International Journal of Scientific Research in Engineering and Management (IJSREM)
Volume: 08 Issue: 09 | Sept - 2024 SJIF Rating: 8.448 ISSN: 2582-3930

Introduction to Machine Learning (ML) its Types, Applications and Challenges: A Review

Mrs. Anush Sharma

Assistant Professor, Dept. of CSE

HIET GROUP OF INSTITUTIONS, Shahpur, HP, India

Shashi Bala

Student of CSE

HIET GROUP OF INSTITUTIONS, Shahpur, HP, India

ABSTRACT

Machine Learning (ML) is a branch of Artificial Intelligence (AI). ML enables the system to learn and improve the performance of system from experience. It has revolutionized many fields by enabling computers to learn from data without explicit programming. The are four main types of ML: supervised machine learning, unsupervised machine learning, semi-supervised machine learning and reinforcement machine learning.

ML offers multiple benefits across various industries. One of the major advantages is its ability to process and analyse vast amount of data. ML has found applications from diverse sectors. The applications of machine learning include fraud detection, autonomous vehicles, traffic prediction, language translation, disease diagnosis, drug discovery etc.

Although ML has various applications but it faces several challenges also. The quantity and quality of data that is given to the ML model affect the performance of the model, as ML models are only as good as the data was given to the model and they are trained on. Data may be incomplete, inconsistent or skewed. Incomplete and biased datasets can lead to the skewed results unreliable models. Another significant challenge in ML is the "black box" nature of many models. As the complexity raises in model, they become increasingly difficult to interpret and explain. The lack of interpretability can be problematic in decision making processes such as disease diagnosis, criminal justice etc. Ensuring the fairness and removing the bias in ML models is another complex task that goes beyond technical solutions. This requires careful consideration of social and legal factors.

KEYWORDS: Machine Learning, unsupervised, inconsistent, complexity.

1. INTRODUCTION

We live in a world where data is all around us. Every object in our environment, including cell phones, social media, voice and face recognition, self-driving automobiles, genomic sequencing, computer games language interpretation etc. all of our daily activities are done online. Data extraction may be used to produce a variety of intelligent applications in various fields such as science, medicine, data governance, manufacturing, financial modelling, education, cybersecurity, marketing etc. Thus, information management technologies that are able to extract useful information from data are needed. AI and ML, particularly ML have progressed and developed essential tools for intelligently analyse such data and create the accompanying real-world applications. For example, ML has become the preferred approach for creating useful software for language processing, speech recognition and computer vision. ML has also had a significant impact on a variety of sectors that deals with data-intensive problems such as supply chain management, consumer service, diagnosis, diagnosis of complex system failures (Pugliese et al., 2021).

In the modern world, a vast amount of data is created and kept in data repositories all over the world due to the falling cost of data storage, and quick growth in data processing speed and the growing integration of internet. Furthermore, data storage is getting cheaper, which implies it will soon become more accessible. The availability of this enormous dataset has allowed for the growth of a new area of study known as "Machine Learning (ML)". For example, when we watch a movie or web series, a recommended system will provide recommendations for movies that we are watching. ML techniques have been found to be beneficial in the fields of cybersecurity also (Tufail et al., 2023).

People have used a variety of tools to complete different jobs more quickly before ML. Different machines have been invented as a result of human innovation. By enabling individuals to satisfy a variety of requirements, such as computing, travel and industry, these technologies made life easier for humans. Among these, ML is the one. Arthur Samuel defines ML as the branch of research that enables computers to learn without the need of explicit programming. ML trains machines to handle data more effectively. Numerous programmers and mathematicians use a variety of techniques to solve problems with large amount of data (Mahesh, 2020).

2. TYPES OF MACHINE LEARNING: -

There are mainly four types of machine learning. These are: supervised machine learning, unsupervised machine learning, semi supervised machine learning and reinforcement machine learning. When learning under supervision, the data has characteristic referred to as a target variable and an independent variable. Unsupervised learning, on the other hand, lacks a goal variable. When goal labels are present in some training data, this is known as semi supervised learning. In reinforcement learning data with less labels are introduced into the model. But we give out the rewards and punishments on the basis of the output of algorithm.

