International Transactions on Artificial Intelligence (ITALIC)
Vol. 1 No. 2, May 2023, pp. 213~229
P-ISSN: 2963-6086, E-ISSN: 2963-1939, DOI: 10.34306

An Overview of Concepts, Applications, Difficulties, and Unresolved Issues in Fog Computing and Machine Learning

Oscar Jayanagara[1], Dewi Sri Surya Wuisan[2]
[1]University of Pelita Harapan, Tangerang, Indonesia
[1]University of Pelita Harapan, Tangerang, Indonesia
e-mail: oscar.fe@uph.edu, dewi.wuisan@uph.edu

Article Info

Article history:

Received March 31, 2023
Revised April 13, 2023
Accepted April 14, 2023

Keywords:

Big Data
Machine Learning
Fog Computing
Applications
Internet of Things

ABSTRACT

Numerous fog computing apps and services are emerging as a result of the large volumes of data produced by systems based on fog computing. Additionally, the crucial field of machine learning (ML), which has made significant advancements in several academic areas, incorporating speech recognition, robotics, and neuromorphic computing, in addition to computer graphics and natural language processing (NLP).. The aim of current research provided insight into providing a list of the fog computing-related ML operations. The security, capacity, and latency standards for networks must be met by many IoT applications. Cloud computing does not, however, satisfy these needs. Today's technology can satisfy these objectives, and edge computing is one such option. The model enables traffic analysis and sensor data analysis. The management of resources, accuracy, and security are three areas of fog computing that we highlight in this thorough assessment of the most recent advancements in ML approaches. Additionally highlighted is the function of ML in edge computing. Additional viewpoints on the ML domain are presented, including those on the different kinds of application support, techniques, and datasets. Finally, open questions and research difficulties are highlighted.

Corresponding Author:

Oscar Jayanagara
University of Pelita Harapan, Tangerang, Indonesia
Email: oscar.fe@uph.edu

1. INTRODUCTION

As the digital era dawns, more businesses and people are using computers and other sophisticated gadgets on a regular basis [1]. Utilizing software and sensors, electronic devices are utilized to produce data [2]. As a result, a lot of businesses must take on the duty of regularly keeping enormous volumes of data [3]. The move in cloud computing,

providing advantages in terms of features that are pay-per-use, scalable, and accessible, has resulted in the need for a flexible infrastructure for information technologies enterprises [4]. The terms "platform as a service," "software as a service," and "infrastructure as a service" are a few examples of popular services that have been available thanks to cloud computing and are heading in the direction of "Anything" as a Service [5]. The cloud cannot transfer or process all massive data produced by sensors, though [6]. Additionally, a number of Internet of Things (IoT) apps need quicker processing than the cloud can currently provide, therefore these applications will not be able to run [7]. The fog computing paradigm, which makes use of idle computing capacity from devices near users to enable storage, networking, and processing at the edge, is used to overcome this issue [8]. Data that serves many uses, such as processing, analysis, and storage, must be sent to the cloud in order to accomplish various goals achieved through fog computing; benefits include increased effectiveness and smaller data sizes [9]. Although security and regulatory concerns are sometimes cited as justifications, this is frequently done for performance-related reasons [10]. IoT data analysis techniques have just started using AI algorithms [11].

An unreliable low-bandwidth connection network, insufficient onboard memory, poor power to process, as well as diverse hardware are only a few of the unfavorable characteristics that a low-layer network demonstrates [12]. These characteristics are different from those of cloud architecture [13]. Machine learning (ML) is considered to be by far widely used. Artificial Intelligence algorithms were utilized in several ways, and computer technologies in several fields, such as advances in hardware, cloud computing, GPU computing, and artificial intelligence (AI), have evolved in the past ten years [14]. The application of ML to address networking issues including resource allocation, routing, security, and traffic engineering has been explored in a number of earlier works [15]. As a result, ML is an important technology in autonomous intelligent/smart environments in terms of administration and operation [16].

Furthermore, ML is important for IoT since without ML, IoT cannot be used for functional, monitoring, or preprocessing tasks like routing or anomaly detection [17]. For the distribution and implementation of IoT applications, it is crucial to consider fog, cloud, and edge computing in the context of ML [18]. Weka, Scikit-learn, and high-level fog nodes, however, are distinct utilizable libraries and frameworks to develop numerous applications for Artificial Intelligence [19]. Using modern technology, such as IOS XR from Cisco, it is simple to incorporate competences for data analysis discovered on networks such as switches and routers. sensors, low-level fog nodes, and actuators were used as a framework for an inquiry into machine learning (ML) [20]. Data aggregation, the use of duty-cycle scheduling, medium access control, clustering, and routing are a few functional tasks that The use of machine learning carry out and enhance (MAC) [21].

Since the bulk of these activities are rapidly developing into dynamic, diverse, and complex structures, managing important processes at fog nodes is challenging [22]. Additionally, in order to attract more users, fog node services need to be more diverse and effective [23]. ML can be advantageous for fog computing (node-server) in a variety of ways because it has been effectively applied to the fog computing paradigm in many earlier sorts of study [24]. For instance, by combining ML with fog computing, users can benefit from deep analytics [25]. ML can offer workable solutions, which allows for the development of effective intelligent fog computing applications [26]. The information and features concealed in acquired data can be mined using these technologies. The following lists the main contributions made in this work.

- We look at studies that used machine learning (ML) to tackle the resource management, accuracy, and security issues associated with the fog computing paradigm.
- We emphasize the use of machine learning (ML) in various edge computing applications.
- Managing resources, being accurate, and maintaining security are some of the difficulties and unresolved problems in fog computing.

2. THE COMPREHENSIVE THEORETICAL BASIS

When it comes to implementing the fog computing concept, which allows the cloud platform to be stretched and brought closer to end users' devices, Cisco is regarded as a pioneer in the field [27]. This method addresses a number of problems, including latency sensitivity, support for geographical dispersion, and IoT applications that are cognizant of quality of service (QoS) [28]. Through the use of network edges as a backup for computing resources, fog computing is a unique extension of the cloud platform with a new paradigm architecture [29]. Fog computing offers application services and data storage, much like the cloud platform [30]. Table 1 lists the various characteristics of a fog system.

