Review On Machine Learning Techniques, Applications, And Challenges

V. Sriharidevi, Dr.R.A. Roseline
[1]*Research scholar, Department of Computer Applications, Government Arts College, Coimbatore-18.*
[2]*Head of the Department, Department of Computer Applications,*
Government Arts College, Coimbatore -18

***Abstract*— The evolution of artificial intelligence in the current digital era has originated in pattern recognition from Machine Learning (ML) and computational learning theory, and has evolved into a core area of computer science and technology. It utilizes data-driven algorithms to generate predictions and improve system performance. As a rapidly developing technology within the IT industry, Machine Learning has significantly impacted sectors such as healthcare, engineering, and agriculture. This paper presents an overview of Machine Learning algorithms aimed at enhancing application intelligence and examines both their practical applications and the challenges encountered in their implementation.**

***Index Terms*— Machine Learning techniques, supervised learning, unsupervised learning, semi-supervised learning, Reinforcement learning, applications, challenges.**

I. INTRODUCTION

This Machine Learning institutionalized towards the end of 1950 is a constituent part of Artificial Intelligence. Over the course of the period, it reoriented its progress towards designing computationally viable and resilient algorithms to effectively perform analytical tasks, including clustering, classification, within the spheres of medicine, engineering and computing contexts. ML exemplifies as one among the groundbreaking and far-reaching technologies of the present era. ML allows the computer possessing the potential to facilitate autonomous learning and respond adaptively, devoid of human input. It formulates model through the intelligent, automated examination of intricate data, ensuring both speed and precision. It demonstrates the capability to derive insights from abstracted data to formulate solutions that augment operational performance. Its diversified potential enables widespread adoption in numerous areas ranging from healthcare, engineering, data automation, spam filtration, diagnostic systems, cybersecurity, and sound processing [1][2].

In the late 1950s, Machine Learning marked a pivotal advancement in AI systems [3]. Its trajectory shifted progressively to an algorithm framework that demonstrates both scalability and robustness. Over the past decade, methodologies have been widely adapted into address a diverse range of computational tasks, notable classifications, regression, and density evaluation. The techniques that have been estimated to have substantial utility across a variety of application domains, including bioinformatics, speech automation, social media spam mail detection, system version, fake analysis, and online advertising. Conceptual foundational algorithms and techniques stem from an interdisciplinary synthesis of statistics, applied mathematics, neuroscience research, and the computer science field, underscoring the cross-disciplinary character of modern learning research.

The subsequent canonical definitions articulate the central tenets of ML.

1) The construction of computer system-based models that replicate learning dynamics provides perfect knowledge, advances the efficiency and accuracy of the intelligent system performance [4].

2) The formulation of Machine Learning frameworks enables systematic solutions to knowledge acquisition, simultaneously advancing the functionality of computational systems [5].

In this current era, system software or systems that display intelligent behavior are mainly referred. whereas Machine Learning is a subset of artificial intelligence, meanwhile, deep learning is a subset of

Machine Learning. Circular diagram of Machine Learning, artificial intelligence, and deep learning shown in Figure.1. DL is included in both ML and the general field of AI, while ML is the process for acquiring knowledge from data or experience, which automates the

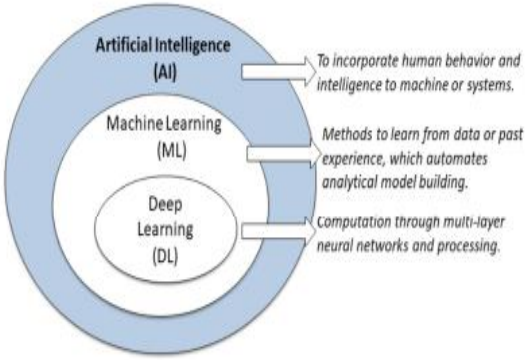

Figure 1. The diagram shows the relationship between Deep Learning, Machine Learning, and Artificial Intelligence.

In the generation of modelling the analysis, AI often integrates human intelligence and conduct into AI algorithms or systems. DL also describes data-driven learning methods that do computation using hierarchical layers present in neural networks for processing [6].

Deep learning technique is a basic technique in artificial intelligence, which refers to the process of processing data, methods, and model are automatically ready to study for complex data and structured data, with abstract data and more hierarchical data are taken from the original input data [7,8]. In the late 1950s, the pioneer and father of Machine Learning as "a field of study that enables computers to learn without being explicitly programmed." Machine Learning emphasizes domains such as pattern recognition, natural language processing, cognitive computing, image processing, knowledge representation, road traffic signals and pathways, cognitive radio, intrusion detection, optical networks, and others. Machine Learning challenges and tasks are generally grouped into three main categories due to the signals and feedback provided to the learning system [9].

Machines can easily compute knowledge; however, humans will find it challenging because of the vast amount of information available for some activities [10]. Because environments change over time, Machine Learning reduces the need to redesign systems owing to environmental changes [11]. Every day, humans acquire and discover new information, tasks, and data; Machine Learning makes it simple to keep track of this new information. In the world of information technology, Machine Learning is one of the newest and fastest-growing technologies. Machine Learning handles a variety of issues in fields such as engineering, agriculture, and medicine.

A variety of Machine Learning applications in numerous spheres of life are covered in this section. Machine Learning is well-suited for medical applications, particularly for cancer diagnosis, heart and lung prognosis, and prediction.

During covid-19 ML played a major role in identifying the problem for all kinds of patients, like children, medically diagnosed patients, and older people. It is also used to discover drug details. Neural networks, network Machine Learning approaches, weather and climate prediction and forecasting, and energy management with an emphasis on both solar and wind energy management are all areas of study in the discipline of network Machine Learning [12].

There are some notable reasons why Machine Learning is important:

(1) Optimizing system architecture for high-volume data ingestion is crucial to ensure accurate, timely, and scalable output generation. [13].

(2) Machine Learning is famous for pattern recognition. The patterns can be achieved using both small and large datasets, grouped, classified, and clustered for processing. [14].

