Huge amounts of data are circulating in the digital world in the era of the Industry 5.0 revolution. Machine learning is experiencing success in several sectors such as intelligent control, decision making, speech recognition, natural language processing, computer graphics, and computer vision, despite the requirement to analyze and interpret data. Due to their amazing performance, Deep Learning and Machine Learning Techniques have recently become extensively recognized and implemented by a variety of real-time engineering applications. Knowledge of machine learning is essential for designing automated and intelligent applications that can handle data in fields such as health, cyber-security, and intelligent transportation systems. There are a range of strategies in the field of machine learning, including reinforcement learning, semi-supervised, unsupervised, and supervised algorithms. This study provides a complete study of managing real-time engineering applications using machine learning, which will improve an application’s capabilities and intelligence. This work adds to the understanding of the applicability of various machine learning approaches in real-world applications such as cyber security, healthcare, and intelligent transportation systems. This study highlights the research objectives and obstacles that Machine Learning approaches encounter while managing real-world applications. This study will act as a reference point for both industry professionals and academics, and from a technical standpoint, it will serve as a benchmark for decision-makers on a range of application domains and real-world scenarios.

1. Introduction

1.1. Machine Learning Evolution. In this digital era, the data source is becoming part of many things around us, and digital recording [1, 2] is a normal routine that is creating bulk amounts of data from real-time engineering applications. This data can be unstructured, semi-structured, and structured. In a variety of domains, intelligent applications can be built using the insights extracted from this data. For example, as in [3] author used cyber-security data for extracting insights and use those insights for building intelligent application for cyber-security which is automated and driven by data. In the article [1], the author uses mobile data for extracting insights and uses those insights for building an intelligent smart application which is aware of context. Real-time engineering applications are based on tools and techniques for managing the data and having the capability for useful knowledge or insight extraction in an intelligent and timely fashion.
Machine Learning is a stream in Artificial Intelligence, which is gaining popularity in recent times in the field of computing and data analysis that will make applications behave intelligently [4]. In industry 4.0 (fourth industrial revolution) machine learning is considered one of the popular technologies which will allow the application to learn from experience, instead of programming specifically for the enhancement of the system [1, 3]. Traditional practices of industries and manufacturing are automated in Industry 4.0 [5] by using machine learning which is considered a smart technology and is used for exploratory data processing. So, machine learning algorithms are keys to developing intelligent real-time engineering applications for real-world problems by analyzing the data intelligently. All the machine learning techniques are categorized into the following types (a) Reinforcement Learning (b) Unsupervised Learning (c) Semi-Supervised Learning, and (d) Supervised Learning.

Based on the collected data from google trends [6], popularity of these techniques is represented in Figure 1. In Figure 1 the y-axis indicated the popularity score of the corresponding technique and the x-axis indicated the time period. As per Figure 1, the popularity score of the technique is growing day by day in recent times. Thus, it gives the motivation to perform this review on machine learning’s role in managing Real Time Engineering Applications. We may use Google Trends to find out what the most popular web subjects are at any given time and location. This could help us generate material and give us suggestions for articles that will most likely appeal to readers. Just make sure the content is relevant to our company or industry. We can look into the findings a little more carefully and investigate the reasons that may have influenced such trends because Google Trends can supply us with data about the specific regions in which our keywords drew substantial interest. With this level of data, we can figure out what’s working and what needs to be improved.

Machine learning algorithms’ performance and characteristics and nature of the data will decide the efficiency and effectiveness of the solution based on machine learning. The data-driven systems [7, 8] can be effectively built by using the following ML areas like reinforcement learning, association rule learning, reduction of dimensionality and feature engineering, data clustering, regression, and classification analysis. From ANN, a new technology is originated from the family of machine learning techniques called Deep Learning which is used for analyzing data intelligently [9].

Every machine learning algorithm’s purpose is different even various machine learning algorithms applied over the same category will generate different outcomes and depends on the nature and characteristics of data [10]. Hence, it’s challenging to select a learning algorithm for generating solutions to a target domain. Thus, there is a need for understanding the applicability and basic principle of ML algorithms in various Real Time Engineering Applications.

A comprehensive study on a variety of machine learning techniques is provided in this article based on the potentiality and importance of ML that can be used for the augmentation of application capability and intelligence. For industry people and academia this article will be acting as a reference manual, to research and study and build intelligent systems which are data-driven in a variety of real-time engineering applications on the basis of machine learning approaches.

1.2. Types of Machine Learning Techniques. Figure 2 shows the Machine Learning Timeline chart. There are 4 classes of machine learning approaches (a) Reinforcement Learning, (b) Semi-Supervised Learning (c) Unsupervised, and (d) Supervised Learning as shown in Figure 3. With the applicability of every ML technique in Real Time Engineering applications, we put down a brief discussion on all the four types of ML approaches as follows:

(i) Reinforcement learning: in an environment-driven approach, RL allows machines and software agents to assess the optimal behavior automatically to enhance the efficiency in a particular context [11]. Penalty or rewards are the basis for RL, and the goal of this approach is to perform actions that minimize the penalty and maximize the reward by using the extracted insights from the environment [12]. RL can be used for enhancing sophisticated systems efficiency by doing operational optimization or by using automation with the help of the trained Artificial Intelligence models like supply chain logistics, manufacturing, driving autonomous tasks, robotics, etc.

(ii) Semi-supervised: as this method operates on both unlabelled and labeled data [3, 7] it is considered a hybrid approach and lies between “with supervision” and “without supervision” learning approach. The author in [12] concludes that the semi-supervised approach is useful in real-time because of numerous amounts of unlabeled data and rare amounts of labeled data available in various contexts. The semi-supervised approach achieves the goal of predicting better when compared to predictions based on labeled data only. Text classification, labeling data, fraud detection, machine translation, etc., are some of the common tasks.

(iii) Unsupervised: as in [7], the author defines that unsupervised approach as a process of data-driven, with minimum or no human interface, it takes datasets consisting of unlabeled data and analyzes them. The unsupervised approach is widely used for purpose of exploring data, results grouping, identifying meaningful structures and trends, and extracting general features. Detecting anomalies, association rules finding, reducing dimensionality, learning features, estimating density, and clustering are the most usual unsupervised tasks.

(iv) Supervised: as in [7], author defines the supervised approach as a process of making a function to learn to map output from input. A function is inferred by using training example collection and training data which is labeled. As in [3], the author states that a supervised learning approach is a task-driven approach, which is to be initiated when certain inputs
Machine Learning is widespread into various domains. Examples include:

- Google AlphaGo program that beat an unhandicapped professional human player
- IBM Watson Defeats two human challengers at Jeopardy
- Andrew Ng & Jeff Dean create a NN that learns to recognize cats by watching unlabelled images
- Yim Kam Ho creates first algorithm for Random Decision forests to enable better prediction performance
- The launch of Kaggle-website for ML competitions
- An ANN, R language and WEKA released in these years
- Arthur Samuel, World first self-learning program developed
- James William Cooley & John Tukey Codeveloped first Fourier Transform

**Figure 1:** World wide trend analysis on machine learning techniques [6].

**Figure 2:** Machine learning time line.

**Figure 3:** Machine learning techniques.
are capable to accomplish a variety of goals. The most frequently used supervised learning tasks are regression and classification.

In Table 1, we summarize various types of machine learning techniques with examples.

Table 2 summarizes the comparison between the current survey with existing surveys and highlights how it is different or enhanced from the existing surveys.

1.3. Contributions. Following are the key contributions to this article:

(i) A comprehensive view on variety of ML algorithms is provided which is applicable to improve data-driven applications, task-driven applications capabilities, and intelligence

(ii) To discuss and review the applicability of various solutions based on ML to a variety of real-time engineering applications

(iii) By considering the data-driven application capabilities and characteristics and nature of the data, the proposed study/review scope is defined

(iv) Various challenges and research directions are summarized and highlighted that fall within this current study scope

1.4. Paper Organization. The organization of the rest of the article is as follows: state of art is presented in the next section which explains and introduces real-time engineering applications and machine learning; in the next section, ML’s role in real-time engineering applications is discussed; and in the coming section, challenges and lessons learned are presented; in the penultimate section, several future directions and potential research issues are discussed and highlighted; and in the final section conclude the comprehensive study on managing Real Time Engineering Applications using Machine Learning.

2. State of the Art

2.1. Real World Issues. Computer systems can utilize all client data through machine learning. It acts according to the program’s instructions while also adapting to new situations or changes. Algorithms adapt to data and exhibit previously unprogrammed behaviors. Acquiring the ability to read and recognize context enables a digital assistant to skim emails and extract vital information. This type of learning entails the capacity to forecast future customer behaviors. This enables you to have a deeper understanding of your customers and to be proactive rather than reactive. Machine learning is applicable to a wide variety of sectors and industries and has the potential to expand throughout time. Figure 4 represents the real-world applications of machine learning.