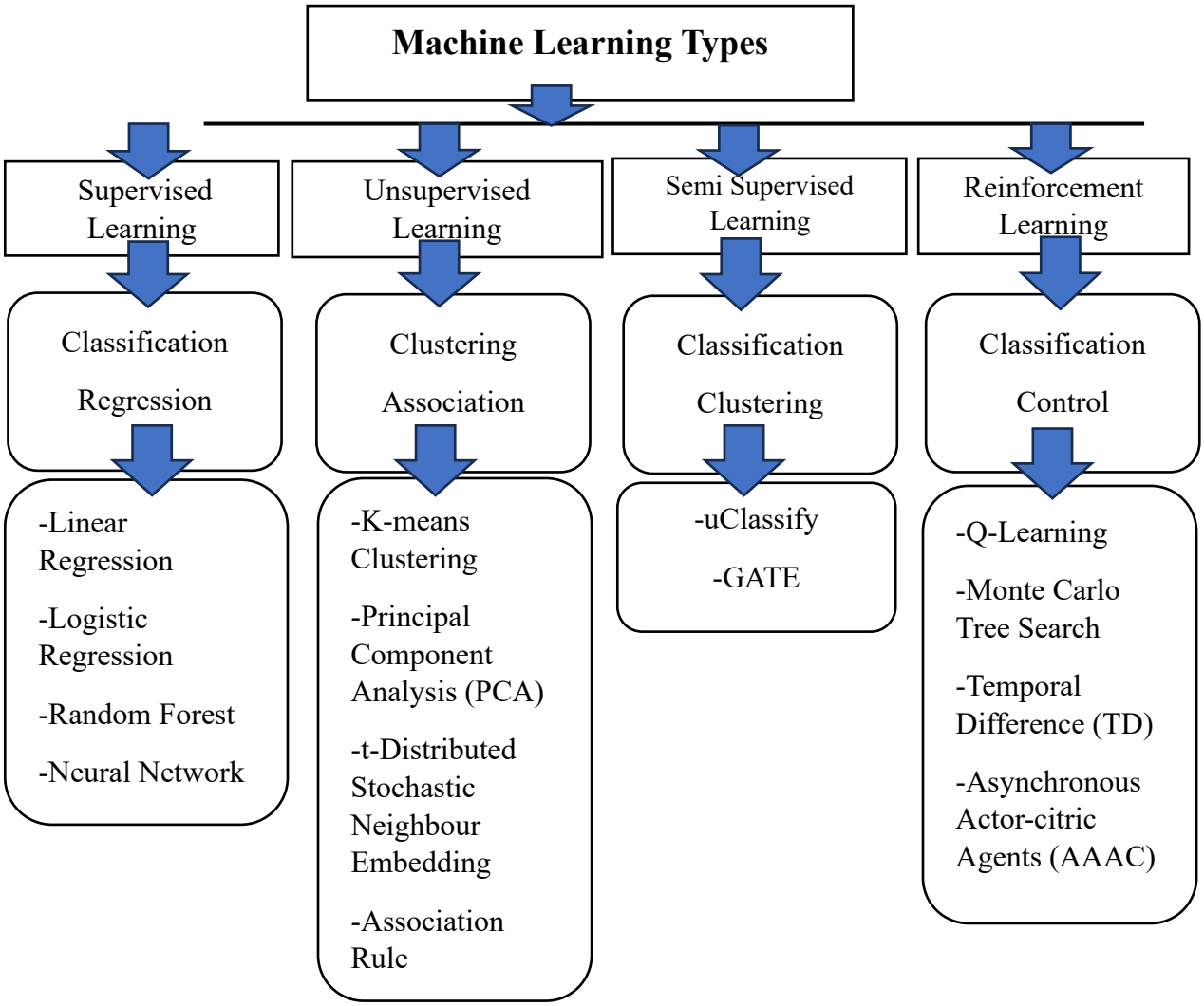

Figure 1: Different types of ML and their algorithms (Taye, 2023)

2.1. SUPERVISED MACHINE LEARNING:

The data for the supervised machine learning system is made up of inputs combined with target variables or labels, during the training phase of the system. The algorithm will search for and take advantage of patterns in the data during the training phase that are connected to the desired outputs. Following training, when new inputs are introduced, a supervised learning algorithm determines which label using the available training data, the new inputs should be placed under. When a supervised machine learning model is given new inputs, its goal is to predict the correct label in a classification problem or the estimated value of the output in a regression issue (Tufail et al., 2023).