Table 1. Characteristics of Fog Computing

Feature	**Description**
Additional Features	Affordable, adaptable, and portable hardware and software deployment is available
Platform	Has a highly virtualized platform.
Operation Areas	Operates locally (from one hop a fog node with a device).
Storage and Services	Owns its own telecommunications, networking, and computer services.
Capability	Supports a wide range due to its capacity to respond immediately, it has a variety of industrial applications.
Heterogeneity	Having extensive and diverse end-user support and situated at the network's edge.

A fog system can be distinguished from cloud computing by a variety of factors in addition to those stated in Table 1. Each feature has pros and cons. Several well-liked elements are shown in Table 2.

Table 2. Characteristics of Fog Computing

Aspects	**Fog Computing vs Cloud Computing**
Usability	Among their primary at the moment, portable and mobile gadgets are used..
Flexibility	A system of fog can install software on devices with minimal specifications, such as alterations as well as internet protocol cameras.
Connectivity	Both facilitate wireless connectivity and machine-to-machine communication.
Distribution Strategies	Both can be densely or sparsely distributed on the basis of geographical location.

Capability	SBoth have the ability to process data produced by a variety of devices.
Resources	The computer resources of a fog system will be less than those of a cloud system (such as memory, processor, and storage), but they can be expanded as needed.

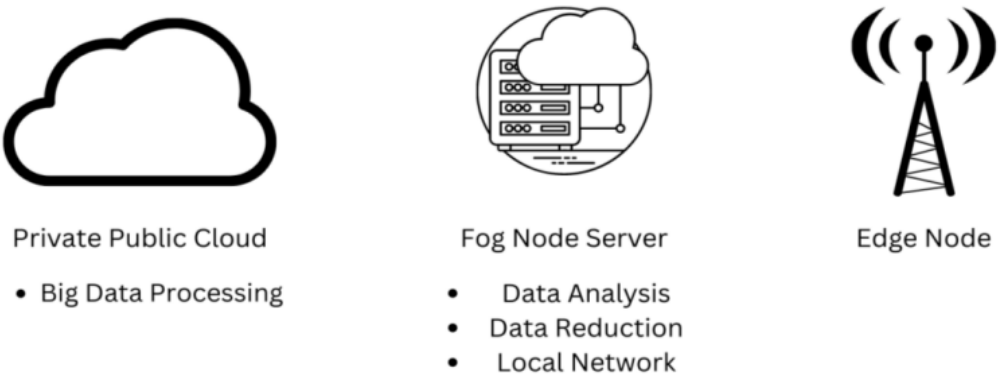

Figure 1. Cloud, fog, and edge computing

In the business world and the academic community, here is no difference between the terms "fog computing" and "edge computing". We distinguish in respect to both terminologies used in this essay. What separates them from edge computing and fog computing is processing of data as well as handling, likewise the location of processing and mental capacity, even if they both lessen network sluggishness and end-to-end delay, which is their shared objective. Alternatively, the primary distinction in contrast to centralized computing, fog computing is decentralized, meaning it does not use it. In a nutshell, the distance between the cloud infrastructure and the source, data processing and storage are carried out via a decentralized computing architecture. A computing facility is "pushed" nearer mobile phones as examples of data sources, actuators as well as sensors as the main idea behind edge computing. Instead of transferring data to the cloud, playing with edges separate responsibilities in local data processing. While this is happening, using its resources, a fog node determines if to examine information from various sources locally or send it to the cloud. Edge computing additionally provides support for a lot of cloud-related services, like SaaS, IaaS, and PaaS. Fog computing, on the other hand, provides all of these services. In conclusion, edge computing completely localizes itself at the edge of a network, but it also extends communication and computation resources there. Observe figure 1.

3. METHOD

3.1. MACHINE LEARNING ON THE CUTTING EDGE

The network of IoT systems is filled with edge networks, hardware, and sensors. Many IoT applications have to meet requirements for network security, capacity, and latency. These demands aren't met by cloud computing, though. A type of edge computing current technological solution that can fulfill these demands. For instance, edge networks can be used for data commutation by automobiles, including automobiles the road drive in concert with improved user competency. Another example is apps with high bandwidth requirements for virtual reality and augmented reality purchase content in a network edge. Analyzing data from sensors and traffic is made possible by the model. The classification of data derived from features retrieved from data sources has been accomplished in ML using a variety of techniques. It is possible to keep track of how the results are used for traffic engineering, disease detection, imaging recognition, and intrusion detection.

A. Utilizing Edge Computing

John Eskaliber et al. suggested a novel framework using several wearable gadgets. Even though these devices have lower resources, these authors advocated for the usage of edge computing devices. Additionally, they provided details on a revolutionary healthcare computing architecture that included edge network decentralized services. K-means clustering was utilized to identify Parkinson's illness in addition to healthcare monitoring techniques that leverage speech signal recognition. Similarly, a PreCog system was described that uses edge devices to cache and prefetch images to detect them quickly. The system being proposed is made up of several components that work together such as devices, edge servers, and cloud servers. Because of the complexity of computing and the amount of data that goes into image recognition, PreCog uses both local computer resources on the devices and cloud computing resources. Unlike the edge computing methods outlined above, which call for finishing computations on the edge server, this procedure does not do so.

Important portions of the suggested model are kept in the edge server and devices maintain a recognition cache. A HiCH, or a healthcare IoT network's hierarchical computer architecture, was proposed; in addition, the gadgets download a portion among the learned classifiers scheduled to be used in the future. The various layers of a fog network now use different ML architectural methodologies. For instance, sensor devices can carry out a variety of tasks like sensing and monitoring; the cloud is in charge of extensive training processes, whereas the edge is responsible for regional making decisions and overseeing systems.

The authors developed a system that was centered on the identification of arrhythmia and was according to the Monitor-Analyze-Plan-Execute-Knowledge model from IBM. The results show that, despite acceptable accuracy, Response time, bandwidth use, and storage efficiency of HiCH functions were better than those of traditional systems. Grassi et al. created ParkMaster, a low-cost crowdsourcing architecture. ParkMaster is in charge of the visual analytics used to assess parking availability. ParkMaster takes a street-side video and then calculates the quantity of automobiles identified through the use of ML techniques, as opposed to the traditional system of centralized surveillance that makes use of cell phones in a vehicle. The findings processed from several cars are uploaded to the ParkMaster cloud. Data is processed, and each driver is given a parking place recommendation. In smartphone edge computing situations, Want et al. developed a system that was intended to function as a service suggestion that relies on QoS prediction. The recommended approach takes mobility into account, in contrast to some context-aware service recommendation systems. Using algorithms for collaborative filtering that are based on mobility data, it then recommends services to other users. High prediction accuracy may be attained by the proposed method, as shown by the findings of several experiments using data from Shanghai Telecom.