2.2. Introduction to Cyber Security. For both, services and information, internet is most extensively exploited. In article [13], author summarizes that since 2017 as an information source Internet is utilized by almost 48% of the whole population in the world. As concluded in the article [14], this number is hiked up to 82% in advanced countries.

The interconnection of distinct devices, networks, and computers is called the Internet, whose preliminary job is to transmit information from one device to another through a network. Internet usage spiked due to the innovations and advancements in mobile device networks and computer systems. As internet is the mostly used by the majority of population as an information source so it’s more prone to cyber criminals [15]. A computer system is said to be stable when it’s offering integrity, availability, and confidentiality of information. As stated in the article [16], with intent to disturb normal activity, if an unauthorized individual enters into the network, then the computer system will be compromised with integrity and security. User assets and cyberspace can be secured from unauthorized individual attacks and access with the help of cyber security. As in article [17], the primary goal of cyber security is to keep information available, integral, and confidential.

2.3. Introduction to Healthcare. With advancements in the field of Deep Learning/Machine Learning, there are a lot of transformations happening in the areas like governance, transportation, and manufacturing. Extensive research is going on in the field of Deep Learning over the last decade. Deep Learning has been applied to lots of areas that delivered a state-of-the-art performance in variety of domains like speech processing, text analytics, and computer vision. Recently researchers started deploying Deep Learning/Machine Learning approaches to healthcare [18], and they delivered outstanding performances in the jobs like brain tumor segmentation [19], image reconstruction in medical images [20, 21] lung nodule detection [22], lung disease classification [23], identification of body parts [24], etc.

It is evident that CAD systems that provide a second opinion will help the radiologists to confirm the disease [25] and deep learning/machine learning will further enhance the performance of these CAD systems and other systems that will provide supporting decisions to the radiologists [26].

Advancement in the technologies like big data, mobile communication, edge computing, and cloud computing is also helping the deployment of deep learning/machine learning models in the domain of healthcare applications [27]. By combining they can achieve greater predictive accuracies and an intelligent solution can be facilitated which is human-centered [28].

2.4. Introduction to Intelligent Transportation Systems. In transit and transportation systems, after the deployment of sensing technologies, communication, and information, the resultant implementation is called an intelligent transportation system [29]. An intelligent transportation system is an intrinsic part of smart cities [30], which have the following services such as autonomous vehicles, public transit system management, traveler information systems, and road traffic management. These services are expected to contribute a lot to the society by curbing pollution, enhancing energy efficiency,
transit and transportation efficiency is enhanced and finally, traffic and road safety is also improved.

Advances in technologies like wireless communication technology, computing, and sensing are enabling intelligent transportation systems applications and also bear a lot of challenges due to their capabilities to generate huge amounts of data, independent QoS requirements, and scalability.

Due to the recent traction in deep learning/machine learning models, approaches like RL and DL are utilized to exploit patterns and generate decisions and predictions accurately [31–33].

2.5. Introduction to Renewable Energy. Sustainable and alternative energy sources are in demand due to the effect created by burning fossil fuels in the environment and fossil fuel depletion. As in article [34], the energy market biomass, wind power, tidal waves, geothermal, solar thermal, and solar photovoltaic are growing as renewable energy resources. There will be instability in the power grids due to various reasons like when demand is more than the supply of the energy and when supply is more than the demand of the energy. Finally, environmental factors affect the energy output of the plants based on the renewable energy. To address the management and optimization of energy, machine learning is used.

2.6. Introduction to Smart Manufacturing. Manufacturing has been divided into a number of categories, one of the categories in which computer-based manufacturing is performed is called Smart Manufacturing, which performs workers’ training, digital technology, and quick changes in the design and with high adaptability. Other responsibilities include recyclability of production effectively, supply chain optimization, and demand-based quick changes in the levels of production. Enabling technologies of Smart Manufacturing are advances in robotics, services and devices connectivity in the industry, and processing capabilities in the big data.

2.7. Introduction to Smart Grid. The basic structure of the electrical power grid has remained same over time, and it has been noticed that it has become outdated and ill-suited, unable to meet demand and supply in the twenty-first century. Even though we are in the twenty-first century, electrical infrastructure has remained mostly unaltered throughout time. However, as the population and consumption have grown, so requires power.

2.7.1. Drawbacks

(i) Analyzing the demand is difficult
(ii) Response time is slow

The new smart grid idea has evolved to address the issues of the old outdated electrical power system. SG is a large energy network that employs real time and intelligent monitoring, communication, control, and self-healing technologies to provide customers with a variety of alternatives while guaranteeing the stability and security of their electricity supply. By definition, SGs are sophisticated cyber-physical system. The functionality of this modern SG can be broken down into four parts.

This contemporary SG’s functionality may be split down into four components:

(1) Consumption: electricity is used for a variety of reasons by various industries and inhabitants
(2) Distribution: the power so that it may be distributed more widely
(3) Transmission: electricity is transmitted over a high-voltage electronic infrastructure
(4) Generation: during this phase, electricity is generated in a variety of methods

ML and DL functionalities in the context of SG include predicting about

(1) Stability of the SG
(2) Optimum schedule
(3) Fraud detection
(4) Security breach detection
(5) Network anomaly detection
(6) Sizing
(7) Fault detection
(8) Energy consumption
(9) Price
(10) Energy generation