(3). Insights and cognitive approaches present in Machine Learning help to achieve data extraction and interpretation for more analysis, decision making, personalization, security, forecasting, and prediction in ml [15].

II. MACHINE LEARNING TECHNIQUES

There are four categories of Machine Learning techniques and algorithms, with subcategories presented in Figure 2. We go into each kind of learning approach and its uses in brief in the sections that follow [16].

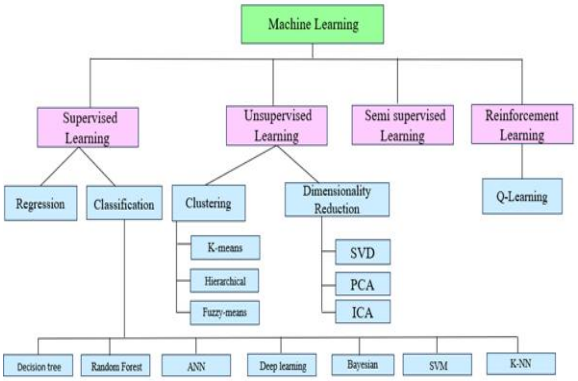

Figure 2. Classification of Machine Learning Techniques.

Four major types of Machine Learning Techniques are seen in the figure 2. Each type of learning technique in real-world situations is briefly discussed.

Figure 3 above states the step-by-step process of the fundamental working process of a Machine Learning prediction model.

In which phase 1 is to train the model on historical data, phase 2 consists of producing its outcomes for the most recent test data.

Which means Phase 1 processes historical data for predictive, and Phase 2 generates a new predictive model to give the new tested predicted data [17].

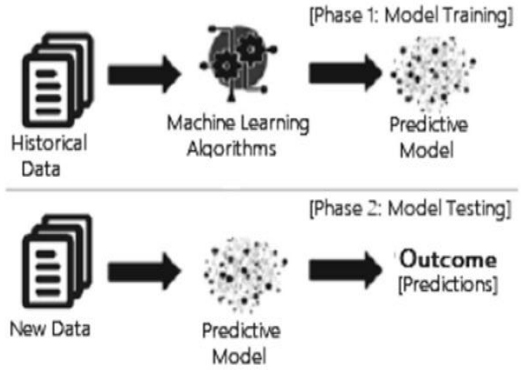

Figure 3. A structure of a Machine Learning model showing a predictive model with processed output.

1. Supervised Learning

Supervised learning is a process where an algorithm learns from a labeled dataset.

These datasets consist of input paired with the correct target output. The primary goal is to establish a functional mapping between these inputs and outputs

1.1 Regression

1.2 Classification

1.1 Regression

A regression algorithm is a type of supervised learning that involves predicting a continuous output variable. It is mainly used for prediction and forecasting. Most of the Machine Learning applications use prediction.

For example, financial, marketing, agriculture, engineering, and health care. Forecasting is used for weather and climate conditions for geographical field. Regression is essentially finding a relationship that is associated between the dependent variable and the independent variable.

1.2. Classification

Classification is a type of supervised learning in which an algorithm identifies new data points by applying the knowledge it has gained from the input data. We have two types of classification algorithms: binary (single-class) and multi-class, where the algorithm predicts between more than two classes [18].

These classification algorithms have seven types, which can be mapped according to the working methodology as follows,

(a)Artificial Neural Networks (ANNs)

Artificial Neural Networks (ANNs)intensively imitate the human brain and resemble it, stimulating the function of biological neurons. which was mainly used for network application areas or connectivism systems used in timeline process algorithms.

These ANNs are very strong and effective in finding patterns that are more complex and countless for human beings to extract and process. In the fig 4, the 3 layers are present in it. First layer is the input layer, where the data is ready for processing, second layer is the hidden layer, which consists of many patterns and more complex data, third layer is the output layer, which has predicted output that matches the original output.

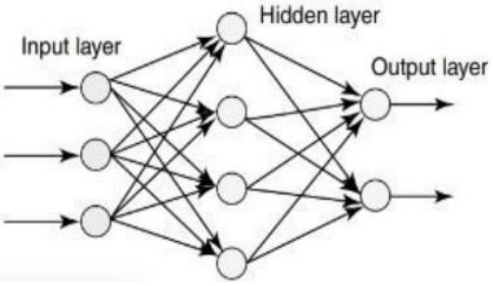

Figure 4. Artificial Neural Networks

Forecasting, prediction, and face recognition are a few examples of ANNs. In addition to the computational

power required for more sensitive tasks, the peer problem facing ANNs is the time required to train the model. The main composition of an ANNs are interconnected processing nodes, said to be neurons, which receive input signals, perform a weighted sum of the inputs, and applied for a non-linear activation function to produce an output signal [18].

(b) Deep Learning (DL)

Deep learning is a part of Machine Learning that involves the use of neural networks, in Machine Learning that allows a machine system to train itself to do the task. since it is a task-driven approach.

These deep learning methods, as a combination of many interconnected layers, are input layers, inner layers, or hidden layers, and an output layer that examines intricate details by utilizing many layers. Like ANN methodology, it uses a multi-layer perception for data learning. DL information processing and communication systems function similarly to the human brain. Deep learning suitable for both supervised learning and unsupervised learning methods. Above mention artificial intelligence algorithm and deep learning are neural-based methods [18].

(c) Support Vector Machine (SVM)

SVM is used for linear and non-linear classification. This algorithm is used to calculate margin values between two classes. A Hyperplane is used to separate two different classes in a dataset. This class has positive and negative values with a maximum margin size of classes [19]. Figure 5 shown picture represent of SVM.