B. Using Machine Learning in Edge Computing has Advantages

To get the foundation of the edge computing infrastructure, Leonhart presented an intelligent intrusion detection system. This system was used to enhance IBM's original "self-protecting" technology. By gathering information from sensors and applying the most recent unsupervised ML technique, the suggested IDS cleverly detects issues with gadgets that may be harmful to the overall structure. Finally, it was demonstrated that this idea system created by is capable of spotting discrepancies in the outside world.

Schneible et al. developed a different method that incorporates simulated neural networks for anomaly detection in edge computing environments. The model that is being provided is put together in a traditional neural network at a specific location (like the centralized cloud, for example). This feature causes delays and congestion to be seen close to the line of travel. The new system's core component is federated learning, in which the edge technology each maintains a model for

repeated training and distributes the training data among themselves. The centralized cloud repository then computes results of the training obtained from edge computing devices. The federated learning technique enables increased bandwidth, reduced latency, and optimal utilization of processing resources via edge networks. Abeshu et al. conducted additional research on distributed attack detection issues before contrasting them with non-distributed attack detection issues. It is harder to conduct an investigation of this kind. Due to the complexity and diversity of devices, it is thought that typical machine learning processes will perform with lesser precision and less scalability in edge computing contexts. Due to advancements in GPU technology and deep neural network theory, a novel deep learning (DL)-based approach has become more well-liked. The results demonstrate the superiority of the DL-based mechanism over traditional methods. Along with security issues, privacy is also a key component of the fog computing environment. For processing data pertaining to a fog computing architecture, there are sensors and devices, Yang et al. suggested an ML-based method for protecting privacy. This new system supports the proposed methodology, so a wide range of data sources can be accepted. Compared to a, the system is also a more scalable centralized system since it distributes computationally challenging activities to the network edge. The experiment's findings show that the technology can attain great precision without sacrificing user privacy. Hogan et al. provide an explanation of traffic engineering in edge networks. In edge networks, there may be a variety of end-to-end pathways, each with its own delay and bandwidth. The option indicated, however, properly satisfies consumers' needs and is of the biggest importance. By computing the results using portfolio theory, maximizing anticipated return, and taking into account level of risk, the proposed study offered a solution: a picture of the lifetime throughput that is anticipated. ML was employed to evaluate the risk level for each path while taking into account the model. Using actual latency traces, the proposed solution was then contrasted with existing approaches. As the solution had indicated, the better performance was attained.

4. RESULTS AND DISCUSSION

4.1. FOG COMPUTING RESOURCE MANAGEMENT

There are many different kinds of sensors, objects, and fog computing devices that may be used, and they all generate a lot of data that needs to be processed. In certain circumstances, real-time processing may be required. Making requests will enable devices, sensors, and objects to use resources to their greatest potential. Because of this, managing resources is necessary for it should be used for fog computing properly carried out. The research that employed ML for managing resource usage in fog computing was examined in this area.

A. Increased Computing

Edge SGD is a decentralized ML approach that the authors offered to do a balanced and quicker computation at the edge of a network to avoid sending raw data to the cloud. This computation at the edge of a network solves a significant linear regression problem. Similar to this, a revolutionary IoT-based strategy was put out by taking into account a regional outlook that enables automating the management of the system's components in the computing domain while utilizing ML methods. To autonomously write the score using ML techniques, ML algorithms were used to deploy recognizing and locating on cloud servers comprehend music. To ensure the effective distribution of computer resources, the authors used the suggested architecture. By moving the computing load between the fog nodes and the central server, the authors employed DL to address the huge data challenge. Their findings demonstrated the ability of their suggested approach to handling large amounts of data. For floating automobile data, a similar approach for data distribution was developed (FCD). Data loss can be prevented during times of connectivity disruption thanks to its architecture, which makes the suggested algorithm resistant to the issue of backhaul connectivity. The design of the program

also favors the modeling of distributed data in the fog, enabling the distribution of conditionally constrained Boltzmann machines to receive the data distribution's gathered information technique. Compressed data from mid-infrared spectroscopy using principal component analysis (PCA), wavelength transform, and compressed learning. The benefits of using MIRS for compressed learning, you can maintain limited rural network bandwidths, the decrease of application delay, conserving energy for communication and processing, and the minimization of necessary memory and storage spaces. As a result, large-scale data processing and fog computing greatly advantage of this technology. It was suggested to use cognitive-based communications, which are based on AIbased computer and communication improvements. Implementing ML gave network analytics as well, including network and networking issue applications of cognitive analytics. They include resource allocation enabling the operation of virtualized networks and efficiently using energy as part of network application in their work. By utilizing ML to identify user activity and offer sensor device scheduling with low latency adaptive MAC layers, the authors made an endeavor to conserve energy usage in fog computing, and latency. Regarding IoTbased health surveillance systems, a HiCH was suggested. Using the suggested computational architecture, hierarchical partitioning can be carried out while ML-based data analyses are running.

B. Making Decisions

Through a fog architecture that takes inspiration from nature, it exhibits low-latency Making decisions and managing resources in an adaptable way. ML and SmartFog are used, it is possible to simulate how the human brain works. By fusing Markov logic networks in machine learning, the authors concentrated on situational awareness and the choice of an ideal course of action. Signal data with a short lag time is examined close to a source of signal using various Machine Learning algorithms within the proposed directional mesh network (DMN) structure. The fog device can similarly decide when to upload data to when not to use the cloud's back end. This choice is made with the help of fog devices that are nearby wearables for smart telehealth and low-resource machine learning. By exploiting choices made throughout distinct events, machine learning was used to enable cognitive development across time. By utilizing ML, the authors of these studies enhanced decision-making in a fog computing environment. In addition, the decision-making capacity of fog computing has been improved by using unsupervised ML techniques. The majority of issues addressed have to do with computing burdens, resource limitations, and unforeseen circumstances.