2.8. Introduction to Computer Networks. The usefulness of ML in networking is aided by key technological advancements in networking, such as network programmability via Software-Defined Networking (SDN). Though machine learning has been widely used to solve problems such as pattern recognition, speech synthesis, and outlier identification, its use in network operations and administration has
<table>
<thead>
<tr>
<th>S. no</th>
<th>Reference no.</th>
<th>Year</th>
<th>Citations</th>
<th>No. of references</th>
<th>Cybersecurity</th>
<th>Computer networks</th>
<th>Human machine interaction</th>
<th>Healthcare</th>
<th>Satellite imaging</th>
<th>Industry 4.0/5.0</th>
<th>Computer vision</th>
<th>Smart energy</th>
<th>Smart grids</th>
<th>Renewable energy systems</th>
<th>Electrical power systems</th>
<th>Civil engineering</th>
<th>Energy systems</th>
<th>Design and manufacturing</th>
<th>Techniques</th>
<th>Metrics</th>
<th>Tools</th>
<th>Trustworthiness</th>
<th>Adversarial ML</th>
</tr>
</thead>
<tbody>
<tr>
<td>1</td>
<td>372</td>
<td>2014</td>
<td>47</td>
<td>42</td>
<td>x</td>
<td>x</td>
<td>x</td>
<td>x</td>
<td>x</td>
<td>x</td>
<td>x</td>
<td>x</td>
<td>x</td>
<td>√</td>
<td>x</td>
<td>x</td>
<td>x</td>
<td>x</td>
<td>x</td>
<td>x</td>
<td>x</td>
<td>x</td>
<td>x</td>
</tr>
<tr>
<td>2</td>
<td>373</td>
<td>2020</td>
<td>1</td>
<td>54</td>
<td>x</td>
<td>x</td>
<td>x</td>
<td>x</td>
<td>x</td>
<td>x</td>
<td>x</td>
<td>x</td>
<td>x</td>
<td>x</td>
<td>√</td>
<td>x</td>
<td>x</td>
<td>x</td>
<td>x</td>
<td>x</td>
<td>x</td>
<td>x</td>
<td>x</td>
</tr>
<tr>
<td>3</td>
<td>374</td>
<td>2020</td>
<td>12</td>
<td>668</td>
<td>√</td>
<td>x</td>
<td>x</td>
<td>x</td>
<td>x</td>
<td>x</td>
<td>x</td>
<td>x</td>
<td>x</td>
<td>x</td>
<td>x</td>
<td>x</td>
<td>x</td>
<td>x</td>
<td>x</td>
<td>x</td>
<td>x</td>
<td>x</td>
<td>x</td>
</tr>
<tr>
<td>4</td>
<td>375</td>
<td>2020</td>
<td>17</td>
<td>1.29</td>
<td>x</td>
<td>x</td>
<td>x</td>
<td>x</td>
<td>x</td>
<td>x</td>
<td>x</td>
<td>x</td>
<td>x</td>
<td>x</td>
<td>x</td>
<td>x</td>
<td>x</td>
<td>x</td>
<td>x</td>
<td>x</td>
<td>x</td>
<td>x</td>
<td>x</td>
</tr>
<tr>
<td>5</td>
<td>376</td>
<td>2020</td>
<td>1</td>
<td>2.59</td>
<td>x</td>
<td>x</td>
<td>x</td>
<td>x</td>
<td>x</td>
<td>x</td>
<td>x</td>
<td>x</td>
<td>x</td>
<td>x</td>
<td>x</td>
<td>x</td>
<td>x</td>
<td>x</td>
<td>x</td>
<td>x</td>
<td>x</td>
<td>x</td>
<td>x</td>
</tr>
<tr>
<td>6</td>
<td>377</td>
<td>2021</td>
<td>10</td>
<td>1.40</td>
<td>√</td>
<td>√</td>
<td>√</td>
<td>√</td>
<td>x</td>
<td>x</td>
<td>x</td>
<td>x</td>
<td>x</td>
<td>x</td>
<td>x</td>
<td>x</td>
<td>x</td>
<td>x</td>
<td>x</td>
<td>x</td>
<td>x</td>
<td>x</td>
<td>x</td>
</tr>
<tr>
<td>7</td>
<td>378</td>
<td>2018</td>
<td>91</td>
<td>48</td>
<td>x</td>
<td>x</td>
<td>x</td>
<td>x</td>
<td>x</td>
<td>x</td>
<td>x</td>
<td>x</td>
<td>x</td>
<td>x</td>
<td>x</td>
<td>x</td>
<td>x</td>
<td>x</td>
<td>x</td>
<td>x</td>
<td>x</td>
<td>x</td>
<td>x</td>
</tr>
<tr>
<td>8</td>
<td>379</td>
<td>2020</td>
<td>6</td>
<td>14</td>
<td>x</td>
<td>x</td>
<td>x</td>
<td>x</td>
<td>x</td>
<td>x</td>
<td>x</td>
<td>x</td>
<td>x</td>
<td>x</td>
<td>x</td>
<td>x</td>
<td>x</td>
<td>x</td>
<td>x</td>
<td>x</td>
<td>x</td>
<td>x</td>
<td>x</td>
</tr>
<tr>
<td>9</td>
<td>380</td>
<td>2017</td>
<td>445</td>
<td>73</td>
<td>x</td>
<td>x</td>
<td>x</td>
<td>x</td>
<td>x</td>
<td>x</td>
<td>x</td>
<td>x</td>
<td>x</td>
<td>x</td>
<td>x</td>
<td>x</td>
<td>x</td>
<td>x</td>
<td>x</td>
<td>x</td>
<td>x</td>
<td>x</td>
<td>x</td>
</tr>
<tr>
<td>10</td>
<td>381</td>
<td>2018</td>
<td>40</td>
<td>235</td>
<td>x</td>
<td>x</td>
<td>x</td>
<td>x</td>
<td>x</td>
<td>x</td>
<td>x</td>
<td>x</td>
<td>x</td>
<td>x</td>
<td>x</td>
<td>x</td>
<td>x</td>
<td>x</td>
<td>x</td>
<td>x</td>
<td>x</td>
<td>x</td>
<td>x</td>
</tr>
<tr>
<td>11</td>
<td>382</td>
<td>2020</td>
<td>91</td>
<td>48</td>
<td>x</td>
<td>x</td>
<td>x</td>
<td>x</td>
<td>x</td>
<td>x</td>
<td>x</td>
<td>x</td>
<td>x</td>
<td>x</td>
<td>x</td>
<td>x</td>
<td>x</td>
<td>x</td>
<td>x</td>
<td>x</td>
<td>x</td>
<td>x</td>
<td>x</td>
</tr>
<tr>
<td>12</td>
<td>383</td>
<td>2020</td>
<td>2</td>
<td>62</td>
<td>x</td>
<td>x</td>
<td>x</td>
<td>x</td>
<td>x</td>
<td>x</td>
<td>x</td>
<td>x</td>
<td>x</td>
<td>x</td>
<td>x</td>
<td>x</td>
<td>x</td>
<td>x</td>
<td>x</td>
<td>x</td>
<td>x</td>
<td>x</td>
<td>x</td>
</tr>
<tr>
<td>13</td>
<td>384</td>
<td>2021</td>
<td>4</td>
<td>140</td>
<td>x</td>
<td>x</td>
<td>x</td>
<td>x</td>
<td>x</td>
<td>x</td>
<td>x</td>
<td>x</td>
<td>x</td>
<td>x</td>
<td>x</td>
<td>x</td>
<td>x</td>
<td>x</td>
<td>x</td>
<td>x</td>
<td>x</td>
<td>x</td>
<td>x</td>
</tr>
<tr>
<td>14</td>
<td>385</td>
<td>2019</td>
<td>17</td>
<td>53</td>
<td>3.0</td>
<td>x</td>
<td>x</td>
<td>x</td>
<td>x</td>
<td>x</td>
<td>x</td>
<td>x</td>
<td>x</td>
<td>x</td>
<td>x</td>
<td>x</td>
<td>x</td>
<td>x</td>
<td>x</td>
<td>x</td>
<td>x</td>
<td>x</td>
<td>x</td>
</tr>
<tr>
<td>15</td>
<td>386</td>
<td>2019</td>
<td>35</td>
<td>2.61</td>
<td>x</td>
<td>x</td>
<td>x</td>
<td>x</td>
<td>x</td>
<td>x</td>
<td>x</td>
<td>x</td>
<td>x</td>
<td>x</td>
<td>x</td>
<td>x</td>
<td>x</td>
<td>x</td>
<td>x</td>
<td>x</td>
<td>x</td>
<td>x</td>
<td>x</td>
</tr>
<tr>
<td>16</td>
<td>387</td>
<td>2020</td>
<td>16</td>
<td>1.50</td>
<td>x</td>
<td>x</td>
<td>x</td>
<td>x</td>
<td>x</td>
<td>x</td>
<td>x</td>
<td>x</td>
<td>x</td>
<td>x</td>
<td>x</td>
<td>x</td>
<td>x</td>
<td>x</td>
<td>x</td>
<td>x</td>
<td>x</td>
<td>x</td>
<td>x</td>
</tr>
<tr>
<td>17</td>
<td>388</td>
<td>2019</td>
<td>1</td>
<td>25</td>
<td>x</td>
<td>x</td>
<td>x</td>
<td>x</td>
<td>x</td>
<td>x</td>
<td>x</td>
<td>x</td>
<td>x</td>
<td>x</td>
<td>x</td>
<td>x</td>
<td>x</td>
<td>x</td>
<td>x</td>
<td>x</td>
<td>x</td>
<td>x</td>
<td>x</td>
</tr>
<tr>
<td>18</td>
<td>389</td>
<td>2018</td>
<td>470</td>
<td>1.21</td>
<td>x</td>
<td>x</td>
<td>x</td>
<td>x</td>
<td>x</td>
<td>x</td>
<td>x</td>
<td>x</td>
<td>x</td>
<td>x</td>
<td>x</td>
<td>x</td>
<td>x</td>
<td>x</td>
<td>x</td>
<td>x</td>
<td>x</td>
<td>x</td>
<td>x</td>
</tr>
<tr>
<td>19</td>
<td>390</td>
<td>2019</td>
<td>63</td>
<td>1.49</td>
<td>x</td>
<td>x</td>
<td>x</td>
<td>x</td>
<td>x</td>
<td>x</td>
<td>x</td>
<td>x</td>
<td>x</td>
<td>x</td>
<td>x</td>
<td>x</td>
<td>x</td>
<td>x</td>
<td>x</td>
<td>x</td>
<td>x</td>
<td>x</td>
<td>x</td>
</tr>
<tr>
<td>20</td>
<td>391</td>
<td>2017</td>
<td>202</td>
<td>2.52</td>
<td>x</td>
<td>x</td>
<td>x</td>
<td>x</td>
<td>x</td>
<td>x</td>
<td>x</td>
<td>x</td>
<td>x</td>
<td>x</td>
<td>x</td>
<td>x</td>
<td>x</td>
<td>x</td>
<td>x</td>
<td>x</td>
<td>x</td>
<td>x</td>
<td>x</td>
</tr>
<tr>
<td>21</td>
<td>392</td>
<td>1997</td>
<td>97</td>
<td>97</td>
<td>x</td>
<td>x</td>
<td>x</td>
<td>x</td>
<td>x</td>
<td>x</td>
<td>x</td>
<td>x</td>
<td>x</td>
<td>x</td>
<td>x</td>
<td>x</td>
<td>x</td>
<td>x</td>
<td>x</td>
<td>x</td>
<td>x</td>
<td>x</td>
<td>x</td>
</tr>
<tr>
<td>22</td>
<td>393</td>
<td>2019</td>
<td>148</td>
<td>114</td>
<td>x</td>
<td>x</td>
<td>x</td>
<td>x</td>
<td>x</td>
<td>x</td>
<td>x</td>
<td>x</td>
<td>x</td>
<td>x</td>
<td>x</td>
<td>x</td>
<td>x</td>
<td>x</td>
<td>x</td>
<td>x</td>
<td>x</td>
<td>x</td>
<td>x</td>
</tr>
<tr>
<td>23</td>
<td>394</td>
<td>2021</td>
<td>1</td>
<td>193</td>
<td>x</td>
<td>x</td>
<td>x</td>
<td>x</td>
<td>x</td>
<td>x</td>
<td>x</td>
<td>x</td>
<td>x</td>
<td>x</td>
<td>x</td>
<td>x</td>
<td>x</td>
<td>x</td>
<td>x</td>
<td>x</td>
<td>x</td>
<td>x</td>
<td>x</td>
</tr>
<tr>
<td>24</td>
<td>395</td>
<td>2018</td>
<td>83</td>
<td>79</td>
<td>√</td>
<td>x</td>
<td>x</td>
<td>x</td>
<td>x</td>
<td>x</td>
<td>x</td>
<td>x</td>
<td>x</td>
<td>x</td>
<td>x</td>
<td>x</td>
<td>x</td>
<td>x</td>
<td>x</td>
<td>x</td>
<td>x</td>
<td>x</td>
<td>x</td>
</tr>
<tr>
<td>25</td>
<td>396</td>
<td>2020</td>
<td>47</td>
<td>103</td>
<td>x</td>
<td>x</td>
<td>x</td>
<td>x</td>
<td>x</td>
<td>x</td>
<td>x</td>
<td>x</td>
<td>x</td>
<td>x</td>
<td>x</td>
<td>x</td>
<td>x</td>
<td>x</td>
<td>x</td>
<td>x</td>
<td>x</td>
<td>x</td>
<td>x</td>
</tr>
<tr>
<td>26</td>
<td>Proposed survey</td>
<td>2021</td>
<td>—</td>
<td>396</td>
<td>√</td>
<td>√</td>
<td>√</td>
<td>√</td>
<td>√</td>
<td>√</td>
<td>√</td>
<td>√</td>
<td>x</td>
<td>x</td>
<td>x</td>
<td>x</td>
<td>x</td>
<td>x</td>
<td>x</td>
<td>x</td>
<td>x</td>
<td>x</td>
<td>x</td>
</tr>
</tbody>
</table>