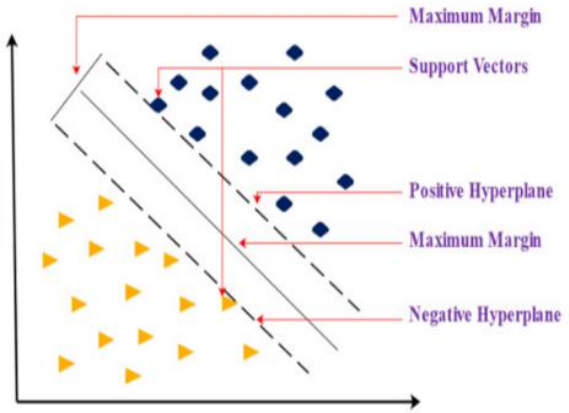

Figure 5: Structure of a support vector machine.

(d)K-Nearest Neighbor (K-NN)

K-Nearest Neighbor is used for both classification and regression in Machine Learning algorithms, a non-formal model, which means instance-based learning. It is a method that stores training data in memory and defers the process of generalization until a new instance needs to be classified. then utilizing their labels or values to predict the new data point's label or value.

Among the many benefits of k-NN are its interpretability and simplicity, its capacity to manage non-linear decision limits, and its resilience to high-dimensional and noisy input.

Among the restrictions are issues with distance and the requirement for relevant data. It is computationally complex when dealing with big datasets [20][21].

(e) Bayesian

A statistical method of education called Bayesian learning learns conditional independence from several statistical methods in order to establish the association within the datasets. Bayesian learning calculates posterior probabilities employing newly acquired information and various prior probability distributions.

The Bayesian approach is predicated on statistical learning. Using conditional independence and several statistical techniques, it ascertains the correlation between two datasets. The Bayesian approach is predicated on statistical learning. Using conditional independence and several statistical techniques, it ascertains the correlation between two datasets [18][22].

(e) Decision trees

Decision tree techniques in supervised learning are used to predict the relationship between input variables and target output variables by sorting.

A decision tree is used for multidimensional analysis with multiple classes. It has 2 tree structures, which has hierarchical structure consisting of nodes, branches, and leaves, which is similar to a tree data structure format.

In a decision tree, each basic first node is called a root node, the second node is said to be branches which has true or false attribute values, where each node assigns the classification [23]. The decision tree structure for casting is shown in Figure 6. Example weather report.

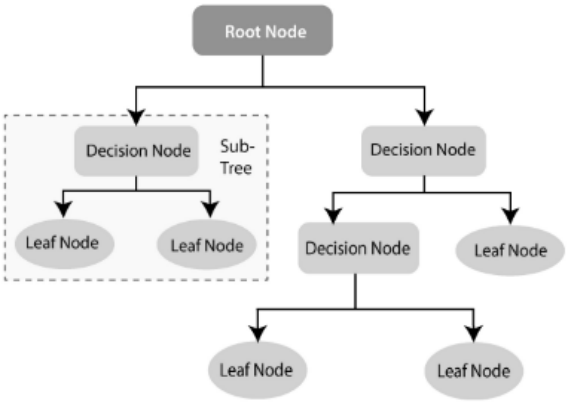

Figure 6: Structure of Decision Tree.

(f)Random Forest.

An assortment of trees forms the Random Forest, which is an ensemble algorithm in a supervised learning technique. Random forest combines many decision tree models. This combined model gives the results better than those from the individual decision tree models [24][25]

2. Unsupervised Learning:

In Machine Learning, the unsupervised learning idea analyzes unlabeled and unclassified data to uncover latent knowledge. Although the algorithms are designed to identify patterns, groups, sorting, sorting order, and other intriguing features in the data set.

Systems (computers)are operated to do itself which we are not instructed to do. Unsupervised learning: the algorithm divides the samples into classes as efficiently as possible using a data-driven approach used for clustering and feature reduction [9][18].

Unsupervised learning algorithms have two different types

2.1 Clustering

2.2 Dimensionality reduction

2.1 Clustering

One of the major areas in unsupervised learning algorithms in Machine Learning. We can view patterns and structures using data exploration. Clustering organizes have into two groups based on their similarity within the data. These clusters have similarity and are grouped according to their data partition. We have methods of clustering.

(a) k-means Clustering

The k-Means algorithm is also a partitional algorithm.is a straightforward and effective clustering algorithm that is more flexible with new samples and ensures convergence.

The data points are grouped using the k-means clustering technique, and each data point is assigned to the cluster with the least distance. The following are K-means working methods. It starts to select a random mean value, and secondly. Find the nearest mean number and add it to a particular cluster. Finally, I continue to do the selection and finding process [26].

(b) Hierarchical clustering

Hierarchical clustering is often intended for larger data sets. It uses either a top-down splitting method or a bottom-up methodology, which is a merging method, where each data point first creates its own cluster [27].

(c) Fuzzy-c

This method finds clusters using similarity metrics, including intensity, connectedness, and distance. When the data sets overlap, this clustering method identifies the best cluster centers. It is a method of thinking that is similar to human reasoning.

This method takes into account every possibility between 0 and 1, i.e., true and false. The network's uncertainty circumstances are handled by fuzzy logic [9].

2.2 Dimensionality reduction

Dimensionality Reduction (DR) technique seeks to minimize the distance in a latent space between distributions of various data sets.

Reduce the amount of data storage space and dimensions. It takes less time to calculate. It is possible to remove duplicated, noisy, and irrelevant data. It is possible to optimize the quality of data. Enhances the efficiency and accuracy of an algorithm. Permit data visualization.

It improves performance and makes the classification process easier. It has 3 different methods [9]

(a)Singular Value Decomposition (SVD)

SVD is an example of the fascinating geometrical characteristic for PCA executed in theory and structural for linear transformations because of its many beautiful algebraic properties [28][29].

(b) Principal Component Analysis (PCA)

Principal component analysis (PCA) is a statistical method for reducing dimensionality. As much of the variation of the original data is preserved as feasible

when converting a sizable collection of correlated variables into a smaller set of uncorrelated variables [22][26].

(c) Independent Component Analysis (ICA)
New version of PCA, often referred to as Independent Component Analysis (ICA). By minimizing higher-order correlations, since PCA is unable to perform, ICA is used to study digital photographs and data [18].