C. Predicting delays and resource provisioning

Parallelization and distribution components of a DL with an edge architecture were offered as optimal resource provisioning using commercially available components. The resources needed by apps are evaluated with these factors in mind. To maximize resource usage, Memory and CPU speed are additional resources. The authors similarly concentrated on the resources being made available in multimodal fog computing. Based on complex ML algorithms, they proposed a productive algorithm. The basic objective of this method is to forecast the resources that fog devices will have access to. It may, nevertheless, also carry out other operations, like examining the veracity rendered of the outcomes as well as enhancing the scope of repeated employment descriptions. The creators instead concentrated on open-loop wireless communication and proactive network association using Machine Learning; this study enables management of anticipatory mobility. Another study was done to estimate usage of links for various picture-processing-related duties, as well as a complete delay, which includes how much time was spent on from data processing to transmit. A group of image-processing services from the GENI infrastructure were used by the authors to develop a computer environment with a realistic fog. The common IoT

communication protocol, Message Queuing Telemetry Transport, was suggested by the authors as a solution to the energy and communication issues between end devices with constrained resources. A total of four ML systems previously forecast data from actual sensors. A proper mix of layers can be used to perform different ML technique steps, as the authors of the study showed, which can help save energy. Numerous layers can be employed to implement the various processes that make up the ML approaches used in ubiquitous computing applications. In research that employed fog computing, ML was utilized to forecast processing delays and supply resources.

The main goal of employing supervised ML algorithms is to improve Fog computing: supply of resources, predicting delays, and energy use. The issues addressed included latency, forecasting time frames required to complete calculations, administrative costs, and energy usage about provisioning of resources, foreseeing delays, and consuming.

4.3. ACCURACY OF FOG COMPUTING

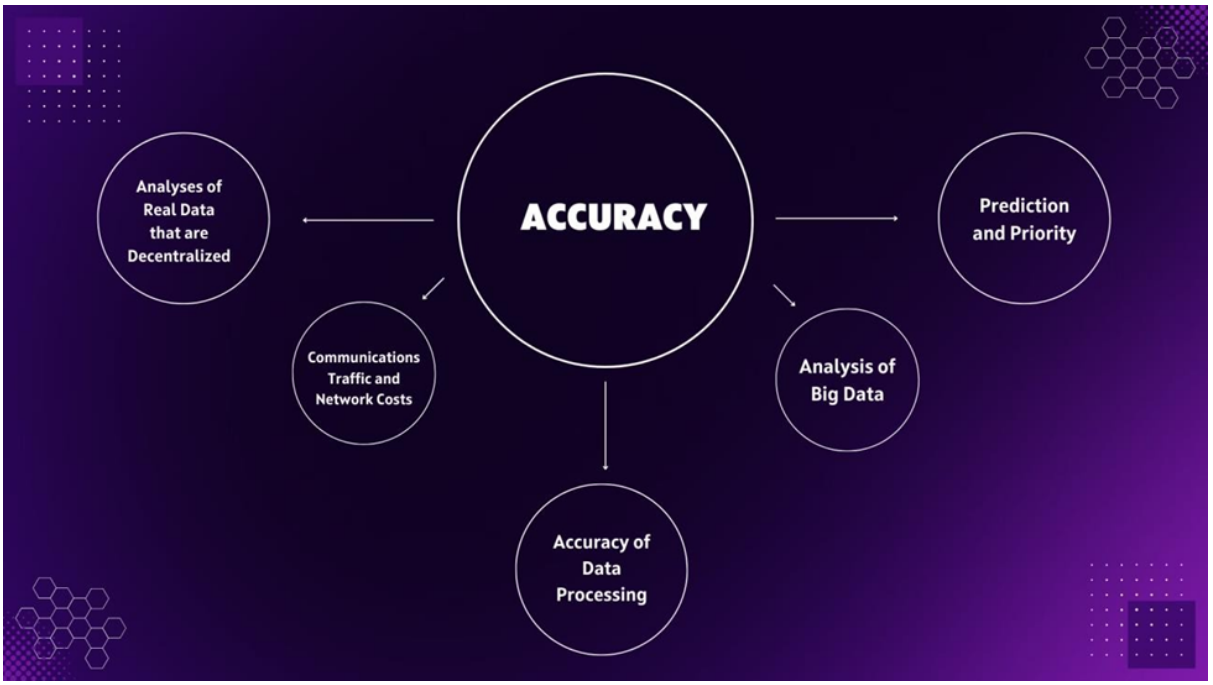

Figure 2. Major Issues in Accuracy

Fog computing helps with accuracy problems since it senses, analyses, and displays information in real-time. Consequently, real-time situational data is widely used in fog computing. ML methods play a crucial role in concerns with accuracy. The accuracy issue can be resolved, especially at the edge of a network, by combining fog computing and ML approaches. Enhancing fog computing accuracy with the use of ML approaches, as depicted in Figure 2, is the main topic that researchers are concentrating on.

A. Analyses of Real Data that are Decentralized

The authors' main goal was to examine trade-offs somewhere between precision and traffic. With the presumption that those facts are transferred wherein partial learning is presented to a number of data gatherers carried out, they presented transfer learning for dispersed hypotheses. The authors set out to address the problem of processing a massive network's edge containing large amounts of data. To analyze pregnancy data in real-time derived through means of IoT devices and gateways, they presented a Machine Learning methodology known to serve as one-dependence estimators. The authors enhanced the processing of big data produced by smart cities and industry. They examined information about mobile

nodes going using edge fog gateways and IoT devices of network architecture using a common distributed ML framework (i.e., HTL). In a separate study by the same researchers, DL was used to investigate how edge analytics, cloud computing, as well as fog computing can handle massive volumes streams of live information produced by cyberphysical systems (CPS).