Table 2: Summary of important surveys on ML.
been limited. The biggest roadblocks are determining what data may be collected and what control actions can be taken on legacy network equipment. These issues are alleviated by the ability to program the network using SDN. ML-based cognition can be utilized to help automate network operation and administration chores. As a result, applying machine learning approaches to such broad and complicated networking challenges is both intriguing and challenging. As a result, ML in networking is a fascinating study area that necessitates a thorough understanding of ML techniques as well as networking issues.

2.9. Introduction to Energy Systems. A set of structured elements designed for the creation, control, and/or transformation of energy is known as an energy system [35, 36]. Mechanical, chemical, thermal, and electro-magnetical components may be combined in energy systems to span a wide variety of energy categories, including renewables and alternative energy sources [37–39]. The progress of energy systems faces difficult decision-making duties in order to meet a variety of demanding and conflicting objectives, such as functional performance, efficiency, financial burden, environmental effect, and so on [40]. The increasing use of data collectors in energy systems has resulted in an enormous quantity of data being collected. Smart sensors are increasingly widely employed in the production and consumption of energy [41–43]. Big data has produced a plethora of opportunities and problems for making well-informed decisions [44, 45]. The use of machine learning models has aided the deployment of big data technologies in a variety of applications [46–50]. Prediction approaches based on machine learning models have gained popularity in the energy sector [51–53] because they make it easier to infer functional relationships from observations. Because of their accuracy, effectiveness, and speed, ML models in energy systems are becoming crucial for predictive modeling of production, consumption, and demand analysis [54, 55]. In the context of complex human interactions, ML models provide insight into energy system functioning [56, 57]. The use of machine learning models is in making traditional energy systems, as well as alternative and renewable energy systems.

3. Recent Works on Real-Time Engineering Applications

3.1. Machine Learning for ITS. Exposure to traffic noise, air pollution, road injuries, and traffic delays are only some of the key issues that urban inhabitants experience on a daily basis. Urban areas are experiencing severe environmental and quality-of-life difficulties as a result of rapid car expansion, insufficient transportation infrastructure, and a lack of road safety rules. For example, in many urban areas, large trucks violate the typical highways, resulting in traffic congestion and delays. In addition, many bikers have frequent near misses as a result of their clothes, posture changes, partial occlusions, and varying observation angles all posing significant challenges to the Machine Learning (ML) algorithms’ detection rates.

Over the last decade, there has been a surge in interest in using machine learning and deep learning methods to analyze and visualize massive amounts of data generated from various sources in order to improve the classification and recognition of pedestrians, bicycles, special vehicles (e.g., emergency vehicles vs. heavy trucks), and License Plate Recognition (LPR) for a safer and more sustainable environment. Although deep models are capable of capturing a wide variety of appearances, adaption to the environment is essential.

Artificial neural networks form the base for deep learning success; in artificial neural networks to mirror an image, the human brain functioning interconnected node system sets are present. The neighboring layer’s nodes will be consisting of connections with weights coming from nodes from other layers. The output value is generated by given input and weight to the activation function in a node. Figure 5 presents the ML mainstream approaches used in ITS.

Figure 6 shows the RL working in intelligent transportation system.
Figures 7–9 present the interaction between ITS and ML and Machine Learning Pipeline.

3.2. Machine Learning for Healthcare. Over time, for the actions performed as a response reward, actions and observations are given as input to policy functions, and the method that learns from this policy function is called RL [58]. There is a wide range of healthcare applications where RL can be used even RL can be used in the detection of disease based on checking symptoms ubiquitously [59]. Another potential use of RL in this domain is Gogame [60].

In semi-supervised learning, both unlabeled data and labeled data are used for training particularly greater doses of unlabeled data and little doses of labeled data are available, and then semi-supervised learning is suitable. Semi-supervised learning can be applied to a variety of healthcare applications like medical image segmentation [61, 62] using various sensors recognition of activity is proposed in [61], in [63] author used semi-supervised learning for healthcare data clustering.

In supervised learning, labeled information is used for training the model to map the input to output. In the regression output value is continuous and in classification output value is discrete. Typical application of supervised learning in the healthcare domain is the identification of organs in the body using various image modalities [19] and nodule classification in the lung images [21].
In unsupervised learning, mapping of input to the output will be done by training the model using unlabeled data:

(i) Similarity is used for clustering
(ii) Feature selection/dimensionality reduction
(iii) Anomaly detection [64]

Unsupervised learning can be applied to a lot of healthcare applications like feature selection [65] using PCA and using Clustering [66] for heart disease prediction.

Various phases in an ML-based Healthcare system are shown in Figure 10.

The four major applications of healthcare that can benefit from ML/DL techniques are prognosis, diagnosis, treatment, and clinical workflow, which are described in Table 3.
3.3. Machine Learning for Cyber Security. Artificial Intelligence and Machine Learning are widely accepted and utilized in various fields like Cyber Security [94–103], design and manufacturing [104], medicine [105–108], education [109], and finance [110–112]. Machine Learning techniques are used widely in the following areas of cyber security: intrusion detection [113–116], dark web or deep web sites [117, 118], phishing [119–121], malware detection [122–125], fraud detection [126–129], and spam classification [130–133]. As time changes there is a need for vigorous and novel techniques to address the issues of cyber security. Machine Learning is suitable for evolutionary attacks as it learns from experiences.

In article [118], the authors analyzed and evaluated the dark web which is a hacker’s social network by using the ML approach for threat prediction in the cyberspace. In article [134], the author used an ML model with social network features for predicting cyberattacks on an organization during a stipulated period. This prediction uses a dataset consisting of darkweb’s 53 forum’s data in it. Advancements in recent areas can be found in [135–138].

Antivirus, firewalls, unified threat management [139], intrusion prevention system [140], and SEIM solutions [141] are some of the classical cyber security systems. As in article [142], the author concluded that, in terms of post-cyber-attack response, performance, and in error rate classical cyber security systems are poor when compared with AI-based systems. As in the article [143], once there is cyber-space damage by the attack then only it’s identified and this situation happens in almost 60%. Both on the cyber security side and attackers’ side, there is a stronger hold by ML. On the cyber security side as specified in this article [144, 145] to safeguard everything from the damage done by the attackers and for detecting attacks at an early stage and finally for performance enhancement ML is used. ML is used on the attacker’s end to locate weaknesses and system vulnerabilities as well as techniques to get beyond firewalls and other defence walls. As in [146], the author concludes that to

Table 3: Neural networks comparison.

