

Information Processing in IoT Based Manufacturing Monitoring System

Richard Essah^{1*}, Abraham Tetteh², Peter Kwaku Baidoo³, Bernice Duah⁴, Ephraim Quaynor Teye⁵

¹Scholar, Department of ICT, University of Education, Winneba, Ghana ^{2,3}Tutor, Department of ICT, Bia Lamplighter College of Education, Ghana

⁴*Tutor, GES Konongo Odumase Senior High School, Ghana*

⁵Division of Academic Affairs, University of Education, Winneba, Ghana

Abstract: The Internet of Things (IoT) is used widely in health care, manufacturing, industry, smart homes, and smart cities, among other areas. The data is collected in the IoT environment by placing the sensors in a structured way in a specific area. It collects data in accordance with the defined service for devices of IoT. For optimal handling of massive data in an IoT environment, the study work provided a new processing information in IoT centered factory system of monitoring. Data management is a critical and necessary activity in IoT systems, and present solutions of big data-centered are sufficient to satisfy every requirement. In the IoT Big data context, it's critical to increase data handling performance because most systems are solely designed for real-time data collecting. The proposed strategy utilizes Hadoop and Apache Kafka to meet the need for real-time data collection as well as offline processing. In comparison to the clustering model of traditional hierarchical and the neural network of back propagation model, the approach proposed performs well in data management and information extraction.

Keywords: Information processing, Internet of Things, manufacturing firms, monitoring system, sensor.

1. Introduction

Industry 4.0, or breakthrough digital technologies usage, is gradually improving product quality, worker safety, defect prediction, and energy utilization and production efficiency (Zhou, Liu & Zhou, 2015; El-Hamdi, Abouabdellah & Oudani, 2019; Chen et. al. 2017). Because they enable leaner and more efficient production, concepts of Industry 4.0 are expected to grow in the next five years by 20% in industrial sectors (Shrouf, Ordieres & Miragliotta, 2014; Wan, Cai & Zhou, 2015). In this context, many manufacturers are concerned with accelerating the integration and usage of secure, dependable artificial intelligence (AI) (European Cyber Security Organisation, 2018). AI-based manufacturing can improve business key performance indicators (KPIs) of manufacturing processes by combining industrial big data heterogeneous analysis, federation, and information modeling (Xu & Hua, 2017; Wan, Yang, Wang & Hua, 2018). In this context, connecting AIbased manufacturing processes to already deployed wireless networks is a tough study subject, especially when core processing occurs outside of industrial premises (Varghese &

Tandur, 2014; Trakadas et. al. 2019).

Most techniques of AI, on the other hand, are centered on models of mathematics that are hard for the average person to understand; as a result, most individuals see technology of AIbased as a black box that they ultimately grow to trust established on their own private experiences. Implementing human-centric AI (HAI) in systems of internet of things (IoT), such that systems of IoT cannot learn only from users nonetheless likewise offer easy explanations for estimations or judgments, is a new research topic (Garca-Magario, Muttukrishnan, & Lloret, 2019). The internet industrial of things (IIoT) is a network physical of equipment, items, or objects (with technology embedded) that are utilized in an industrial setting for remote and sensing control, allowing for deeper integration between the physical and cyber worlds (Moura, Ceotto, Gonzalez & Toledo, 2018). In the era of fifthgeneration (5G), providing high-performance, dependable, and efficient applications must be integrated with taking advantage of 5G network abilities. It is necessary to make the most use of available resources while adhering to tight Quality of Service (QoS) criteria for example jitter, high data rates, and latency ultra-low (Zafeiropoulos et. al. 2020; Zahariadis, Voulkidis, Karkazis & Trakadas, 2017).