B. Communications Traffic and Network Costs

Because there was so much aging data available in the real world, the authors concentrated on computing speed. They suggested combining a very effective DL system and age estimation algorithm working together to optimize the system's performance. A threetiered fog computing system that combines cloud, edge, and fog layers were presented using this particular combination. The authors discussed bandwidth and network latency issues. It was suggested to implement a system of universal healthcare, or UbeHealth. Highperformance computers, deep learning, and big data are used to estimate network traffic. The network and cloudlet layers use the findings to optimize data caching, routing, and rates. Similar to this, the authors presented a distributed learning framework that entails applying distributed ML approaches to adjacent nodes or the nodes directly responsible for producing the data to do partial data analytics on them. The challenge sensors are capable of detecting and processing was the focus of the authors' work. a system of early warning suggested to make it easier to spot animals in the vicinity of a road or railroad so that approaching cars might be warned of potential animal crossings. With machine learning, images taken on devices' edges may be categorized. Additionally, ML enables the forecasting of various timevarying traffic profiles.

C. Accuracy of Data Processing

The authors suggested the development of a novel in-network self-learning algorithm that may be incorporated into a constructive force management utilizing a cooperative Enhancing data processing outcomes with a fog platform. The overall amount of data processing and communication needed across the board in IoT networks can be decreased thanks to this innovative technique. To recognize both facial expression and the emotion expressed via voice, DL algorithms two known as These were VGG-Net and Alex-Net presented. An innovative affective experience management (AIEM) system was put forth. Three factors served as the foundation for the composition and design of this AIEM: accurate emotion recognition management, intelligent real-time management of emotional interactions and the collecting of emotion data.

D. Analysis of Big Data

It was suggested to create a novel system using big data platforms, Cloud, fog, and Internet of Things. In order to solve the sleep apnea condition despite current deficiencies, creative There were chances to offer fresh and inventive services presented in this study. The architecture of a big data analyzer modules' descriptions of the MLlib, the ML module's characteristics were provided by Scikit-learn 0.18.0 libraries and Apache SparkSQL. For the early reduction of customer-side data, another paradigm was put out. With relation to endto-end data minimization applications for businesses, this platform also offers a commercial model. This study's findings demonstrated an improvement in privacy, client trust, secure data exchange, and utilization costs. The main issues in sustainable businesses and medical care are big data reduction and real-time preprocessing

E. Prediction and Priority

It was revealed how cyber-healthcare would be implemented. The fog-based, multilayer framework's main focus is an identification system for a patient's condition, which heavily relies on machine learning (ML) methods. A multiplayer architectural patient-centered IoT ehealth ecosystem including a device, fog

computing, and other components, and a cloud was introduced by the researchers to assist in the management of complicated systems considering data's latency, variety, and speed. Biologically inspired machine learning (ML) unsupervised intelligence technology called hierarchical temporal memory (HTM) is the foundation of this system. The created distributed sparse representations are sent into the HTM ML module. Big data analytics also leveraged GraphLab, a scalable and quick ML platform.

4.3. SECURITY FOR FOG COMPUTING

When users utilize IoT networks, they want to receive safe, dependable services; in other words, there ought to be trust between a fog network's entire set of devices. By making the data anonymous, data services believe that the cloud or fog is an ideal area to examine and determine which data needs to be processed in a certain way. Distributed systems are more vulnerable to cyberattacks than centralized ones, and this vulnerability is exacerbated by developers' propensity to focus on building functional systems before adding security measures. Security of the devices, the network, and the data make up the three security tiers ML solutions that are associated with fog computing.

For the protection of data privacy, several writers have put forward several strategies. To offer a confidentiality shield system founded on ML, Marlin et al. used a linear regression approach. The fact that this mechanism provides a clause allowing for multipurpose aggregated data approach means that it can accommodate a range of data sources. To increase a system's capacity for growth, this novel approach distributes computationally demanding activities to the edge of the network. The results of the experiments showed that the suggested method was capable of achieving a high accuracy of 90% without endangering user privacy. Franklin et al. employed Using simulated and real-world datasets, a neural network, linear regression, boost, ensemble bag, as well as SVM with differential privacy are used to improve the effectiveness of query evaluation. Their findings showed that using a prediction model made it easier to keep data private and eliminate the mean absolute error. A unique Identity-based encryption and identity-based signing are the two cryptographic approaches used by the Fog Security Service (FSS), and has been proposed by the author the fog layer to offer IoT devices end-to-end security. Authentication, confidentiality, and nonrepudiation are just a few of the services that FSS has offered. One network being utilized architecture with a lot of traffic, the FSS has been assessed and utilized in a simulator for OPNET. Strengthening the crucial SCADA-based IoT infrastructure integrity, security, and privacy at the fog layer was the major objective of a novel security "toolbox" that was developed. Two essential components of the toolkit are integrated: a cryptographic-based access method and signature techniques at the fog layer technique at the cloud level. In order to show that the suggested platform is suitable for use in a practical application, the authors provided a prototype for evaluation of the proposed work.

In conclusion, social IoT system databases need to be secure from hackers. Utilizing cutting-edge differential privacy techniques, which can shield data from flaws and improve query assessment while supporting ML algorithms, efficient privacy protection can be obtained.

4.4. OPEN PROBLEMS AND CHALLENGES

Fog computing uses heterogeneous sensors and many sorts of edge network data as its data sources. A system is made more difficult by the variety of data kinds that can be collected, including confusing and partial data. For the sake of the management of resources, reliability as well as safety, Figure 3 lists the fog computing difficulties. In the next subsections, more information is given. more "intelligence" than existing family planning technology because they can "think" independently without relying on input from computer programmers. Existing technology is still considered "pseudo-artificial intelligence" because it relies on information and actions in a goal-oriented direction determined by the programmer.