<table>
<thead>
<tr>
<th>Machine learning applications in</th>
<th>Domain</th>
<th>Areas</th>
<th>References</th>
</tr>
</thead>
<tbody>
<tr>
<td>Prognosis</td>
<td></td>
<td>In clinical practice, the process of forecasting the disease development is called prognosis. Prognosis can predict various things like survival chances, corresponding disease related health issues detection, will the disease be stable, improve or become worse and finally, prognosis can forecast the symptoms of the disease</td>
<td>[67–69]</td>
</tr>
<tr>
<td>Diagnosis</td>
<td>Health records in electronic format</td>
<td>Patient’s full history of medication is generated by healthcare services and hospitals which is stored in terms of records in electronic formats. This generated data will be consisting of unstructured and structured data. For aiding the process of diagnosis ML-based methods are applied over electronic health records for fetching the features</td>
<td>[70–73]</td>
</tr>
<tr>
<td>Treatment</td>
<td>Analysis of image in the medical fields</td>
<td>Image retrieval, image registration, image reconstruction, image segmentation, image classification, image detection, image enhancement</td>
<td>[74–93]</td>
</tr>
<tr>
<td>Clinical workflow</td>
<td>Interpretation of an image</td>
<td>Radiologists are responsible for analyzing the images and the results of the study will be written in the radiology reports</td>
<td>[94–97]</td>
</tr>
</tbody>
</table>
further enhance the classification performance ML approaches are combined.


3.5. Machine Learning for Smart Manufacturing. The following table shows the ML applicability to Smart Manufacturing. State-of-art works are presented in Table 5.

3.6. Machine Learning for Smart Grids. This subsection discusses machine learning applicability to smart grids. State-of-the-art works are presented in Table 6.

3.7. Machine Learning for Computer Networks

3.7.1. Traffic Prediction. As networks are day by day becoming diverse and complex, it becomes difficult to manage and perform network operations so huge importance is given to traffic forecast in the network to properly manage and perform network operations. Time Series Forecasting is forecasting the traffic in near future.

3.7.2. Traffic Prediction. To manage and perform network operations, it’s quite important to perform classification of the network traffic which includes provisioning of the resource, monitoring of the performance, differentiation of the service and quality of service, intrusion detection and security, and finally capacity planning.

3.7.3. Congestion Control. In a network, excess packets will be throttled using the concept called congestion control. It makes sure the packet loss ratio is in an acceptable range, utilization of resources is at a fair level, and stability of the network is managed.

Table 7 presents ML state of art systems in networking.

3.7.4. Machine Learning for Civil Engineering. The first uses of ML programs in Civil Engineering involved testing different existing tools on simple programs [210–213], more difficult problems are addressed in [214–216].

3.7.5. Machine Learning for Energy Systems. Hybrid ML models, ensembles, Deep Learning, Decision Trees, ANFIS, WNN, SVM, ELM, MLP, ANN are among the ten key ML models often employed in energy systems, according to the approach.

Table 8 presents ML state of art systems in the Energy Systems domain.


While machine learning offers promise and is already proving beneficial to businesses around the world, it is not without its hurdles and issues. For instance, machine learning is useful for spotting patterns, but it performs poorly at generalizing knowledge. There is also the issue of “algorithm weariness” among users.

In ML, for model training, decent amount of data and resources that provide high performances are needed. This challenge is addressed by involving multiple GPU’s. In Real Time Engineering Applications, an ML approach is needed which is modeled to address a particular problem robustly. As the same model designed to address one task in real-time engineering application cannot address all the tasks in a variety of domains, so there is a need to design a model for each task in the Real Time Engineering Applications.

ML approaches should have the skill to prevent issues in the early stages as this is an important challenge to address in most real-time engineering applications. In the medical domain, ML can be used in predicting diseases and ML techniques can also be used for forecasting the detection of terrorism attacks. As in [243], the catastrophic consequences cannot be avoided by having faith blindly in the ML predictions. As in article [244], author states that ML approaches are used in various domains, but in some domains as an alternative to accuracy and speed ML approaches require correctness at very high levels. To convert a model into trustworthy, there is a need to avoid a shift in dataset, which means the model is to be trained and tested on the same dataset which can be ensured by avoiding data leakages [245].

Moving object’s location can be identified by using the enabling technologies like GPS and cell phones and this information to be maintained securely as tamper-proof is one of the crucial tasks for ML. As in article [246], author states that an object’s location information from multiple sources is compared and tries to find the similarity, and as in article [247] author confirms that due to network delays the location change of the objects there is always ambiguity in the location information gathered from multiple sources and the trustworthiness of such information needs to be addressed using ML techniques.

In a connected web system, to have interaction between consumers and service providers with trustworthiness an ontology of trust is proposed in the article [248]. In text classification also trustworthiness is used. As in article [249] author states that in semantic and practical terms where the meaning of the text is interpreted trustworthiness can be fused. In article [250] author validates the software’s trustworthiness using a metric model. As in article [251], the author states that in companies and data centers the consumption of power can be mitigated by utilizing ML approaches for designing strategies that are power-aware. To reduce the consumption in its entirety, it’s better to turn off the machines dynamically. Which machine to be turned off will be decided by the forecasting model and it’s very
important to have trust in this forecasting model before setting up the machine to be switched off. Fatigue in the alarm is generating false alarms at higher rates. This will reduce the response time of the security staff and this issue is an interesting area in cyber security [252, 253].

Some concerns associated with machine learning have substantial repercussions that are already manifesting now. One is the absence of explainability and interpretability, also known as the “black box problem.” Even its creators are unable to comprehend how machine learning models generate their own judgments and behaviors. This makes it

<table>
<thead>
<tr>
<th>Reference</th>
<th>Machine learning technique</th>
<th>Purpose</th>
</tr>
</thead>
<tbody>
<tr>
<td>[147]</td>
<td>Statistical machine learning techniques</td>
<td>Short and medium forecasting</td>
</tr>
<tr>
<td>[148]</td>
<td>Autoregressive integrated moving average and autoregressive moving average</td>
<td>Wind power forecasting and forecasting of wind speed</td>
</tr>
<tr>
<td>[149]</td>
<td>Kalman filter model is used</td>
<td>Wind-generated power and wind speed forecasting through online</td>
</tr>
<tr>
<td>[150]</td>
<td>Review on two machine learning techniques is done</td>
<td>Wind speed forecasting</td>
</tr>
<tr>
<td>[151]</td>
<td>ANN</td>
<td>TIME SERIES PREDICTION</td>
</tr>
<tr>
<td>[152]</td>
<td>ANN variant is used called recurrent multi-layer perceptron</td>
<td>Used for the prediction of long-term power generation</td>
</tr>
<tr>
<td>[153]</td>
<td>SVM is used</td>
<td>To measure the wind speed</td>
</tr>
<tr>
<td>[154]</td>
<td>Fuzzy models are used</td>
<td>For the prediction of wind power generation.</td>
</tr>
<tr>
<td>[155]</td>
<td>Numerical weather prediction model is used</td>
<td>Wind power consumption and generation forecast</td>
</tr>
<tr>
<td>[156]</td>
<td>Ensemble model is used</td>
<td>Wind power consumption and generation forecast</td>
</tr>
<tr>
<td>[157]</td>
<td>ANN and k-nearest neighbor approaches are used</td>
<td>Wind power generation forecast</td>
</tr>
<tr>
<td>[158]</td>
<td>Particle swarm optimization, k-NN and SVM are used</td>
<td>Wind power generation forecast</td>
</tr>
<tr>
<td>[159]</td>
<td>Techniques considered are (i) Random forest (ii) Regression trees (iii) ANN (iv) MLP (v) SVM</td>
<td>Review on machine learning techniques for wind power generation forecasting</td>
</tr>
</tbody>
</table>

<table>
<thead>
<tr>
<th>Reference</th>
<th>Machine learning technique</th>
<th>Purpose</th>
</tr>
</thead>
<tbody>
<tr>
<td>[160]</td>
<td>ANN is used</td>
<td>Solar energy generation</td>
</tr>
<tr>
<td>[161]</td>
<td>SVM is used</td>
<td>Power generation using solar</td>
</tr>
<tr>
<td>[162]</td>
<td>Ensemble method</td>
<td>Forecasting solar power generation</td>
</tr>
<tr>
<td>[163]</td>
<td>Statistical methods</td>
<td>Review on solar energy power generation</td>
</tr>
</tbody>
</table>