AI, in this setting, is crucial to IIoT-enabled cybersecurity linked manufacturing setting, since it allows for precise threat detection and mitigation (Chatzigiannakis et. al. 2019; Fotiadou et. al. 2020; Lagutin et. al. 2020; Lagutin et. al. 2019). Simultaneously, presenting AI will result in a more safer and productive working environment, releasing routine procedures of human employees and allowing intelligent robots and machineries to do reasoning, heavy tasks, decision-making, and permitting human employees to concentrate on creativity (Yao, Zhou, Zhang, & Bor, 2017; Shin & Park, 2019). However, there are a few challenges and limitations to adopting and integrating AI-based innovation in the manufacturing domain that need to be appropriately addressed so as to realize its complete possibility without risking humans' indispensable role or critical procedures and data's security (Bresniker et. al. 2019; Zeadally, Adi, Baig & Khan, 2020). Researchers are familiar

^{*}Corresponding author: richardeessah84@gmail.com

with the phrase "data," as research into data management is still blooming with innovative technology.

Because of ever-increasing users' and services' availability, data amount has increased dramatically in recent years. The Internet of Things is explained by data vast volume created by actuators and sensors in a real-time context (IoT). IoT data collection architecture incorporates a variety of sources such as web resources, software applications, and other sources. All of these sources generate large amounts of data, necessitating a huge system of storage. Virtual sensors, as well as physical sensors, have just been advanced and are centered on data fusion mix from physical sensors and are employed in the cloud setting. The acquired data is referred to as sensor raw data, and it is processed, stored, and collected into meaningful info that aids in the resolution of data-related issues. In order to handle enormous amounts of heterogeneous data, sensor wireless networks are used for Internet of Things actualization, and sensor networks of large wireless scale are utilized for data management in computing cloud environments. The Internet of Things (IoT) is used widely in health care, manufacturing, industry, smart homes, and smart cities, among other areas.

The data is collected in the IoT environment by placing the sensors in a structured way in a specific area. It collects data in accordance with the defined service for devices of IoT. However, these sensors have constraints for example energy management, distance, and sensitivity, they gather data from the setting and send it to a common central node for analysis, after which the essential info is transferred to additional nodes. Because several devices of IoT send data bundles to a main server or central node, a unit storage is required to enhance and store the data before sending it to the cloud. The user is encouraged to have proper understanding of the sensor data so as to shun inaccurate data occurrence. The advancement of data gathering devices and applications enhances people's lives while also rapidly increasing the amount of data available. As a result, it is critical to handle, store, and analyze such data via the internet of things, according to a present study trend.

Big data, on the other hand, is an important process that processes acquired data and converts it into useful information in order to gain process knowledge. Big data is a vast volume, diversity, and velocity of information asset that necessitates information efficient system processing to enhance the automated process of decision-making. Analytics of big data is data analysis extension to analyze and handle real-time data huge amounts with many architectures. It serves as a boundary, providing support regarding forecasting, productivity, and innovation for the exponentially rising data in order to acquire the required solutions. Because data becomes outdated in a little period of time, it is critical to deal with it on a good platform while it is lively. Management of big data is a difficult undertaking since it necessitates an effective management data strategy, yet it is still necessary for some applications such as forecasting of weather. Traditional systems have difficulty processing enormous amounts of data, making it difficult to analyze and process. Regarding time and cost roles, the database management system lags behind while managing massive data.

When opposed to traditional data sets, big data is unusual in the storage process since the information is stored each byte, requiring the complete dataset to be analyzed in order to retrieve the required information. The raw data is transformed into useful data in the form of volumes that can be recognized as patterns. Big data analytics include efficient operations, smart decision-making, product development, and service system of well-functioning. Data cleansing, data capturing, data processing, data association, data distributing, data indexing, data moving, data mining, data displaying, and data analyzing are all key elements of big data processing, which are frequently utilized for data processing of real-time. The analysis aids both academic and technical users in gaining a previous data understanding. As stated in the design analysis, it converts data into features and predicts data into appropriate applications. Volume, diversity, and velocity are three essential concepts that are commonly mentioned in big data research.