C.Semi-Supervised Learning
Machine Learning techniques advantage of using supervised and unsupervised learning in a single algorithm is called semi-supervised learning.
Semi-supervised learning has both methods built using a combination of labeled and unlabeled data, focus with minimum cost and effort in obtaining labels for instances [30].
Semi-supervised learning can do training and transudative support vector machine. It betters the prediction results than it could with just the labeled data.
They can deal with high-dimensional data, rough data that is loosely coupled data, and low-dimensional data.
Machine translation, fraud detection, data labeling, and text classification all make extensive use of these techniques [31].
Semi – supervised learning has tremendous practical value, for example

- Speech recognition,
- Natural language parsing,
- Spam filtering.
- Video surveillance.

D. Reinforcement Learning.
Reinforcement learning technique, learn in the absence of a training dataset, and interacts with the environment.
The decision-making method is taken in sequential order. Increase in performance. The action taken trial and error search between two different ways between the environment and the agent [32].

1. Positive outcomes rewards to increase in efficiency in the network.
2. A negative positive outcome degrades the performance in the network.

In the figure 7. The sequence of decision-making made the agent and the environment.

Figure 7. Reinforcement Learning structure

Formally, the action takes place between the agent and the environment. State and rewards go to the input of the agent from the environment.
The processes of input and output are connected. Raw data is used as the first input. after receiving the output. For the following procedure, the output is returned to the input. This loop keeps going until the desired outcome is obtained

III. SUMMARY OF MACHINE LEARNING ALGORITHM

S. no	Types	Purpose	Advantage	Disadvantage	Real-world examples
1	Decision Tree	Terms and conditional rules-based decision	Prior Domain expertise is not required to build a decision tree It simplifies complex decisions by breaking the hierarchy into smaller parts.	It can produce only a single output attribute It is sensitive to small variations in the dataset, and it is unstable for continuous data, the tree leads to a complicated structure	Bank Loan Approval
2	Support Vector Machine	Error detection classification	It is a straightforward model that works with both the numeric and categorical data types.SVM effectively handles the nonlinear data points in SVM, overfitting is minimal due to the margin maximization	The computational cost is high in SVM Varied kernels produce datasets with different performance outcomes Choosing an appropriate kernel requires expert knowledge While dealing with large datasets, the training time is expensive and slow.	Email spam detection.

3	K-Nearest Neighbors (KNN)	Similarity Prediction	The KNN mechanism is easy to implement and analyze the training phase in KNN is rapid. It requires no training time, allowing effortless data updating	KNN demands high memory storage Outliers affect the predictions of KNN Prediction time is long in the case of KNN	Movie recommendation systems.
4	Bayesian	Probabilistic Reasoning	The calculation process in this model is uncomplicated. Its high-speed accuracy makes it the best choice for large-scale datasets.	When the variables are interdependent, accuracy drops. It identifies groupings in less complicated ways and widely used for clustering large datasets.	Medical diagnosis for heart, brained
5	k-Means	Cluster	It is an effective algorithm for clustering tasks. It can easily integrate fresh examples This algorithm ensures convergence during optimization	Predefined cluster count is required beforehand. It is incompatible with a categorical algorithm It cannot recognize the cluster with indefinite shapes.	Retail customer segmentation
6	Principal Component Analysis (PCA)	Sequential decision-making process	PCA is a fast and user-friendly algorithm It converts complex datasets into visual data PCA effectively eliminates redundant and correlated variables.	PCA demands high memory for processing. Storing the correlation matrix requires extensive space PCA reduces the interpretability of original variables.	Image processing Facial recognition systems.
7	Reinforcement Learning	Sequential decision making	It adapts to achieve optimal decision patterns Reinforcement learning is not likely influenced by bias RL performs well without heavy labelled datasets.	RL often faces the problem of finding local minima Implementation of RL is costly and resource-intensive for straightforward, simple tasks, RL is unsuitable.	Fully automated Self-driving cars.
8	Regression	Continuous value Predication	It shows the importance of each feature. Regression performs consistently even with varying data volumes. Effective for diagnosing and refining errors.	Regression focuses on the average relationship between variables The model is heavily affected by extreme values The independent assumption often fails in real datasets	American house prices Prediction
9	Naive Bayes	Text classification	Naive Bayes trains swiftly due to its minimal data requirements. It works best when the predictors are mutually independent.	Native Bayes treats all features as independent, unlike its unrealistic assumption Smoothing is needed for some datasets in certain cases	Psychological sentiment analysis of product reviews given in online marketing feedback
10	Artificial Neural Network (ANN)	Pattern Recognition	ANN efficiently learns complex relationships and simple interactions This network is robust to noise and imperfect data ANN can handle linear and nonlinear data patterns	It is obliged to solve the local minima problems It often overfits while handling small datasets Understanding the internal workings of an ANN is quite challenging.	Handwritten digit recognition (the MNIST dataset).

IV APPLICATIONS OF MACHINE LEARNING

In this 5th generation of computers, automated learning has become popular in many application areas owing to its capacity to learn from the past while making educated decisions. We can discuss a few applications of Machine Learning techniques as follows.

1. Forecasting and predictive modeling: A crucial field of Machine Learning application has contingent upon discovering and capitalizing on connections amongst explanatory factors and predicted variables from prior incidents to foresee an unanticipated result [34].

As an instance, identifying suspects or offenders following a criminal act that has occurred, or spotting fraud with credit cards in real time. Furthermore,

merchants may use Machine Learning algorithms for handling inventory, avert delayed delivery issues, improve logistics and warehouse management in e-commerce, and obtain a greater grasp of the preferences and actions of their consumers.

Artificial neural networks, deep learning, and all other methods in Machine Learning are typically employed in the field. Since accurate forecasting and prediction provide benefits, it could benefit corporations, industries, and almost any organization to make better decisions. This spans a wide range of industries, notably social networking, healthcare, banking and financial services, e-commerce, telecommunications, government entities, sales, and marketing [35,36].