<table>
<thead>
<tr>
<th>Reference</th>
<th>Machine learning technique</th>
<th>Purpose</th>
</tr>
</thead>
<tbody>
<tr>
<td>[164–166]</td>
<td>RNN, SVM</td>
<td>Rainfall prediction</td>
</tr>
<tr>
<td>[167]</td>
<td>RNN, SVM</td>
<td>Forecast values of rainfall depth</td>
</tr>
<tr>
<td>[168]</td>
<td>Ensemble learning</td>
<td>Forecast the hydro energy consumption</td>
</tr>
<tr>
<td>[169–171]</td>
<td>ANN</td>
<td>Hydropower plant management</td>
</tr>
</tbody>
</table>

<table>
<thead>
<tr>
<th>Reference</th>
<th>Machine learning technique</th>
<th>Purpose</th>
</tr>
</thead>
<tbody>
<tr>
<td>[172]</td>
<td>SVM</td>
<td>To estimate the products, optimize cost</td>
</tr>
<tr>
<td>[173]</td>
<td>Least square SVM, backpropagation NN</td>
<td>Cost estimation of the product life-cycle</td>
</tr>
<tr>
<td>[174]</td>
<td>Genetic algorithm based ML techniques are used</td>
<td>For decision support systems</td>
</tr>
<tr>
<td>[175]</td>
<td>NN and ML techniques are used</td>
<td>Rotating equipment’s life is predicted</td>
</tr>
</tbody>
</table>

<table>
<thead>
<tr>
<th>Reference</th>
<th>Machine learning technique</th>
<th>Purpose</th>
</tr>
</thead>
<tbody>
<tr>
<td>[176]</td>
<td>Markov decision is used</td>
<td>For predicting plant operation and health management</td>
</tr>
</tbody>
</table>

<table>
<thead>
<tr>
<th>Reference</th>
<th>Machine learning technique</th>
<th>Purpose</th>
</tr>
</thead>
<tbody>
<tr>
<td>[177]</td>
<td>Various ML techniques are used</td>
<td>Discussed about hadoop framework and cloud computing utilization for managing data</td>
</tr>
<tr>
<td>[178]</td>
<td>Various ML techniques are used</td>
<td>How cloud can be incorporated into various product life cycles can be incorporated</td>
</tr>
</tbody>
</table>

<table>
<thead>
<tr>
<th>Reference</th>
<th>Machine learning technique</th>
<th>Purpose</th>
</tr>
</thead>
<tbody>
<tr>
<td>[179]</td>
<td>Various ML techniques are used</td>
<td>Hardware and software changes in the manufacturing can be tracked</td>
</tr>
</tbody>
</table>
difficult to correct faults and ensure that a model's output is accurate and impartial. When it was discovered that Apple’s credit card algorithm offered women much lesser credit lines than men, for instance, the corporation was unable to explain the reason or address the problem.

This pertains to the most serious problem affecting the field: data and algorithmic bias. Since the debut of the technology, machine learning models have been frequently and largely constructed using data that was obtained and labeled in a biased manner, sometimes on intentionally. It has been discovered that algorithms are frequently biased towards women, African Americans, and individuals of other ethnicities. Google’s DeepMind, one of the world’s leading AI labs, issued a warning that the technology poses a threat to queer individuals.

This issue is pervasive and well-known, yet there is resistance to taking the substantial action that many experts in the field insist is necessary. Timnit Gebru and Margaret Mitchell, co-leaders of Google’s ethical AI team, were sacked in retaliation for Gebru’s refusal to retract research on the dangers of deploying huge language models, according to tens of thousands of Google employees. In a survey of researchers, policymakers, and activists, the majority expressed concern that the progress of AI by 2030 will continue to prioritize profit maximization and societal control over ethics. The nation as a whole is currently debating and enacting AI-related legislation, particularly with relation to immediately and blatantly damaging applications, like facial recognition for law enforcement. These discussions will probably continue. And the evolving data

<table>
<thead>
<tr>
<th>Reference</th>
<th>DL/ML technique</th>
<th>Application</th>
</tr>
</thead>
<tbody>
<tr>
<td>[180]</td>
<td>Weighted regularized extreme learning machine</td>
<td>Prediction of wind speed can be improved</td>
</tr>
<tr>
<td>[181]</td>
<td>Fuzzy logic, NN</td>
<td>Predict solar irradiance</td>
</tr>
<tr>
<td>[182]</td>
<td>RFR, GBR, SVR</td>
<td>Solar radiation can be forecasted</td>
</tr>
<tr>
<td>[183]</td>
<td>Gradient boost, random forest, regression tree</td>
<td>Solar irradiance is forecasted</td>
</tr>
<tr>
<td>[184]</td>
<td>RNN</td>
<td>Predict power and wind speed</td>
</tr>
</tbody>
</table>

<table>
<thead>
<tr>
<th>Reference</th>
<th>ML techniques</th>
<th>Applications</th>
</tr>
</thead>
<tbody>
<tr>
<td>[191, 192]</td>
<td>Supervised: MLP-NN</td>
<td>Prediction of network traffic</td>
</tr>
<tr>
<td>[193, 194]</td>
<td>Supervised: KBR, LSTM-RNN, MLP-NN</td>
<td>Prediction of traffic volume</td>
</tr>
</tbody>
</table>

<table>
<thead>
<tr>
<th>Reference</th>
<th>ML techniques</th>
<th>Applications</th>
</tr>
</thead>
<tbody>
<tr>
<td>[195–197]</td>
<td>Supervised: SVM</td>
<td>Classification of traffic based on host behavior</td>
</tr>
<tr>
<td>[198]</td>
<td>Unsupervised: HCA</td>
<td>Classification of traffic based on host behavior</td>
</tr>
<tr>
<td>[199]</td>
<td>Supervised: AdaBoost</td>
<td>Classification of traffic based on host behavior</td>
</tr>
<tr>
<td>[200–202]</td>
<td>Supervised k-NN, NBKE, BAGGING</td>
<td>Supervised flow feature based traffic classification</td>
</tr>
<tr>
<td>[203–205]</td>
<td>Unsupervised DBSCAN, AutoClass,k-means</td>
<td>UnSupervised flow feature based traffic classification</td>
</tr>
<tr>
<td>[206]</td>
<td>Supervised k-NN,Linear-SVM, Radial-SVM, DT, RF, extended tree, AdaBoost, Gradient-AdaBoost, NB, MLP</td>
<td>NFVand SDN-based traffic classification</td>
</tr>
</tbody>
</table>

<table>
<thead>
<tr>
<th>Reference</th>
<th>ML techniques</th>
<th>Applications</th>
</tr>
</thead>
</table>

<table>
<thead>
<tr>
<th>Table 6: ML state-of-the-art systems in smart grids domain.</th>
</tr>
</thead>
<tbody>
<tr>
<td>Reference</td>
</tr>
<tr>
<td>-----------</td>
</tr>
<tr>
<td>Energy forecasting</td>
</tr>
<tr>
<td>[180]</td>
</tr>
<tr>
<td>[181]</td>
</tr>
<tr>
<td>[182]</td>
</tr>
<tr>
<td>[183]</td>
</tr>
<tr>
<td>[184]</td>
</tr>
</tbody>
</table>

<table>
<thead>
<tr>
<th>Securing smart grids</th>
<th></th>
<th></th>
</tr>
</thead>
<tbody>
<tr>
<td>[185]</td>
<td>SVM</td>
<td>Stealthy attacks can be detected</td>
</tr>
<tr>
<td>[186]</td>
<td>ANN</td>
<td>Consumption of energy can be analyzed</td>
</tr>
<tr>
<td>[187]</td>
<td>DCNN</td>
<td>Theft of electricity can be detected by analyzing data</td>
</tr>
<tr>
<td>[188]</td>
<td>RNN</td>
<td>Smart grids’ false attacks can be identified</td>
</tr>
<tr>
<td>[189]</td>
<td>RNN</td>
<td>Attacks on network and frauds in the networks based on blockchain can be identified</td>
</tr>
<tr>
<td>[190]</td>
<td>Kalman filter, chi-square detector, and cosine similarity matching</td>
<td>Attacks on communication system can be identified</td>
</tr>
</tbody>
</table>

<table>
<thead>
<tr>
<th>Table 7: ML state-of-the-art systems in computer networking domain.</th>
</tr>
</thead>
<tbody>
<tr>
<td>Reference</td>
</tr>
<tr>
<td>-----------</td>
</tr>
<tr>
<td>Traffic prediction</td>
</tr>
<tr>
<td>[191, 192]</td>
</tr>
<tr>
<td>[193, 194]</td>
</tr>
</tbody>
</table>

<table>
<thead>
<tr>
<th>Traffic classification</th>
<th></th>
<th></th>
</tr>
</thead>
<tbody>
<tr>
<td>[195–197]</td>
<td>Supervised: SVM</td>
<td>Classification of traffic based on host behavior</td>
</tr>
<tr>
<td>[198]</td>
<td>Unsupervised: HCA</td>
<td>Classification of traffic based on host behavior</td>
</tr>
<tr>
<td>[199]</td>
<td>Supervised: AdaBoost</td>
<td>Classification of traffic based on host behavior</td>
</tr>
<tr>
<td>[200–202]</td>
<td>Supervised k-NN, NBKE, BAGGING</td>
<td>Supervised flow feature based traffic classification</td>
</tr>
<tr>
<td>[203–205]</td>
<td>Unsupervised DBSCAN, AutoClass,k-means</td>
<td>UnSupervised flow feature based traffic classification</td>
</tr>
<tr>
<td>[206]</td>
<td>Supervised k-NN,Linear-SVM, Radial-SVM, DT, RF, extended tree, AdaBoost, Gradient-AdaBoost, NB, MLP</td>
<td>NFVand SDN-based traffic classification</td>
</tr>
</tbody>
</table>

privacy rules will soon influence data collecting and, by extension, machine learning.

