learning algorithms. In *2016 3rd International Conference on Computing for Sustainable Global Development (INDIACom)* (pp. 1310-1315). Ieee.