Today, transportation networks are a crucial component of every country's growth in revenue. However, traffic congestion contributes to catastrophic problems in numerous municipalities around the globe, resulting in traffic signal monitoring, delays in dedicating Google Path Finder, an increase in petrol and gasoline costs, greater carbon dioxide and nitrogen gas emissions, accidents, crises, and a decline in the general quality of life in cosmopolitan society [30,40,41]. Consequently, a public transport system with intelligence that detects air quality levels, water supply systems, and public movements is proactively predicted and forecasts the conditions present in the areas. This can be accurate while using Machine Learning and deep learning modelling, which could assist in mitigating the difficulties

2. Cybersecurity: Broadly speaking, cybersecurity is to solve the issues for network broadcasting, hardware and software systems, and secure and independent data transfer using codes and algorithms, which is used in the internet based online l assaults. Subsequently serves as one of the most crucial. Machine Learning with cybersecurity technology keeps to learns perpetually by analyzing data to make patterns, installing security software within the system. Detecting the network traffic, detecting threats from within, predicting the precise positions of dangerous online neighborhoods, shielding users while they surf, or spying on suspicious activity to safeguard data in the cloud. Clustering algorithms, for instance, are capable of helping find cyber-anomalies and policy breaches.

Machine Learning classification models that employ safety characteristics can help identify numerous types of breaches or assaults. Enormous security datasets can additionally be used with an assortment of computational security models. Furthermore, an indispensable element of a rule-based security system could comprise protective rules produced by learned association rule algorithms [37]. Consequently, we might infer that by utilizing the numerous learning methodologies covered in the section, professionals in cybersecurity might become more vigilant in effectively averting breaches and attacks. Cybersecurity experts might be a more proactive in effectively avoiding breaches and assaults with the adoption of Machine Learning initiatives and algorithms.

3. Internet of Things (IoT) A further vital component of Industry sophisticated ones by facilitating transmission of information and task automation, eliminating the requirement for human oversight. IoT gets seen as the technology of the future, given that it possesses the potential to boost virtually all facets of our lives, notably smart homes, smart government, education, communication, transportation, retail, agriculture, health care, and business [37][38]. Smart cities that deploy technology to boost municipal amenities and elevate the standard of living for residents are one of the fundamental IoT application areas. Processing data from different devices with different needs and backgrounds is still a significant difficulty. Machine Learning has become a potential option for handling high-dimensional and large-scale data in Internet of Things networks in order to process heterogeneous data. By offering valuable insights for a range of applications, Machine Learning may help IoT networks. There are still several obstacles to ML's enormous potential in the context of IoT. Numerous problems could be tackled with Machine Learning approaches, notably parking availability, traffic prediction in metropolitan and cosmopolitan cities, calculating the overall energy consumption of the inhabitants' resource management system over a period of time, and valuing people in making swift and contextually sensitive decisions.

4. Medical diagnosis during COVID-19: In several medical domains, including sickness prediction, health information collection, data regularity detection, handling patients, etc., Machine Learning might assist with solving analytical and predictive issues [42,43].

The infectious circumstances termed coronavirus disease (COVID-19) has been tied to international health organization. During COVID-19, learning tactics are becoming more and more popular. Throughout the COVID-19 pandemic, methodologies for learning have been utilized to distinguish between individuals at risk, their death rate, and other oddities [44].

Subsequently proves beneficial for finding and curing diseases, forecasting COVID-19 outbreaks, as well as discovering further about the genesis of the virus may foresee the locations and times at which the COVID-19 virus seems most likely to spread, probably informing those regions about getting ready accordingly. Deep learning is seen as a vital method for prospective applications, specifically in conjunction with the COVID-19 pandemic, and it likewise offers intriguing solutions to medical image processing problems. With everything accounted for, in the field of medical science and healthcare, advanced Machine Learning, with its subset used to diagnose pandemic issues, helps to take decision.

5. E-commerce-based financial services recommendation: Among the many crucial components for nearly all e-commerce sites in recent years is product recommendations, which also happens to be another of the most prevalent and frequently utilized. In the current scenario, to attract customers, Machine Learning techniques are used to gauge their products kept in priority to purchases and offer psychological behavior product suggestions for their next purchase. It will affect previous product purchase may be in the same day or two or three days, or sometimes later.

By way of example, e-commerce companies might swiftly place product recommendations and offers through an investigation into surfing statistics and the rate of click-through for a particular product. Through the use of automated forecasting that utilizes methods involving Machine Learning, many online marketing companies like Alibaba… [45], may mitigate shortages of goods, enhance logistics and storage, and optimize their inventory. The potential growth of marketing and sales is in gathering, evaluating, and applying consumer data, such as repetition of product, timeline for buying each product, and feedback. Moreover, businesses use packages and organize to fulfill the requirements of their customers, conserving existing customers while drawing fresh ones.

6. Speech recognition and language processing: Consequently, language processing or Natural language processing makes it possible for computers perform things such as reading text, listening to language, interpreting it, gauging emotion, and identifying its key elements for Machine Learning [46,47]. Chatbots, virtual personal assistants, speech recognition, text descriptions, linguistic or natural language translation, and other activities are examples of tasks associated with natural language processing, occasionally referred to as deep learning for text, data mining for analyzing the issues and intelligences of AI gives response which seeks public sentiment, perspectives of view from materials found in websites, analyses, social networks, communities, media, and numerous other sources.

Companies and brands, for instance, utilize sentiment research on social networking sites or the internet more generally to figure out brand, product, or service. Typically, sentiment analysis is conceived that investigates messages like awards and reviews of a product. Emotional feelings in marketing trends are such as satisfaction, good, and excellent are coined.

7. Image processing and pattern recognition for communication

Recognition of pictures [48], which accurately recognizes, comes from Machine Learning techniques. Picture recognition, for instance, is extensively utilized in order to identify people or characters in a picture, tag concepts come from internet-based communication medias and in the medical area, an X-ray or scan can access for is cancerous. Speech recognition [49] is also exceptionally prevalent and often employs language and sound models [50], akin to Google Assistant, Cortana, and other gadgets that employ Machine Learning techniques. Pattern recognition [51], for instance, is the automated detection of regularities and patterns in data in picture analysis. Selection of features, clustering, classification, and sequence labeling constitute some of the Machine Learning techniques employed within the discipline.