5. Machine Learning Applications

Because of its ability to make intelligent decisions and its potential to learn from the past, machine learning techniques are more popular in industry 4.0.

Here we discuss and summarize various machine learning techniques application areas.

5.1. Intelligent Decision-Making and Predictive Analytics.

By making use of data-driven predictive analytics, intelligent decisions are made by applying machine learning techniques [2, 254]. To predict the unknown outcomes by relying on the earlier events by exploiting and capturing the relationship between the predicted variables and explanatory variables is the basis for predictive analytics [7], for example, credit card fraud identification and criminal identification after a crime. In the retail industry, predictive analytics and intelligent decision-making can be used for out-of-stock situation avoidance, inventory management, behavior, and preferences of the consumer are better understood and logistics and warehouse are optimized. Support Vector Machines, Decision Trees, and ANN are the most widely used techniques in the above areas [8, 10]. Predicting the outcome accurately can help every organization like social networking, transportation, sales and marketing, healthcare, financial services, banking services, telecommunication, e-commerce, industries, etc., to improve.

5.2. Cyber-Security and Threat Intelligence.

Protecting data, hardware, systems, and networks is the responsibility of cyber-security and this is an important area in Industry 4.0 [5]. In cyber-security, one of the crucial technologies is machine learning which provides protection by securing cloud data, while browsing keeps people safe, foreseen the bad people online, insider threats are identified and malware is detected in the traffic. Machine learning classification models [255], deep learning-based security models [9], and association rule learning techniques [3] are used in cyber-security and threat intelligence.

5.3. Smart Cities.

In IoT, all objects are converted into things by equipping objects with transmitting capabilities for transferring the information and performing jobs with no user intervention.

Some of the applications of IoT are business, healthcare, agriculture, retail, transportation, communication, education, smart home, smart governance [2], and smart cities [256, 257]. Machine learning has become a crucial technology in the internet of things because of its ability to analyze the data and predict future events [1]. For instance, congestion can be predicted in smart cities, take decisions based on the surroundings knowledge, energy estimation for a particular period, and predicting parking availability.

---

Table 8: ML state-of-the-art systems in energy systems domain.

<table>
<thead>
<tr>
<th>ML technique</th>
<th>Reference</th>
<th>Application</th>
</tr>
</thead>
<tbody>
<tr>
<td>ANN</td>
<td>[217]</td>
<td>In the buildings of industry savings of the energy are verified and measured</td>
</tr>
<tr>
<td></td>
<td>[218]</td>
<td>Forecast solar radiation and predict wind speed</td>
</tr>
<tr>
<td></td>
<td>[219]</td>
<td>Electricity cost is forecasted</td>
</tr>
<tr>
<td></td>
<td>[220]</td>
<td>Power generation plans are created and scheduled and fluctuations in the wind power are controlled</td>
</tr>
<tr>
<td></td>
<td>[221]</td>
<td>Various capacities of the renewable energy generation are optimized</td>
</tr>
<tr>
<td>MLP</td>
<td>[222]</td>
<td>Plants are ranked</td>
</tr>
<tr>
<td></td>
<td>[223]</td>
<td>Forecast solar radiation</td>
</tr>
<tr>
<td></td>
<td>[224]</td>
<td>Predict solar power generation</td>
</tr>
<tr>
<td></td>
<td>[225]</td>
<td>Predict load</td>
</tr>
<tr>
<td></td>
<td>[226]</td>
<td>Solar irradiation is forecasted</td>
</tr>
<tr>
<td>SVM</td>
<td>[227]</td>
<td>Forecast price of the electricity in the market</td>
</tr>
<tr>
<td></td>
<td>[228]</td>
<td>Estimate the power quality</td>
</tr>
<tr>
<td></td>
<td>[229]</td>
<td>Disturbances in the power quality will be classified</td>
</tr>
<tr>
<td>WNN</td>
<td>[230]</td>
<td>Time series forecast</td>
</tr>
<tr>
<td></td>
<td>[231]</td>
<td>Predict the speed of the wind</td>
</tr>
<tr>
<td></td>
<td>[232]</td>
<td>In forecasting the wind power fluctuations can be mitigated</td>
</tr>
<tr>
<td>ANFIS</td>
<td>[233]</td>
<td>A protection system is presented</td>
</tr>
<tr>
<td></td>
<td>[234]</td>
<td>Demand of power is forecasted</td>
</tr>
<tr>
<td></td>
<td>[235]</td>
<td>Solar radiation is forecasted</td>
</tr>
<tr>
<td>Decision tree</td>
<td>[236]</td>
<td>Blackout risk is forecasted</td>
</tr>
<tr>
<td></td>
<td>[237]</td>
<td>Cost minimization in energy systems</td>
</tr>
<tr>
<td>Deep learning</td>
<td>[238]</td>
<td>Estimation of state-of-charge of battery</td>
</tr>
<tr>
<td></td>
<td>[239]</td>
<td>Predicting the electricity demand in the households</td>
</tr>
<tr>
<td></td>
<td>[240]</td>
<td>Energy consumption is forecasted</td>
</tr>
<tr>
<td>Ensemble model</td>
<td>[241]</td>
<td>Building electricity demand is forecasted</td>
</tr>
<tr>
<td></td>
<td>[242]</td>
<td>Predict buildings cooling load</td>
</tr>
</tbody>
</table>
5.4. Transportation and Traffic Prediction. Generally, transportation networks have been an important part of every country's economy. Yet, numerouscities across the world are witnessing an enormous amount of traffic volume, leading to severe difficulties such as a decrease in the quality of life in modern society, crises, accidents, CO₂ pollution increased, higher fuel prices, traffic congestion, and delays [258]. As a result, an ITS, that predicts traffic and is critical, and it is an essential component of the smart city. Absolute forecasting of traffic based on deep learning and machine learning models can assist to mitigate problems [259–261]. For instance, machine learning may aid transportation firms in identifying potential difficulties that may arise on certain routes and advising that their clients choose an alternative way based on their history of travel and pattern of travel by taking variety of routes. Finally, by predicting and visualizing future changes, these solutions will assist to optimize flow of the traffic, enhance the use and effectiveness of sustainable forms of transportation, and reduce real-world disturbance.

5.5. Healthcare and COVID-19 Pandemic. In a variety of medical-related application areas, like prediction of illness, extraction of medical information, data regularity identification, management of patient data, and so on, machine learning may assist address diagnostic and prognostic issues [262–264]. Here in this article [265], coronavirus is considered as an infectious disease by the WHO. Learning approaches have recently been prominent in the fight against COVID-19 [266, 267]. Learning approaches are being utilized to categorize the death rate, patients at high risk, and various abnormalities in the COVID-19 pandemic [268]. It may be utilized to fully comprehend the virus's origins, predict the outbreak of COVID-19, and diagnose and treat the disease [269, 270]. Researchers may use machine learning to predict where and when COVID-19 will spread, and then inform those locations to make the necessary preparations. For COVID 19 pandemic [271–273], to address the medical image processing problems, deep learning can provide better solutions. Altogether, deep and machine learning approaches can aid in the battle against the COVID-19 virus and pandemic, and perhaps even the development of intelligent clinical judgments in the healthcare arena.

