Machine Learning and its Applications- A Review Study

Dr. Chhinder Kaur[1], Richa Chandel[2], Dr. Tejinder Pal Singh Brar[3], Shikha Sharma[4]
[1,2,3,4]Department of Computer Applications, Chandigarh Group of Colleges, Landran, Mohali

Abstract: **In today's age, machine learning has become a widely-used technological tool. Out of all the aspects of artificial intelligence, machine learning has received the most attention and has led to significant advancements in digitalization solutions as the world prepares for augmented and virtual reality technology to become more widespread. Researchers are actively working towards making machines more intelligent and sophisticated, taking inspiration from the innate intelligence of humans. Artificial intelligence (AI) refers to the use of computers and other electronic devices to perform tasks that usually require human intelligence. A key area of focus in machine learning research and development is to improve accuracy, which involves analyzing data to gain valuable insights and creating algorithms based on this analysis. From healthcare to finance, transportation to agriculture, and beyond, the possibilities are vast and continually expanding. It will be exciting to see how machine learning continues to shape and transform various fields in the coming years. Its application involves analyzing and comprehending complex data sets, and making predictions or decisions based on the analysis. This article aims to provide a high-level overview of machine learning, its subfields, and some of the ways in which this technology is being used to enhance knowledge and improve work performance**

.

Keywords: Digitalization, Machine Learning, Artificial Intelligence, Deep Learning, Bioinformatics

I. INTRODUCTION

Artificial intelligence (AI) is the use of computers and other electronic devices to complete tasks that would typically require human intelligence. In the language of computer science, AI refers to the study of "intelligent agents," which are devices capable of observing their surroundings and taking actions to maximize their chances of successfully accomplishing a specific set of tasks. Informal definitions of artificial intelligence include the "ability to learn and solve problems" and the "capacity to understand, similar to how human brains learn from other human brains." The field of machine learning has experienced numerous breakthroughs since the 1950s, followed by a period of disillusionment known as "AI winter." However, recent massive efforts have led to new methods and successes in machine learning, resulting in a rebound in the field (Shinde, 2018).

It is important to learn the fundamentals of Machine Learning (ML) at this time. The term "machine learning" was created to describe the field of algorithms and statistical models that allow computers to perform tasks without explicit programming, instead relying on patterns and instructions. ML is a subset of artificial intelligence that grants systems a range of perceptual abilities and the capability to learn and adapt over time. In terms of task execution, Machine Learning primarily relies on various computer algorithms that allow machines to acquire available data, use it for learning, and gain more expertise over time. With this information, machines can make judgments and predictions.

Imagine a computer system that has been trained to analyze patient records for the purpose of detecting or predicting cancer. As the system gains experience by accurately analyzing medical records from a larger patient population, its performance will naturally improve.The accuracy of the algorithm can be measured by the percentage of correct predictions and detections, which can be verified by an expert oncologist.

Figure 1 provides a visual representation of the connections between Artificial Intelligence, Machine Learning, and Deep Learning.AI methods involve training algorithms on data to recognize patterns and make predictions, using techniques like machine learning, deep learning, and reinforcement learning. These methods enable computers to perform tasks that typically require human intelligence.

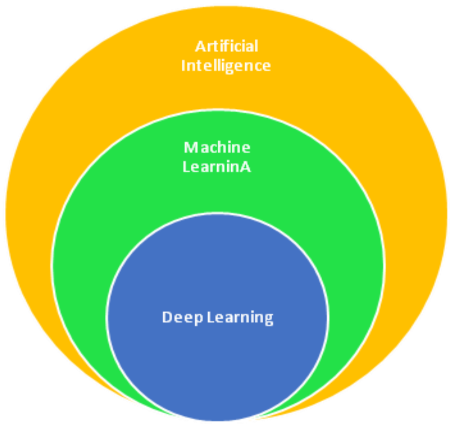

Figure 1: Artificial intelligence methods

II. LITERATURE REVIEW

In 2006, Hua conducted a study that utilized neural networks to forecast Hong Kong hotel occupancy rates. The study showed that the neural networks outperformed both the naive extrapolation model and multiple regression. This study demonstrated the potential of neural networks for predicting hotel occupancy rates. Additionally, in 2011, Hong described machine learning methods using GA-SVR for real-valued gas molecules. The experimental results showed that SVR models outperformed both ARIMA and BPNN models. To predict future tourism demand, Zarandi (2012) recommended using Chaotic Genetic Algorithms (CGAs) such as SVRCGA, which are capable of resolving the premature local optimum problem. Based on the findings, the SVRCGA model was found to be superior to its competitors.