2.2. UNSUPERVISED MACHINE LEARNING:

Unsupervised machine learning examines datasets without the need for human intervention. The method ideally divides the samples into several classes in unsupervised learning based on the characteristics of the training data only, devoid of matching labels. Typically, unsupervised machine learning is demonstrated by the automated friend recognition for a user on social networking platform like Facebook, Google or figuring out how many mails sent to a particular recipient and arranged into collective groups (Pugliese et al., 2021).

2.3. SEMI SUPERVISED MACHINE LEARNING:

The combination of supervised and unsupervised machine learning results in semi supervised machine learning. It can be useful in data mining and machine learning applications where unlabelled data is already available and obtaining labelled data is time consuming. In more popular supervised machine learning techniques, an algorithm is "labelled", meaning that each record contains information about the outcome (Mahesh, 2020).

2.4. REINFORCEMENT MACHINE LEARNING:

Reinforcement machine learning is another type of ML. It allows machines and software agents to automatically choose the best course of action in a given situation or environment to enhance its efficiency. This kind of learning is focused on rewards and penalty and its major objective is to apply the knowledge from environmental activities to take the actions to raise the rewards or to reduce the risk. It is effective tool for developing AI models that might aid in boosting automation or maximize the proficient functioning of advanced robotics tasks involving autonomous driving, supply chain logistics, and manufacturing. However, it is not recommended to solve the basic or straightforward situations (Sarker, 2021).

Table1: Types of ML algorithms and real-world application examples (Pugliese et al., 2021)

S. No.	Machine learning type	Description	Model building	Examples
1.	Supervised learning	Data is labelled with classes and outcomes (task-driven approach)	Model learns from labelled data. This type of learning comprises classification, regression, naïve bayes classification, random forests and neural networks.	E-mail data (i.e. automatic answering of incoming messages, mail organization into folders, spam filtering), face and speech recognition, information retrieval, data centre optimization.

2.	Unsupervised learning	Data is not labelled. The algorithm optimally separates the samples into different classes on the basis of features of the training data alone, without corresponding labels (data-driven approach)	Models learn from unlabelled data. This type of learning comprises k-means clustering, hierarchical clustering and principal component analysis (PCA).	Organising large computer clusters, social network analysis, DNA classification, computational biology and market segmentation.
3.	Semi supervised learning	The algorithms work with the labelled and unlabelled data.	Models are built using combined labelled and unlabelled data. This type of learning includes both classification and clustering.	Text document classifier, text filtering, sematic scene classification, fraud detection.
4.	Reinforcement learning	The algorithm operates sequentially to automatically evaluate the optimal behaviour in a particular context or environment to improve the efficiency (environment-driven approach)	Models are based on reward or penalty. This type of learning uses classification.	Traffic forecasting service, computer games, machinery applications, autonomous driving tasks, medicine, surgery.

3. APPLICATIONS OF ML:(Sarker, 2021)

3.1. Predictive analysis and decision making: ML enables intelligent decision-making through analytics across various industries, using algorithms. It also plays crucial role in cybersecurity and employing techniques like clustering, classification etc.

3.2. Internet of Things (IoT): IoT transform everyday objects into smart devices, with smart cities as a key application area enhancing urban services. ML enable predictive modelling for tasks like traffic management and energy usage optimization in smart cities.

3.3. E-Commerce: ML enables e-commerce companies to analyse consumer behaviour and make personalized product recommendations, optimizing inventory management and enhancing customer experience. This technology helps business retain customers and attract new ones by providing tailored content and offers.

3.4. Traffic prediction: ML enables accurate traffic prediction for intelligent transportation systems in smart cities. These models optimize routes, reduce congestion and improve overall transportation efficiency and sustainability.

3.5. Speech, pattern and image recognition: These are the key applications of ML, enabling automated identification in face, voice and pattern analysis and recognition. These techniques use classification, clustering and other methods to interpret complex data for various real-world applications.

3.6. Healthcare: ML aids healthcare, particularly in combating COVID-19, trough disease prediction outbreak forecasting. It enables informed clinical decisions and pandemic management via risk assessment and medical image processing.