5.6. Product Recommendations in E-commerce. One of the most prominent areas in e-commerce where machine learning techniques are used is suggesting products to the users of the e-commerce. Technology of machine learning can help e-commerce websites to analyze their customers' purchase histories and provide personalized product recommendations based on their behavior and preferences for their next purchase. By monitoring browsing tendencies and click-through rates of certain goods, e-commerce businesses, for example, may simply place product suggestions and offers. Most merchants, such as flipkart [274] and amazon, can avert out-of-stock problems, manage better inventory, optimize storage, and optimize transportation by using machine learning-based predictive models. Future of marketing and sales is to improve the personalized experience of the users while purchasing the products by collecting their data and analyzing the data and use them to improve the experience of the users. In addition, to attract new customers and also to retain the existing ones the e-commerce website will build packages to attract the customers and keep the existing ones.

5.7. Sentiment Analysis and NLP (Natural Language Processing). An act of using a computer system to read and comprehend spoken or written language [275] is called Natural Language Processing. Thus, NLP aids computers in reading texts, hearing speech, interpreting it, analyzing sentiment, and determining which elements are important, all of which may be done using machine learning techniques. Some of the examples of NLP are machine translation, language translation, document description, speech recognition, chatbot, and virtual personal assistant. Collecting data and generating views and mood of the public from news, forums, social media, reviews, and blogs is the responsibility of sentiment analysis [276] which is a subfield of NLP. In sentiment analysis, texts are analyzed by using machine learning tasks to identify the polarity like neutral, negative and positive and emotions like not interested, have interest, angry, very sad, sad, happy, and very happy.

5.8. Image, Speech and Pattern Recognition. Machine Learning is widely used in the image recognition [277] whose task is to detect the object in an image. Some of the instances of image recognition are social media suggestions tagging, face detection, character recognition and cancer label on an X-ray image. Alexa, Siri, Cortana, Google Assistant etc., are the famous linguistic and sound models in speech recognition [286,282]. The automatic detection of patterns and data regularities, such as picture analysis, is characterized as pattern recognition [278]. Several machine learning approaches are employed in this field, including classification, feature selection, clustering, and sequence labeling.

5.9. Sustainable Agriculture. Agriculture is necessary for all human activities to survive [279]. Sustainable agriculture techniques increase agricultural output while decreasing negative environmental consequences [279–281]. In articles [281–284] authors convey those emerging technologies like mobile devices, mobile technologies, Internet of Things can be used to capture the huge amounts of data to encourage the adoption of practices of sustainable agriculture by encouraging knowledge transfer among farmers. By using technologies, skills, information knowledge-intensive supply chains are developed in sustainable agriculture. Various techniques of machine learning can be applied in processing phase of the agriculture, production phase and preprocessing phase, distribution phases like consumer analysis, inventory management, production planning, demand
estimation of livestock, soil nutrient management, weed
detection, disease detection, weather prediction, irrigation
requirements, soil properties, and crop yield prediction.

5.10. Context-Aware and Analytics of User Behavior.
Capturing information or knowledge about the surrounding
is called context-awareness and tunes the behaviors of the
system accordingly [285, 286]. Hardware and software are
used in context-aware computing for automating the col-
clection and interpreting of the data.

From the historical data [1, 287] machine learning will
derive knowledge by using their learning capabilities which
is used for bringing tremendous changes in the mobile app
development environment.

Smart apps can be developed by the programmers, using
which uses can be entertained, support is provided for the
user and human behavior is understood [288–290] and can
build a variety of context-aware systems based on data-
driven approaches like context-aware smart searching, smart
interruption management, smart mobile recommendation,
etc., for instance, as in [291] phone call app can be created by
using association rules with context awareness. Clustering
approaches [292] are used and classification methods
[10, 293, 294] are used for predicting future events and for
capturing users’ behavior.

6. Challenges and Future Research Directions

In this review, quite a few research issues are raised by
studying the applicability of variety of ML approaches in the
analysis of applications and intelligent data. Here, oppor-
tunities in research and potential future directions are
summarized and discussed.

Research directions are summarized as follows:

(i) While dealing with real-world data, there is a need
for focusing on the in-detail study of the capturing
techniques of data

(ii) There is a huge requirement for fine-tuning the
preprocessing techniques or to have novel data
preprocessing techniques to deal with real-world
data associated with application domains

(iii) Identifying the appropriate machine learning
technique for the target application is also one of the
research interests

(iv) There is a huge interest in the academia in existing
machine learning hybrid algorithms enhancement
or modification and also in proposing novel hybrid
algorithms for their applicability to the target ap-
plications domain

Machine learning techniques’ performance over the data
and the data’s nature and characteristics will decide the
efficiency and effectiveness of the machine learning solu-
tions. Data collection in various application domains like
agriculture, healthcare, cyber-security etc., is complicated
because of the generation of huge amounts of data in very
less time by these application domains. To proceed further
in the analysis of the data in machine learning-based
applications relevant data collection is the key factor. So,
while dealing with real-world data, there is a need for fo-
cusing on the more deep investigation of the data collection
methods.

There may be many outliers, missing values, and am-
biguous values in the data that is already existing which will
impact the machine learning algorithms training. Thus,
there is a requirement for the cleansing of collected data
from variety of sources which is a difficult task. So, there is a
need for preprocessing methods to be fine-tuned and novel
preprocessing techniques to be proposed that can make
machine learning algorithms to be used effectively.

Choosing an appropriate machine learning algorithm
best suited for the target application, for the extraction
insights, and for analyzing the data is a challenging task,
because the characteristics and nature of the data may have
an impact on the outcome of the different machine learning
techniques [10]. Inappropriate machine learning algorithm
will generate unforeseen results which might reduce the
accuracy and effectiveness of the model. For this purpose,
the focus is on hybrid models, and these models are fine-
tuned for the target application domains or novel techniques
are to be proposed.

Machine learning algorithms and the nature of the data
will decide the ultimate success of the applications and their
Corresponding machine learning-based solutions. Machine
Learning models will generate less accuracy and become
useless when the data is the insufficient quantity for training,
irrelevant features, poor quality, and non-representative and
bad data to learn. For an intelligent application to be built,
there are two important factors i.e., various learning tech-
niques handling and effective processing of data.

Our research into machine learning algorithms for in-
telligent data analysis and applications raises a number of
new research questions in the field. As a result, we highlight
the issues addressed, as well as prospective research possi-
bilities and future initiatives, in this section.

The nature and qualities of the data, as well as the
performance of the learning algorithms, determine the ef-
ectiveness and efficiency of a machine learning-based so-
lution. To gather information in a specific domain, such as
cyber security, IoT, healthcare, agriculture, and so on. As a
result, data for the target machine learning-based applica-
tions is collected. When working with real-world data, a
thorough analysis of data collection methodologies is re-
quired. Furthermore, historical data may contain a large
number of unclear values, missing values, outliers, and data
that has no meaning.

Many machine learning algorithms exist to analyze data
and extract insights; however, the ultimate success of a
machine learning-based solution and its accompanying
applications is largely dependent on both the data and the
learning algorithms. Produce reduced accuracy if the data is
bad to learn, such as non-representative, poor-quality, ir-
relevant features, or insufficient amount for training. As a
result, establishing a machine learning-based solution and
eventually building intelligent applications, correctly pro-
cessing the data, and handling the various learning algo-
rithms is critical.
7. Conclusion and Future Scope

In this study on machine learning algorithms, a comprehensive review is conducted for applications and intelligent analysis of data. Here, the real-world issues and how solutions are prepared by using a variety of learning algorithms are discussed briefly. Machine Learning techniques’ performance and characteristics of the data will decide the machine learning model’s success. To generate intelligent decision-making, machine learning algorithms need to be acquainted with target application knowledge and trained with data collected from various real-world situations. For highlighting the applicability of ML approaches to variety of issues in the real world and variety of application areas are discussed in this review. At last, research directions and other challenges are discussed and summarized. All the challenges in the target applications domain must be addressed by using solutions effectively. For both industry professionals and academia, this study will serve as a reference point and from the technical perspective, this study also works as a benchmark for the decision makers on a variety of application domains and various real-world situations. Machine Learning’s application is not restricted to any one sector. Rather, it is spreading across a wide range of industries, including banking and finance, information technology, media and entertainment, gaming, and the automobile sector. Because the breadth of Machine Learning is so broad, there are several areas where academics are trying to revolutionize the world in the future.

Data Availability

No data were used to support this study.

Conflicts of Interest

The authors declare that they have no conflicts of interest to report regarding the present study.

Authors’ Contributions

All authors substantially contributed to the conception and design of the article and interpretation of the relevant literature.

References


M. Lei, L. Shiyian, J. Chuanwen, L. Hongling, and Z. Yan, "A review on the forecasting of wind speed and generated


consumption,” *Sustainable Energy, Grids and Networks*, vol. 6, pp. 91–99, 2016.