The direction of stock prices can be predicted using a Fuzzy Multi-Agent System with four layers (FMAS), as proposed by (Shahrabi, 2013). The authors found that FMAS was a useful method for their research, achieved by cooperation of a group of smart bots. Similarly, (Pai, 2014) presented a Modular Genetic-Fuzzy Forecasting System using a genetic fuzzy expert system for estimating tourist demand. This technique is superior to other methods for estimating tourism demand, as concluded by the authors due to its high level of accuracy in predictive power. Moreover, (Zhang, 2018) compared traditional approaches to statistical time series forecasting with Machine Learning techniques and found that the latter was more accurate.

III. WORKING AND LEVELS

Different types of Machine Learning algorithms can be defined in terms of the following scenarios involving the availability of training data, test data, and instructional strategies.

A. *Supervised Learning*

One of the most common learning strategies supervised learning takes as input a data set with the known outputs for each associated input. As can be seen in Figure 2, the Machine Learning model makes use of them in an effort to construct a connection between the input and output. Again, the two main categories of supervised learning techniques are "regression" and "classification" issues. The machine learning model translates the continuous output function against the input variables in the case of a regression problem. The model must, for instance, determine an individual's age from a single photograph. Classification problems, on the other hand, have the machine learning model attempt to classify the input variables into distinct groups. To give just one example, one may utilize the local median income to predict home values.

B. *Unsupervised Learning*

When compared to other learning algorithms, unsupervised learning emerges victorious since it is used to tackle issues for which the outcomes and effects of variables are unknown or only partially known. As can be seen in Figure 3, the structures can be deduced by clustering the provided data according to the relationship among the variables of data. Without any prior knowledge of the training data, the primary goal of this machine-learning model is to create a cluster of unsorted information based on similarities, patterns, and differences. The machine is only able to uncover the unlabeled data's underlying structure on its own. Two subcategories of Unsupervised Learning tasks are Clustering and Association. (Foxon,2021). Data points in a clustering problem are grouped according to the correlations between the variables, such as clients being categorized based on how they purchase goods. It's a common technique for analyzing data statistically in many fields. To solve an association problem, we look for commonalities between items in the dataset, such as assuming that customers of a certain store also purchased a certain other store's wares.

C. *Semi-supervised Learning*

Algorithms for semi-supervised learning draw from both supervised and unsupervised approaches. In areas of machine learning where we already have unlabeled data and were collecting the labeled data from this is a laborious procedure (Zhu, 2009), this can be a huge help. Semi-supervised learning problems involve a large amount of unlabeled input data alongside a smaller amount of labeled data. Example: we have a series of photographs from a photo album, but only some of them contain titles (such as "elephant," "car," and "bridge").

D. *Reinforcement Learning*

Machine learning can also take the form of reinforcement learning, in which a "reinforcement agent" actively seeks to increase the frequency with which the desired outcome is achieved. At first, the trainee has no idea what to do while awaiting any given circumstance. Though the learner's actions may not directly alter the environment, they may have long-term consequences. After the learning algorithm chooses an output for a given input, it receives feedback from the environment that indicates how well the result meets the learner's objectives. The reinforcement learning algorithm is useful for sequential tasks, in which the learner interacts with an environment by selecting appropriate actions (the outputs) in response to specific observations (the inputs). Two things—a search strategy based on trial and error, and a time lag between actions—are important to the success of a reinforcement learning algorithm (Sutton, 1992).In Figure 4, the agent is shown taking in some form of input I being in a state (s), undergoing a change (r), and being influenced (a) by a function of its surrounding environment. The agent uses these inputs to infer a behavior, designated by B, and then uses B to guide the action that produces the result.