4. ISSUES AND CHALLENGES IN ML:(Issues in Machine Learning, n.d.)

4.1. Inadequate and poor quality of data: inadequate and poor quality of training data can severely impact the ML model performance, leading to inaccurate predictions. Ensuring high quality, clean and diverse datasets is essential for effective ML outcomes.

4.2. Monitoring and Maintenance: regular monitoring and maintenance are crucial to keep ML models effective, as changes in data or user expectations may require adjustments. Continuous vigilance ensures models stay relevant and accurate.

4.3. Overfitting and Underfitting: Overfitting happens when a model learns noise from the dataset, impacting its performance, while underfitting occurs when a model is too simple to capture underlying patterns. Solutions include adjusting model complexity, refining features or modifying constrains.

4.4. Lack of skilled resources: The ML industry faces challenges dues to shortage of skilled professionals, requiring investment in training and education to bridge the gap.

4.5. Data bias: It can lead to errors by giving disproportionate weight to certain elements in a dataset. Detecting and mitigating this bias necessitates thorough examination, regular analysis, and strategies to promote data diversity.

5. Discussion

Machine Learning (ML) has emerged as a transformative technology in the field of Artificial Intelligence (AI). ML revolutionizing various industries and offer innovative solutions to complex problems. This study outlines the four main types of ML: supervised learning, which uses labelled data for tasks like classification and regression; unsupervised learning, which discovers patterns in unlabelled data; semi supervised learning, a hybrid of supervised and unsupervised learning; reinforcement learning. The diversity in these approaches demonstrates the flexibility of ML in adapting various problem and data types. The applications of ML ranging across various sectors such as healthcare, transportation, finance and technology.

However, ML has numerous benefits but it faces several challenges also. The performance of ML model depends on the quality and quantity of training data. The incomplete or biased datasets leads to unreliable results. As model become more complex, they often become less interpretable. To ensure fairness and removing bias in ML model is another complex task that requires careful consideration of social, ethical and legal factors. Moreover, while ML models can excel at specific tasks, they often struggle to generalize their knowledge to new and unseen situations. As the fields of AI and ML continue to evolve, several areas are likely to see significant development.

6. CONCLUSION

Machine Learning (ML) has become an innovative technology in many scientific and industrial sectors. ML has opened up new possibilities for data analysis, prediction and decision-making processes. Different types of machine learning have unique approaches to solving complex problems and extracting valuable insight from data.

The applications of ML are vast and diverse, ranging from fraud detection and autonomous vehicles to language translation and disease diagnosis. Its ability to analyse and process enormous amount of data has made it an invaluable

tool in fields such as cybersecurity, finance, healthcare, and many other sectors. The predictive capacity of ML models offers powerful benefits, allowing organizations to anticipate future trends and behaviours based on historical data.

From above discussion we can conclude that ML represents a powerful set of tools and techniques that are reshaping how we approach complex problems across numerous domains. While challenges remain, the benefits of ML are immense. As the research progress and new applications emerge, ML plays an important role in shaping our technological future and driving innovation across industries.

7. REFERENCES

- Mahesh, B. (2020). Machine Learning Algorithms - A Review. *International Journal of Science and Research (IJSR)*, *9*(1), 381–386. https://doi.org/10.21275/art20203995
- Pugliese, R., Regondi, S., & Marini, R. (2021). Machine learning-based approach: Global trends, research directions, and regulatory standpoints. In *Data Science and Management* (Vol. 4, pp. 19–29). KeAi Communications Co. https://doi.org/10.1016/j.dsm.2021.12.002
- Sarker, I. H. (2021). Machine Learning: Algorithms, Real-World Applications and Research Directions. In *SN Computer Science* (Vol. 2, Issue 3). Springer. https://doi.org/10.1007/s42979-021-00592-x
- Taye, M. M. (2023). Understanding of Machine Learning with Deep Learning: Architectures, Workflow, Applications and Future Directions. In *Computers* (Vol. 12, Issue 5). MDPI. https://doi.org/10.3390/computers12050091
- Tufail, S., Riggs, H., Tariq, M., & Sarwat, A. I. (2023). Advancements and Challenges in Machine Learning: A Comprehensive Review of Models, Libraries, Applications, and Algorithms. In *Electronics (Switzerland)* (Vol. 12, Issue 8). MDPI. https://doi.org/10.3390/electronics12081789
- *Issues in Machine Learning*. (n.d.). https://iabac.org/blog/issues-in-machine-learning