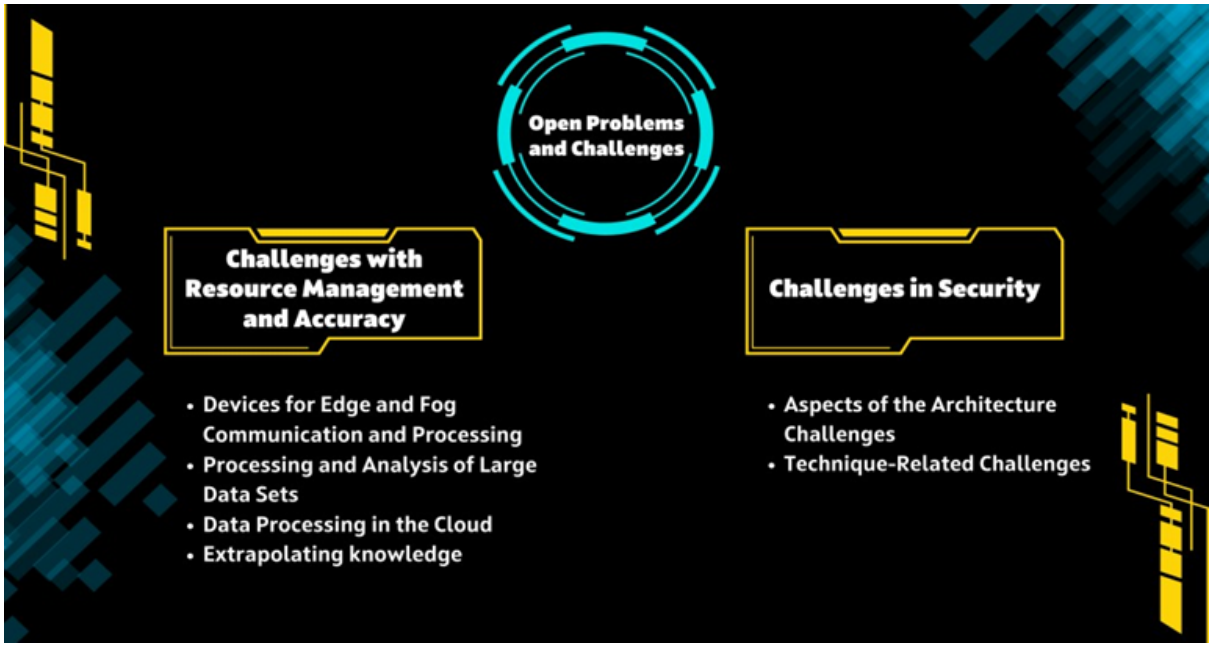

Figure 3. Fog computing's difficulties and unsolved issues

A. Challenges with Resource Management and Accuracy

Managing at the edge of a network, data processing is a challenge given the size of the majority of applications worldwide (fog). Although shifting processing to a network's edge and preprocessing, rather than using centralized processing, have helped combating fog with computation many processing obstacles, there are still several issues that need to be resolved. The management of data processing has benefited greatly from ML. The system will be vastly enhanced if machine learning is properly applied to fog computing. We then list the difficulties and unresolved problems in fog computing's processing and computation management.

1) Devices for Edge and Fog Communication and Processing

The action of storing, networking of data and analytics have all been made possible by the introduction of the use of fog computing fresh paradigm. Nearer to the edge of apps and devices, analytics is carried out. To improve requirements for processing and communication with respect to latency, energy efficiency, and dependability difficulties, DL, Time-sensitive applications in fog computing networks should use ML and AI. Due to edge nodes' limited processing capabilities, data outsourcing to a fog network may result in decreased adaptability and accuracy levels. To handle the erratic loads of distributed applications, a fog architecture should be scalable and adaptable. IoT cloud designs that are effective for these new difficulties are needed as the number of sensors and data needs rise, increasing communication and energy limits. With the growing amount of data collected in the wild, existing DL platforms' slow computational speeds cause processing and communication issues. As soon as processing is carried out with respect to sensors, edge devices both the end-to-end latency and energy use can be decreased. Using an algorithm that learns on its own can greatly enhance the results of analyzing data while also lowering the overall volume of data processing and communication needed throughout the entire Internet of Things (IoT) network. If cloud and fog computing platforms for big data and the Internet of Things are combined with technologies, new prospects may be possible for the creation of new, cutting-edge services.

2) Processing and Analysis of Large Data Sets

The Internet of Things (IoT) generates a volume of data that is substantial that must be processed and analyzed. The absence of computer cognition and expertise in specific fields is a significant obstacle. The fact that the current inspection system does not ensure good performance and efficient processing is another difficulty linked to the implementation of many sensors. Real-time data analysis is needed to process and analyze huge data volumes generated at a network's border. AML's aone-dependence estimators that have been verified must be used to carry out such an investigation. Combining fog computing with deep learning can address the aforementioned issues. IoT and mobile device dissemination will generate a significant volume of data, which will be too much for centralized cloud-based analysis to handle. One of the biggest challenges that still have to be solved is the reliability from largescale unlabeled inputs, in multilayer representation, extraction of multiscale features that are both scalable. The issue of processing resources will be solved by a fog network's huge data reduction, which will also significantly cut the cost and energy of employing cloud-based analysis.

3) Cloud-based Data Processing

Processor delay is a result of centralized cloud computing. Contrarily, the spread of live data in the mist as opposed to residing in the cloud centrally significantly increases the latency of data processing, especially when ML analytics are used to make predictions. Similarly to this, because there is no Internet connectivity and there is a massive volume of data generated by numerous apps, the transmission is currently the transfer of unprocessed raw data to a centralized cloud not viable. Telehealth, which produces huge data in the field of medicine, is one of the aforementioned uses. It takes time to transfer these enormous amounts of data to the cloud, which could delay processing, which is not acceptable, especially for medical data. Another significant factor that should be taken into account when processing data is energy. Given that big data requires a lot of processing power, the use of clouds for processing enormous volumes of data will use a lot of energy. As a result, when the management data, services, and applications in computing are relocated employing ML and DL, from the core "cloud" to the fog approaches, the capability to handle key events will become better to provide real-time processing.

4) Extrapolating knowledge

polating knowledge Cloud computing can be used to derive knowledge from the raw data that sensors have collected. This strategy may prove difficult, though, given the quick growth in sensor numbers and the massive amount of data they will generate. With the use of machine learning (ML), these problems will be solved by a decentralized system of computing the analysis of data. The knowledge extraction process can also be greatly aided and the user experience improved by combining compute nodes for the cloud and fog with DL methods. To determine a patient's medical status in healthcare systems, it is essential to make the shift the extraction of quantitative vital signs and the shift from clinic-centric therapy to patientcentric healthcare.

B. Challenges in Security

In terms of architecture and technical concepts, this Section highlights the difficulties and unresolved problems associated with fog computing security. to gain a thorough understanding of the sorts of security risks that face fog computing and the obstacles they face.