E. *Multitask Learning*

The goal of Multi-Task Learning is to facilitate the development and progress of other students. When this algorithm is trained on a specific task, it recalls the steps required to solve the problem and the learning algorithm's response to those steps in order to draw the appropriate conclusion. Following this, the algorithm employs the same strategies to solve more problems with a similar structure. Sharing one's own knowledge and experiences with other students can help the entire class learn more quickly and efficiently, which is a key part of any learning algorithm (Caruana, 1997).

F. *Ensemble Learning*

In Ensemble Learning, many separate learners are combined into one. This learner could be a Naive Bayes, a neural network, or a decision tree, among other options. Since the 1990s, ensemble learning has been increasingly popular. As opposed to a single learner, a group of students is always preferable when completing a task (Opitz, 1999).

G. *Neural Network Learning*

Neurons, the cell-like structures found in our brains, serve as the biological inspiration for the Neural Network. Understanding how neurons work is essential for decoding neural networks. As shown in Figure 5, a neuron is composed of four main parts: the soma, the nucleus, the dendrites, and the axon. The Dendrites are responsible for picking up electrical impulses and relaying them to the Soma, where they can be processed. The result is then floated along the axon to the terminals of the dendrites, where it is picked up by the next neuron in the circuit.

Neural networks are interconnected networks of neurons that allow for the transmission of electrical signals throughout the brain. A similar three-layer structure is also included in an artificial neural network, as seen in Figure 6. The input layer is analogous to dendrites, the hidden layer, processes the input in a way similar to the soma or axon, and the output layer, which is analogous to axons in transporting the result of the processing (, S. B.,2014).Supervised networks, unsupervised networks, and reinforcement networks are the three most common varieties of the artificial neural network (Kumar, Y., 2020).

H. Instance-Based Learning

The learner is sufficiently primed for it to be able to acquire specific patterns, which it then attempts to exercise on the data that is provided to it. Therefore, it bears this moniker. This type of learner is so sluggish that it waits for the test set to arrive before analyzing it with the training data. One drawback is that as data sizes grow, so does their complexity.

IV. MACHINE LEARNING ALGORITHMS

The core of Machine Learning is found in its algorithms. They are crucial to the overall operation of the system and, with sufficient input data, are capable of making accurate predictions in a short amount of time.

a. Supervised Machine Learning Algorithms

Supervised machine learning algorithms are a type of machine learning algorithm that learn from labeled training data. In supervised learning, the training data consists of a set of examples that are labeled with the correct output. The goal of the algorithm is to learn a function that maps inputs to the correct output based on the examples in the training data.

Some common types of supervised learning algorithms include:

1. Linear regression:This algorithm is used for classification tasks, where the goal is to predict

2. Support vector machines: This algorithm is also used for classification tasks, but it works byfinding a decision boundary that maximally separates the different classes.

3. Decision trees: This algorithm creates a tree-like model of decisions and their possible consequences, used for classification and regression tasks.

4. Random forests: This is an ensemble method that combines multiple decision trees and outputs the class that is the mode of the class's output by individual trees.

5. Naive Bayes: This algorithm is used for classification tasks and makes predictions based on the probability of each class given the features of the example.

6. K-nearest neighbors: This algorithm makes predictions based on the K examples in the training set that are closest to the input example.

There are many other types of supervised learning algorithms as well. The choice of which algorithm to use depends on the nature of the problem and the available data.

b. Unsupervised Machine Learning Algorithms

Unsupervised machine learning algorithms are a type of machine learning algorithm that work without being provided with explicit labeled training examples. Instead, unsupervised learning algorithms try to find patterns and relationships in the data by themselves. Some common applications of unsupervised learning include anomalydetection, density estimation, and clustering. Some popular unsupervised learning algorithms include:

1. **K-means clustering:** a method for partitioning a dataset into a set of K clusters, where each data point is assigned to the cluster with the nearest mean.
2. **Hierarchical clustering**: a method for finding a hierarchy of clusters by building a tree-like structure, where the clusters at each level are combined based on some similarity measure.
3. **DBSCAN:** a density-based clustering algorithm that can find arbitrarily shaped clusters and is robust to noise.
4. **Anomaly detection:** a method for identifying unusual patterns in a dataset, which may be indicative of some type of error or fraud.
5. **Autoencoders:** a neural network architecture for learning a low-dimensional representation of an input dataset, which can be used for dimensionality reduction, feature extraction, and generating synthetic examples.
6. **Generative models:** a type of model that can generate synthetic examples that are similar to a given training dataset, by learning the underlying distribution of the data.