8. smartphone apps and user behavior analysis: context–awareness is the ability of a system to learn

about what happens in the what is happening at any given moment and change its behavior accordingly [50,51]. Context-aware computing automatically retrieves, assesses, and modifies data using both hardware and software [52,53]. The potential of AI, especially Machine Learning algorithms that can learn from contextual data, has led to a significant expansion in the environment for developing mobile apps [55].

9. Sustainable agriculture: Agriculture is indispensable for the ongoing existence of all living human beings. Sustainable agriculture provides a way to increase agricultural productivity while monitoring adverse effects on the environment [54]. Cooperative bank for fertilizers and supply chains relies on information, skills, technology, and other expertise. By harnessing the rising count of values gathered by emerging technologies, such as mobile devices, internet-based devices, and IoT-based active screening and system are used, knowledge transfer aids farmers in making superior choices about advancing the use of sustainable agriculture methods [56,57].
Pre-production (for forecasting crop yield, soil attributes, drip irrigation specifications, etc.); production (for forecasting weather, identifying diseases, weeds, managing soil nutrients, managing livestock, etc.); processing (for estimating demand, production planning, etc.); and distribution are just a few of the stages of sustainable agriculture where Machine Learning can be applied to inventory control, consumer analysis, etc[58].

V. CHALLENGES OF MACHINE LEARNING

This field provides an assortment of innovative studies that have not been previously discussed

1. Resource Management:
Effectively optimizes assignments over time while consuming the least amount of network energy by leveraging local knowledge and application limits. Each sensor node learns the minimum number of resources required to carry out its planned tasks and optimizes its reward prospects by determining the optimal settings for the intended use. It consists of five fundamental tasks: (a) creating a single measurement from several measurements, (b) sending a message to the next level, (c) receiving incoming messages, (d) taking samples for readings, and (e) applying sleep mode to the radio. To optimize network longevity, these steps must be taken in the right sequence. Nevertheless, since events lack a set schedule, the network is unable to determine priority. In one instance, a node may not be able to forecast when an object will get close enough to begin sampling. This difficulty is addressed by the DIRL method, which generates the required knowledge of priority based on a predefined Q-learning process that uses cost values and rewards [59].

2. Analyzing Animal Behavior Classification:
Classification of animals can be done using a decision tree algorithm [60], which makes perfect and accurately classify the behavior of a herd of animals as either passive or active using information such as action, movement happen in the body and neck pitch angle. The genetic approach to defining animal behavior is useful since it is simple to use and requires few key characteristics [61].

3 Clock Synchronization Using Self-Organizing Maps:
Sensor nodes must synchronize their clocks, given that a significant number of operations rely upon coherence. It is thought that self-organizing maps (SOM) provide reliable clock synch in large-scale networks. Absent the assistance of a central timing device, the nodes are capable of providing near-optimal estimations of the current time despite having limited computation and storage capabilities. However, the approach relies on the supposition that each node is uniformly dispersed of transmission power is monitored with in the area [62].

4. Air pollution detection:
A neural network-based approach has been proposed for assessing air pollution that can be detected by following levels using environmental data, i.e., PM 2.5, PM 10, and NO2. Sensor nodes are used to measure temperature and humidity. Neural networks serve a purpose in this technique to determine air quality and gas concentration. introduced a new standard utilizing a neural network method for handling lighting in smart buildings. They developed a new mathematical formula using a radial basis function (RBF) neural network [63].

VI. CONCLUSION

This review paper gives an overview of Machine Learning , including its algorithms and methods. This review of Machine Learning methodology and various domains like medical science, network, and agriculture. Fine-tuned technology in this current world at present is machine language, which provides problem-solving and deals with various issues. The content given includes Machine Learning algorithms like supervised, unsupervised, semi-supervised learning, and reinforcement learning.

REFERENCES

[1] J. Bi, H. Yuan, and M. Zhou, "Temporal prediction of multiapplication consolidated workloads in distributed clouds," IEEE Transactions on Automation Science and Engineering, vol. 16, no. 4, pp. 1763–1773, 2019.

[2] J. Bi, H. Yuan, L. Zhang, and J. Zhang, "SGW-SCN: An integrated machine learning approach for workload forecasting in geo-distributed cloud data centers," Information Sciences, vol. 481, pp. 57–68, 2019.

[3] T. O. Ayodele, "Introduction to machine learning," in New Advances in Machine Learning. Rijeka, Croatia: InTech, 2010.

[4] A. H. Duffy, "The 'what' and 'how' of learning in design," IEEE Expert, vol. 12, no. 3, pp. 71–76, 1997.

[5] P. Langley and H. A. Simon, "Applications of machine learning and rule induction," Communications of the ACM, vol. 38, no. 11, pp. 54–64, 1995.

[6] I. H. Sarker, "AI-driven cybersecurity: An overview, security intelligence modeling and research directions," SN Computer Science, vol. 2, no. 1, p. 21, 2021.

[7] I. H. Sarker, "Machine learning: Algorithms, real-world applications, and research directions," SN Computer Science, vol. 2, no. 3, pp. 1–21, 2021.

[8] I. H. Sarker, "Deep learning: A comprehensive overview on techniques, taxonomy, applications and research directions," SN Computer Science, vol. 2, no. 6, pp. 1–20, 2021.

[9] M. Pundir and J. K. Sandhu, "A systematic review of quality of service in wireless sensor networks using machine learning: Recent trends and future vision," Journal of Network and Computer Applications, vol. 188, p. 103084, 2021.

[10] I. H. Witten and E. Frank, Data Mining: Practical Machine Learning Tools and Techniques. San Francisco, CA, USA: Morgan Kaufmann, 2005.