1) Aspects of the Architecture Challenges

- Trust

  As a result of their frequent deployment and lack of stringent oversight and security, fog computing devices are vulnerable to all kinds of security risks. Consequently, building trust at the level of the fog is the main obstacle. Public key infrastructure is one method that can help to partially resolve this problem. Furthermore, fog might show possibilities for the trusted execution environment. Fog's open network environment makes it simple for malicious software to proliferate to intelligent gadgets and seriously endanger user data. In order to maintain a trustworthy execution environment, harmful procedures must be kept under control and prevented from spreading.

- Authentication

  Since services are provided to end users on a large scale by front fog nodes, authentication is a significant issue of security in fog computing. To connect to the network, a device must first self-authenticate a network of fog before it can use any of the network's functions. The entry of unauthorized nodes must be stopped with this action. The fact that the devices connected to a network have varied limitations, such as those related to power, processing, and storage, also presents a difficult task.

- Security in Wireless

  Wireless sensors and Internet of Things (IoT) devices make up fog computing platforms. The quantity and visibility of wireless devices makes it challenging to maintain fog network security. In the absence of concealment and security, WiFi networks provide attackers unmatched ability to intercept critical information while it is being transmitted. In order to reduce unauthorized access points, packet sniffing, as well as other problems, encryption and authentication processes should be used for Wireless connectivity both indoors and outside the fog platform with end-user devices.

- Privacy of End Users

  Users' concerns about the breach regarding them in private, including location, information, including personal data, have increased in the modern day. The use of services like the Internet of Things, and cloud computing by end users draws attention to this information. Because fog nodes near clients can gather less tolerant data as opposed to a distant cloud on the main network, keeping confidentiality in a fog faces unique issues. As a result of the potential collection of sensitive information about an end user's identity, utility use (such as smart grids), or location by fog nodes located close to them, fog computing presents a challenge for privacy protection. Because fog nodes are dispersed across a broad area, centralized control is also made challenging. An untrusted party may be able to access the network through a poorly protected edge node. Once inside the network, the intrusive party has access to the private information shared between entities by users and can mine and steal it.

- Negative Attacks

  Without convenient security measures in place, a network's capabilities may be seriously compromised given the variety of harmful assaults that can be launched against the environment for fog computing. A malicious assault is a DoS attack. It is simple to execute a DoS attack since practically all gadgets connecting are not mutually exclusive inside a verified network. Spoofing multiple devices' addresses and sending phony processing/storage requests are two additional ways the execution of a DoS

attack. Because of the openness of its network, the fog computing environment is incompatible with current protection techniques for additional network types. The scale of a network is the primary difficulty. To circumvent performance improvements and storage and computational constraints, hundreds of hundreds of thousands of nodes making up an Internet of Things network may use services for fog or clouds.

2) Technique-Related Challenges

- Privacy Model

  The provision of health services categorization models can be delegated to fog servers or the cloud for determining decisions in order to reduce local computing costs. Health service providers view categorization models as assets, thus they should be protected as well.

- Numerous and Diverse

  Due to the Variety and quantity of data, Machine Learning is essential for the safety of fog computing. IoT data analysis should be up to date with these sectors' most recent trends because AI and ML are topics that are rapidly expanding. Data analysis in close proximity is crucial for a node, according to an assessment of several ML approaches and many IoT cases. Therefore, research on ML that can analyze a big amount of time series data while not requiring a lot of memory should be done.

- Flexible

  By putting a number of compromises related to Internet of Things (IoT) order of importance, the use of supervised machine learning decreases portions of appropriate security solutions. The way It's handled is a significant issue. A user can choose the best high levels of security degree that is adaptable thanks to flexible security. As determined by the restrictions on edge resources and of the users, the main goal of employing flexible keeping an IoT-based application secure is security.

- Methodology of Detection

  The IoT employs both anomalies and signature attack detection methods. In signature-based detection, an attack is found by gathering particular analyzing and comparing device data to a signature, which is a collection of rules or patterns. When detecting questionable actions or attacks a machine, discovering anomalies entails creating a model using samples of typical conduct as well as deviations based on the model. These methods, however, are unable to identify zero-day attacks. The main issue is detecting assaults whose signatures are not acceptable identified in the predefined collection of laws or guidelines.

- Security Issues

  The majority of currently used cryptosystems were created expressly to encrypt integer information. In addition, they can make a few straightforward calculations that could affect the outcomes and perhaps result in a misdiagnosis. The ability to provide a safe and precise diagnosis is one of the main obstacles in this area. The privacy of sensors has generally been protected in recent investigations by the use of cryptography. Cryptography should safeguard the encryption keys, however, it cannot be used in circumstances where it is necessary to share data with the general public. Differential privacy is currently being actively used to address privacy-related challenges. Release of query results rather than providing clients with datasets is the main idea underlying differentiated privacy. Due to the daily exchange of numerous questions and pieces of information between the system and its users, this process might not be appropriate for

CPS. This situation makes it difficult to establish differential privacy in CPS since a lot of noise must be added to the leaked queries in these circumstances.

5. CONCLUSION

As a result, our research highlights the contribution of ML to security, accuracy, and resource management in three domains of fog computing. Regarding resource management in fog computing, ML has been widely used to handle numerous issues, in contrast to preciseness and safety. Edge computing using Invoking Machine Learning incorporated into the cloud computing layers as well. This study addresses several difficulties and unresolved problems. As a result of fog computing's similarities to cloud computing, security is the more difficult component to address.

The potential for ML to be the dominant technology in many fields is too great. For fog computing applications, it might be regarded as a potent analytical tool. ML literature on its function in systems and services for fog computing is still lacking, despite the most recent application of Machine Learning to fog computing with success. This gap is examined in the current work. As far as we know, there hasn't been any research done on the application in fog computing of Machine Learning.

Supervised learning is the most common ML problem used in fog computing. The healthcare industry has made extensive use of ML and fog computing in terms of applicability. The development of fog computing applications and services is also significantly impacted by ML. As a result, when developing new applications for fog computing, studying and creating should take into account the benefits of Machine Learning to handle a variety of issues and obstacles.

ACKNOWLEDGEMENTS

Regarding the success of the research journal An Overview of Concepts, Applications, Difficulties, and Unresolved Issues in Fog Computing and Machine Learning, we are grateful to Raharja University and our supervisors.