V. REINFORCEMENT LEARNING ALGORITHMS

Reinforcement learning is a type of machine learning in which an agent learns to interact with its environment in order to maximize a reward signal. It is often used in applications where an agent needs to learn how to make decisions in complex, dynamic environments. Some popular reinforcement learning algorithms include:

1. **Q-learning:** a model-free algorithm that learns a policy for selecting actions by estimating the maximum expected reward for each action, based on the observed rewards and transitions in the environment.
2. **SARSA:** a model-based algorithm that estimates the expected reward and value for each state and action, based on the observed transitions and rewards in the environment.
3. **DQN (Deep Q-Network):** a variant of Q-learning that uses a neural network to approximate the action-value function, which allows it to learn from high-dimensional state spaces.
4. **Policy gradient methods:** a class of algorithms that directly optimize the policy for selecting actions by adjusting the parameters of the policy function using gradient descent.
5. **Actor-critic methods:** a class of algorithms that learn both a policy for selecting actions and a

value function that estimates the expected reward for each state or action.

6. **Monte Carlo Tree Search:** a tree-based search algorithm that uses simulated rollouts of the game to evaluate the value of each action and select the best one.

VI. APPLICATIONS TO MACHINE LEARNING

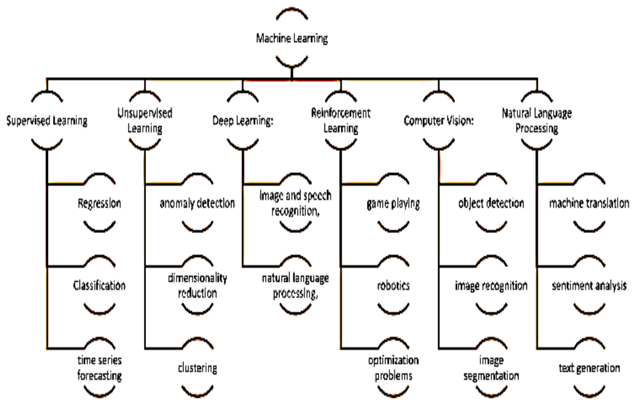

Figure 2:Tools commonly used in Machine Learning (Kumar, Y., 2020)

Figure 2 depicts the many different application domains and sub-domains we found when researching the field of machine learning. Below, we'll go over some of the most prevalent real-world uses of machine learning. Voice assistants take text inquiries as input, interpret them with the use of machine learning models, and then take some sort of action based on that interpretation. These systems incorporate a Natural Language Processing (NLP) module in addition to an Automatic Sound Recognition (ASR) feature that translates analog voice inputs into text. The cameras in smart phones use a neural network educated on millions of photos to do image identification. The camera uses machine learning to figure out what it's looking at, then adjusts the lighting, selects the best exposure, and fine-tunes the colors to best suit the scene. Discussed in (Kumar, Y, 2020) are the numerous forms of machine learning algorithms that may be applied to a wide variety of tasks, including data mining, predictive analytics, image processing, and many more.

With the help of convolution neural networks, Face Unlocking can accurately categorize and recognize a user's face, allowing for foolproof authentication. Automatic face detection from a database of millions of photographs requires first the identification of features in those images and then the training of a neural network.

In order to filter data and give the most relevant items to consumers, Recommendation Engine combines two methods, namely content-based and collaborative filtering. It does this by calculating the degree to which a user's past actions and preferences are similar. E-commerce sites that offer us the newest things based on our browsinghistory are one example, as are food delivery services that show us the restaurants, they recommend based on our order history.

Chatbots employ machine learning to mimic human conversation, whether it's to answer frequently asked questions or to provide personalized guidance through a task. These are always improving in intelligence, which is great for the end user.Google Search is able to provide the finest results because it learns from its past mistakes and uses this information to better tailor its future suggestions.