[11] I. H. Sarker, M. M. Hoque, M. K. Uddin, and A. Tawfeeq, "Mobile data science and intelligent apps: Concepts, AI-based modeling and research directions," Mobile Networks and Applications, pp. 1–19, 2020.

[12] I. H. Sarker, "Machine learning: Algorithms, real-world applications and research directions," SN Computer Science, vol. 2, p. 160, 2021.

[13] P. Lison, An Introduction to Machine Learning. Language Technology Group (LTG), 2015.

[14] J. Brownlee, "Machine learning mastery," 2014. [Online]. Available: http://machinelearningmastery.com/discover-feature-engineering-howtoengineer-features-and-how-to-getgood-at-it

[15] A. Lavecchia, "Machine-learning approaches in drug discovery: Methods and applications," Drug Discovery Today, vol. 20, no. 3, pp. 318–331, 2015.

[16] J. Qin, W. Fu, H. Gao, and W. X. Zheng, "Distributed k-means algorithm and fuzzy c-means algorithm for sensor networks based on multiagent consensus theory," IEEE Transactions on Cybernetics, vol. 47, no. 3, pp. 772–783, 2017.

[17] I. H. Sarker, Machine Learning: Algorithms, Real-World Applications and Research Directions. Singapore: Springer Nature, 2021.

[18] P. Nayak, G. K. Swetha, S. Gupta, and K. Madhavi, "Routing in wireless sensor networks using machine learning techniques: Challenges and opportunities," Information Processing and Management, 2021.

[19] I. Steinwart and A. Christmann, Support Vector Machines. New York, NY, USA: Springer, 2008.

[20] M. F. Aslan, M. F. Unlersen, K. Sabanci, and A. Durdu, "CNN-based transfer learning-

BiLSTM network: A novel approach for COVID-19 infection detection," Applied Soft Computing, vol. 98, p. 106912, 2021. Y. Li, L.E. Parker, Nearest neighbor imputation using spatial–temporal correlations in wireless sensor networks, Inf. Fusion 15 (2014) 64–79.

[21] M. Pundir and J. K. Sandhu, "A systematic review of quality of service in wireless sensor networks using machine learning: Recent trend and future vision," Journal of Network and Computer Applications, vol. 188, p. 103084, 2021.

[22] K. Makantasis, K. Karantzalos, A. Doulamis, and N. Doulamis, "Deep supervised learning for hyperspectral data classification through convolutional neural networks," in Proc. IEEE Int. Geoscience and Remote Sensing Symp. (IGARSS), 2015, pp. 4959–4962.

[23] B. Alotaibi and K. Elleithy, "A new MAC address spoofing detection technique based on random forests," Sensors, vol. 16, no. 3, pp. 1–14, 2016.

[24] W. Elghazel, K. Medjaher, N. Zerhouni, J. Bahi, A. Farhat, C. Guyeux, and M. Hakem, "Random forests for industrial device functioning diagnostics using wireless sensor networks," in Proc. IEEE Aerospace Conf., 2015, pp. 1–9.

[25] L. Wang, "Discovering phase transitions with unsupervised learning," Physical Review B, vol. 94, no. 19, p. 195105, 2016.

[26] X. Xu, R. Ansari, A. Khokhar, and A. V. Vasilakos, "Hierarchical data aggregation using compressive sensing (HDACS) in WSNs," ACM Transactions on Sensor Networks, vol. 11, no. 3, p. 45, 2015.

[27] P. Guo, J. Cao, and X. Liu, "Lossless in-network processing in WSNs for domain-specific monitoring applications," IEEE Transactions on Industrial Informatics, vol. 13, no. 5, pp. 2130–2139, 2017.

[28] G. Gennarelli and F. Soldovieri, "Performance analysis of incoherent RF tomography using wireless sensor networks," IEEE Transactions on Geoscience and Remote Sensing, vol. 54, no. 5, pp. 2722–2732, 2016.

[29] J. Han, J. Pei, and M. Kamber, Data Mining: Concepts and Techniques. Amsterdam, Netherlands: Elsevier, 2011.

[30] I. H. Sarker, A. S. M. Kayes, S. Badsha, H. Alqahtani, P. Watters, and A. Ng, "Cybersecurity data science: An overview from machine learning perspective," Journal of Big Data, vol. 7, no. 1, pp. 1–29, 2020.

[31] Y. S. Abu-Mostafa, M. Magdon-Ismail, and H. T. Lin, Learning from Data. AMLBook, 2012.

[32] A. Dey, "Machine learning algorithms: A review," International Journal of Computer Science and Information Technologies, vol. 7, no. 3, pp. 1174–1179, 2016.

[33] L. Cao, "Data science: A comprehensive overview," ACM Computing Surveys, vol. 50, no. 3, p. 43, 2017.

[34] M. S. Mahdavinejad, M. Rezvan, M. Barekatain, P. Adibi, P. Barnaghi, and A. P. Sheth, "Machine learning for Internet of Things data analysis: A survey," Digital Communications and Networks, vol. 4, no. 3, pp. 161–175, 2018.

[35] B. Ślusarczyk, "Industry 4.0: Are we ready?" Polish Journal of Management Studies, vol. 17, 2018.

[36] A. Zanella, N. Bui, A. Castellani, L. Vangelista, and M. Zorzi, "Internet of things for smart cities," IEEE Internet of Things Journal, vol. 1, no. 1, pp. 22–32, 2014.

[37] Y. Zheng, S. Rajasegarar, and C. Leckie, "Parking availability prediction for sensor-enabled car parks in smart cities," in Proc. IEEE Int. Conf. Intelligent Sensors, Sensor Networks and Information Processing (ISSNIP), 2015, pp. 1–6.

[38] [38] A. Boukerche and J. Wang, "Machine learning-based traffic prediction models for intelligent transportation systems," Computer Networks, vol. 181, 2020.

[39] A. Essien, I. Petrounias, P. Sampaio, and S. Sampaio, "Improving urban traffic speed prediction using data source fusion and deep learning," in Proc. IEEE Int. Conf. Big Data and Smart Computing (BigComp), 2019, pp. 1–8.