REFERENCES

[1] P. Rashi, M. C. Lohani, N. Luftiani, T. Hermansyah, and I. N. Hikam, “New Personalized Social Approach Based on Flexible Integration of Web Services,” *Int. Trans. Artif. Intell.*, vol. 1, no. 1, pp. 1–17, 2022.
[2] A. S. Bist, V. Agarwal, Q. Aini, and N. Khofifah, “Managing Digital Transformation in Marketing:" Fusion of Traditional Marketing and Digital Marketing"," *Int. Trans. Artif. Intell.*, vol. 1, no. 1, pp. 18–27, 2022.
[3] M. R. Anwar, F. P. Oganda, N. P. L. Santoso, and M. Fabio, “Artificial Intelligence that Exists in the Human Mind,” *Int. Trans. Artif. Intell.*, vol. 1, no. 1, pp. 28–42, 2022.
[4] H. T. Sukmana, A. E. Widjaja, and H. J. Situmorang, “Game Theoretical-Based Logistics Costs Analysis: A Review,” *Int. Trans. Artif. Intell.*, vol. 1, no. 1, pp. 43–61, 2022.
[5] M. Wahyudi, V. Meilinda, and A. Khoirunisa, “The Digital Economy’s Use of Big Data,” *Int. Trans. Artif. Intell.*, vol. 1, no. 1, pp. 62–70, 2022.
[6] U. Rahardja, “Camera Trap Approaches Using Artificial Intelligence and Citizen Science,” *Int. Trans. Artif. Intell.*, vol. 1, no. 1, pp. 71–83, 2022.
[7] A. Ramadhan and T. Nurtino, “Integrated Energy System Systems and Game Theory A Review,” *Int. Trans. Artif. Intell.*, vol. 1, no. 1, pp. 84–101, 2022.
[8] C. Sriliasta, D. S. S. Wuisan, and T. Mariyanti, “Functions of Artificial Intelligence, Income Investment Instrument, and Crypto Money in Era of The Fourth Revolution,” *Int. Trans. Artif. Intell.*, vol. 1, no. 1, pp. 117–128, 2022.
[9] P. A. Sunarya, “Machine Learning and Artificial Intelligence as Educational Games,” *Int. Trans. Artif. Intell.*, vol. 1, no. 1, pp. 129–138, 2022.
[10] A. G. Prawiyogi, S. Purnama, and L. Meria, “Smart Cities Using Machine Learning and Intelligent Applications,” *Int. Trans. Artif. Intell.*, vol. 1, no. 1, pp. 102–116, 2022.
[11] S. A. Putra, “Virtual Reality’s Impacts on Learning Results in 5.0 Education: a Meta-Analysis,” *Int. Trans. Educ. Technol.*, vol. 1, no. 1, pp. 10–18, 2022.
[12] A. Rachmawati, “Analysis of Machine Learning Systems for Cyber Physical Systems,” *Int. Trans. Educ. Technol.*, vol. 1, no. 1, pp. 1–9, 2022.
[13] N. N. Azizah and T. Mariyanti, “Education and Technology Management Policies and Practices in Madarasah,” *Int. Trans. Educ. Technol.*, vol. 1, no. 1, pp. 29–34, 2022.
[14] N. Ramadhona, A. A. Putri, and D. S. S. Wuisan, “Students’ Opinions of the Use of Quipper School as an Online Learning Platform for Teaching English,” *Int. Trans. Educ. Technol.*, vol. 1, no. 1, pp. 35–41, 2022.
[15] C. S. Bangun, S. Purnama, and A. S. Panjaitan, “Analysis of New Business Opportunities from Online

Informal Education Mediamorphosis Through Digital Platforms," *Int. Trans. Educ. Technol.*, vol. 1, no. 1, pp. 42–52, 2022.
[16] U. Rahardja, "Using Highchart to Implement Business Intelligence on Attendance Assessment System based on YII Framework," *Int. Trans. Educ. Technol.*, vol. 1, no. 1, pp. 19–28, 2022.
[17] M. K. van der Hulst *et al.*, "A systematic approach to assess the environmental impact of emerging technologies: A case study for the GHG footprint of CIGS solar photovoltaic laminate," *J. Ind. Ecol.*, vol. 24, no. 6, pp. 1234–1249, 2020.
[18] A. Belhadi, S. S. Kamble, S. A. R. Khan, F. E. Touriki, and D. Kumar M, "Infectious waste management strategy during COVID-19 pandemic in Africa: an integrated decision-making framework for selecting sustainable technologies," *Environ. Manage.*, vol. 66, pp. 1085–1104, 2020.
[19] E. Igos, E. Benetto, R. Meyer, P. Baustert, and B. Othoniel, "How to treat uncertainties in life cycle assessment studies?," *Int. J. Life Cycle Assess.*, vol. 24, pp. 794–807, 2019.
[20] S. M. Moni, "A Framework for Life Cycle Assessment (LCA) of Emerging Technologies at Low Technology Readiness Levels." Clemson University, 2020.
[21] V. Prado *et al.*, "Sensitivity to weighting in life cycle impact assessment (LCIA)," *Int. J. Life Cycle Assess.*, vol. 25, pp. 2393–2406, 2020.
[22] S. Cucurachi, C. F. Blanco, B. Steubing, and R. Heijungs, "Implementation of uncertainty analysis and moment-independent global sensitivity analysis for full-scale life cycle assessment models," *J. Ind. Ecol.*, vol. 26, no. 2, pp. 374–391, 2022.
[23] S. M. Moni, R. Mahmud, K. High, and M. Carbajales-Dale, "Life cycle assessment of emerging technologies: A review," *J. Ind. Ecol.*, vol. 24, no. 1, pp. 52–63, 2020.
[24] C. G. Bhat, "Life cycle information models with parameter uncertainty analysis to facilitate the use of life-cycle assessment outcomes in pavement design decision-making." Michigan Technological University, 2020.
[25] S. Merschak, P. Hehenberger, S. Schmidt, and R. Kirchberger, "Considerations of life cycle assessment and the estimate of carbon footprint of powertrains," SAE Technical Paper, 2020.

	**Oscar Jayanagara** University of Pelita Harapan, Tangerang, Indonesia. She can be contacted at email: oscar.fe@uph.edu
	**Dewi Sri Surya Wuisan** University of Pelita Harapan, Tangerang, Indonesia. She can be contacted at email: dewi.wuisan@uph.edu