VII. CHALLENGES IN MACHINE LEARNING

Whilst significant progress has been made in the field of machine learning in recent decades, significant problems still need to be addressed. You'll need a massive dataset in order to train a model to do the work you need it to do. To create an image classifier that can accurately determine if an uploaded photo depicts a cat or dog, for instance, we would need to train our Machine Learning model with thousands of examples. Complex computations are needed even for simple tasks when using machine learning and models. That is why specialized hardware like graphics processing units and tensor processing units is necessary. Scientists are always looking for new ways to improve existing algorithms, making them faster, more accurate, and easier to use in a variety of situations.

VIII. CONCLUSION

In conclusion, machine learning has emerged as a prominent and widely-used aspect of artificial intelligence, driving significant advancements in digitalization solutions. It holds immense potential for augmenting and virtual reality technologies. Researchers are actively striving to enhance the intelligence and sophistication of machines, drawing inspiration from human intelligence. Accuracy improvement is a key focus, achieved through data analysis and algorithm creation. Machine learning's impact spans across various sectors, including healthcare, finance, transportation, and agriculture, with endless possibilities for expansion. As we look ahead, it is exciting to witness the transformative power of machine learning in shaping diverse fields. This review paper provides a comprehensive overview of machine learning, its subfields, and its applications, highlighting its role in enhancing knowledge and work performance.

REFERENCES

[1] Shinde, P. P., & Shah, S. (2018). A review of machine learning and deep learning applications. 2018 Fourth International Conference on Computing Communication Control and Automation (ICCUBEA).

[2] Hua, Z., & Zhang, B. (2006). A hybrid support vector machines and logistic regression approach forforecasting intermittent demand of spare parts. Applied Mathematics and Computation, 181(2), 1035-1048.

[3] Hong, W., Dong, Y., Chen, L., & Wei, S. (2011). SVR with hybrid chaotic genetic algorithms for tourism demand forecasting. Applied Soft Computing, 11(2), 1881-1890.

[4] Zarandi, M. H., Hadavandi, E., &Turksen, I. B. (2012). A hybrid fuzzy intelligent agent-based system for stock price prediction. International Journal of Intelligent Systems, 27(11), 947-969.

[5] Shahrabi, J., Hadavandi, E., &Asadi, S. (2013). Developing a hybrid intelligent model for forecasting problems: Case study of tourism demand time series. Knowledge-Based Systems, 43, 112-122.

[6] Pai, P., Hung, K., & Lin, K. (2014). Tourism demand forecasting using novel hybrid system. Expert Systems with Applications, 41(8), 3691-3702.

[7] Zhang, Y., Wang, Y., Zhou, G., Jin, J., Wang, B., Wang, X., &Cichocki, A. (2018). Multi-kernel extreme learning machine for EEG classification in brain-computer interfaces. Expert Systems with Applications, 96, 302-310.

[8]HTTP://ljournal.ru/WP-content/uploads/2017/03/a-2017-023.pdf. (2017). Foxon, F. (2021). Ammonoid taxonomy with supervised and unsupervised machine learning algorithms.

[9] Zhu, X., & Goldberg, A. B. (2009). Introduction to semi-supervised learning. Synthesis Lectures on Artificial Intelligence and Machine Learning, 3(1), 1-130.

[10] Sutton, R. S. (1992). Introduction: The challenge of reinforcement learning. Reinforcement Learning, 1-3. Caruana, Rich, "Multitask learning.", Machine learning, vol. 28, no. 1, (1997), pp. 41-75.

[11] Opitz, D., & Maclin, R. (1999). Popular ensemble methods: An empirical study. Journal of Artificial Intelligence Research, 11, 169-198.

[12]https://www.researchgate.net/profile/Sajad_Ahmadian2/publication/282818770/figure/fig2/AS:406577333981185@1473946998408/The-structure-of-the-two-layered-feed-forwardneural-network.png

[13] Shama, Vidushi, SachinRai, and Anurag Dev, "A comprehensive study of artificial neural networks", International Journal of Advanced research in computer science and software engineering, (2012), vol. 2, no. 10.

[14] S. B. (2014). A survey: Research summary on neural networks. International Journal of Research in Engineering and Technology, 03(15), 385-389.

[15] Kumar, Y., Kaur, K., & Singh, G. (2020). Machine learning aspects and its applications towards different research areas. 2020 International Conference on Computation, Automation and Knowledge Management (ICCAKM).

.