[40] A. Essien, I. Petrounias, P. Sampaio, and S. Sampaio, "A deep-learning model for urban traffic flow prediction with traffic events mined from Twitter," in World Wide Web Conference, 2020, pp. 1–24.

[41] M. Fatima and M. Pasha et al., "Survey of machine learning algorithms for disease diagnostic," Journal of Intelligent Learning Systems and Applications, vol. 9, no. 1, p. 1, 2017.

[42] M. Nilashi, O. B. Ibrahim, H. Ahmadi, and L. Shahmoradi, "An analytical method for diseases prediction using machine learning techniques," Computers & Chemical Engineering, vol. 106, pp. 212–223, 2017.

[43] S. Gökhan and Y. Nevin, "Data analysis in health and big data: A machine learning medical diagnosis model based on patients' complaints," Communications in Statistics – Theory and Methods, pp. 1–10, 2019.

[44] A. Marchand and P. Marx, "Automated product recommendations with preference-based explanations," Journal of Retailing, vol. 96, no. 3, pp. 328–343, 2020.

[45] D. W. Otter, J. R. Medina, and J. K. Kalita, "A survey of the usages of deep learning for natural language processing," IEEE Transactions on Neural Networks and Learning Systems, 2020.

[46] I. H. Sarker, M. M. Hoque, M. K. Uddin, and A. Tawfeeq, "Mobile data science and intelligent apps: Concepts, AI-based modeling and research directions," Mobile Networks and Applications, pp. 1–19, 2020.

[47] H. Fujiyoshi, T. Hirakawa, and T. Yamashita, "Deep learning-based image recognition for autonomous driving," IATSS Research, vol. 43, no. 4, pp. 244–252, 2019.

[48] C.-C. Chiu, T. N. Sainath, Y. Wu, R. Prabhavalkar, P. Nguyen, Z. Chen, A. Kannan, R. J. Weiss, K. Rao, and E. Gonina et al., "State-of-the-art speech recognition with sequence-to-sequence models," in Proc. IEEE Int. Conf. Acoustics, Speech and Signal Processing (ICASSP), 2018, pp. 4774–4778.

[49] G. López, L. Quesada, and L. A. Guerrero, "Alexa vs. Siri vs. Cortana vs. Google Assistant: A comparison of speech-based natural user interfaces," in Proc. Int. Conf. Applied Human Factors and Ergonomics, 2017, pp. 241–250.

[50] Y. Anzai, Pattern Recognition and Machine Learning. Elsevier, 2012.

[51] A. K. Dey, "Understanding and using context," Personal and Ubiquitous Computing, vol. 5, no. 1, pp. 4–7, 2001.

[52] I. H. Sarker, "Context-aware rule learning from smartphone data: Survey, challenges and future directions," Journal of Big Data, vol. 6, no. 1, pp. 1–25, 2019.

[53] H. Sarker, M. M. Hoque, M. K. Uddin, and A. Tawfeeq, "Mobile data science and intelligent apps: Concepts, AI-based modeling and research directions," Mobile Networks and Applications, pp. 1–19, 2020.

[54] H. Zhu, H. Cao, E. Chen, H. Xiong, and J. Tian, "Exploiting enriched contextual information for mobile app classification," in Proc. 21st ACM Int. Conf. Information and Knowledge Management, 2012, pp. 1617–1621.

[55] R. Sharma, S. S. Kamble, A. Gunasekaran, V. Kumar, and A. Kumar, "A systematic literature review on machine learning applications for sustainable agriculture supply chain performance," Computers & Operations Research, vol. 119, 2020.

[56] N. Adnan, S. M. Nordin, I. Rahman, and A. Noor, "The effects of knowledge transfer on farmers' decision-making toward sustainable agriculture practices," World Journal of Science, Technology and Sustainable Development, 2018.

[57] H. Cobuloglu and I. E. Büyüktahtakın, "A stochastic multi-criteria decision analysis for sustainable biomass crop selection," Expert Systems with Applications, vol. 42, no. 15–16, pp. 6065–6074, 2015.

[58] K. Shah and M. Kumar, "Distributed independent reinforcement learning (DIRL) approach to resource management in wireless sensor networks," in Proc. Int. Conf. Mobile Adhoc and Sensor Systems, 2007, pp. 1–9.

[59] A. Mainwaring, D. Culler, J. Polastre, R. Szewczyk, and J. Anderson, "Wireless sensor networks for habitat monitoring," in Proc. 1st ACM Int. Workshop on Wireless Sensor Networks and Applications, 2002, pp. 88–97.

[60] E. Nadimi, H. T. Søgaard, and T. Bak, "Zigbee-based wireless sensor networks for classifying the behaviour of a herd of animals using classification trees," Biosystems Engineering, vol. 100, no. 2, pp. 167–176, 2008.

[61] L. Paladina, A. Biundo, M. Scarpa, and A. Puliafito, “Self-organizing maps for synchronization in wireless sensor networks,” in Proc. New Technologies, Mobility and Security, 2008, pp. 1–6.

[62] O. Postolache, J. Pereira, and P. Girão, “Smart sensors network for air quality monitoring applications,” IEEE Transactions on Instrumentation and Measurement, vol. 58, no. 9, pp. 3253–3262, 2009.

Sriharidevi V is a Research Scholar, PG and Research Department of Computer Applications. She did her M.Phil. in Computer Science with specialization in the area of Computer Network at Periyar University Distance Education. She has published one paper in international journal. Her research interests include wireless sensor network, Machine Learning.

Dr.R A Roseline is Associate Professor and Head, PG and Research Department of Computer Applications. She has 32 years of experienced in Higher education. She did her PhD in Computer Science in the area of Wireless Sensor Networks at Bharathiar University. She received M.Phil in Computer Science from Periyar University specializing in Mobile Adohc Networks. She has published about 30 papers in international journals. She has completed a UGC Minor Project in Pollution Monitoring using Sensors networks with funding of 1.7 lacs. She has presented papers in international conference at Australia and Malaysia