In which the variety indicates data sources' heterogeneous nature for example machines, sensors, and extra applications of data generating, while the volume signifies the rising data nature regarding bytes. Velocity refers to the degree upon which data is produced and defines how well the generated data meets the requirements. Other important big data characteristics, as well as velocity, diversity, and volume, are veracity, which ensures that data is transferred authentically from multiple sites and regions, and value, which is applied to signify stored data significance and its merit computing. In a big data architecture, collected data is kept and processed, and the time period for storing and processing the data is set regarding validity. Equally, in a big data environment, the information flow is dynamic because most applications are used to handle real-time acquired data, so it's important to talk about how the data varies regarding variability. Another important big data feature is venue, which specifies the particular area where data is kept and retrieved. These locations are commonly referred to as data centers.

Because raw data is meaningless until it is correctly analyzed, it is necessary to explain data nature regarding vagueness, and then vocabulary is applied to explain other notations of grammar and data readability. All of these major aspects work together to help you handle your data more effectively in big data analysis. The four categories of big data analytics are predictive analysis, descriptive analysis, prescriptive analysis, and diagnostic analysis. In this case, descriptive analysis is a first-stage processing data procedure that specifies past data in order to establish the structure of data. Descriptive analysis organizes data using appropriate data mining approaches and helps the extraction of relevant information from the identified patterns. Future predictions based on probability could be conceivable in descriptive analysis, giving the user a clear picture of what will happen next. In predictive analysis case, existing data is applied to forecast future events. It's akin to a model forecasting, which employs numerous techniques of data mining, as well as artificial intelligence, to extract vital information from current and historical data.

In diagnostic analysis, the problems' root cause is determined in order to gather the necessary information about the scheme's behavior. Constant efforts are made in diagnostic analysis to identify faults so that they can be prevented or corrected in the future, improving efficiency of data handling in environment of big data. When dealing with large amounts of data, prescriptive analysis is used to make decisions centered on raw data study. This prescriptive analysis offers a healthier answer to the challenges in management of big data by offering crucial historical data and prediction specifics using a predictive analysis method. Figure 1 portrays analysis of big data classifications.



2. Motivation

The necessity for system of IIoT-based that monitors production act and security is the driving force for this research. This system, which allows for important processes' digitization in vital infrastructures such as plants of manufacturing, must be dependable, secure, privacy-preserving, and resilient. Together, it must make it easier for humans and machines to engage, as well as encourage peer-to-peer collaboration. It should also gather, store, and analyze data for manufacturing organizations to improve. AI, in this setting has the ability to efficiently meet numerous of these requests, allowing it to be integrated into the fabric of a reliable IIoT Big data system.

3. Literature Review

The IoT's challenges and issues this section examines the big data environment by examining current models of research. Big data IoT issues arise first and foremost from the huge amount of data collected in the environment of IoT via radio devices frequency and sensors (Cai et.al. 2017). Because the Internet of Things has hundreds of devices that generate data on their own, the amount of data collected rises. This large amount of data causes bandwidth challenges when storing, processing, and transferring in a real-time context. To process the information, a large quantity of bandwidth is necessary, which is not always reliable. Similarly, concerns relating to data storage necessitate a high storage space amount for data management and ensure that recovery and backup of vital data is possible in the future (Ngu et. al. 2017) In contrast to storage, time processing will rise when a big data amount is processed in real time via the system, potentially affecting the system's service quality. Data gathering in an IoT system is not the same as it is in other systems (Pandian et. al. 2019). Diverse sensors' types are used in design IoT dependent on the application, which offers dissimilar data forms, which is regarded a tough procedure in big data IoT.

Because the data is unstructured, organized, and semi-

structured (Condry et al. 2016), the system must store it and gather it in an individual location storage, which takes up additional capacity. Nearly 75% of the data collected cannot be handled using traditional methods, necessitating the development of an effective methodology to process both organized and unstructured data. Another essential problem in big data IoT (Bashar et al. 2019) is data transmission speed, which defines directly velocity, which is one main key features in analysis of big data. In a real-time context, processing data to extract vital information necessitates data analytics highspeed, which is challenging to do (Yi, Xu & Xu, 2019). Another problem with IoT big data is time series (Bestak & Smys, 2019) for effective data analysis. Sensors in an environment of IoT are applied to gather data at specific period intervals for specific applications, and the collected data can be meaningless at times if there is a change. As a result, the major concerns must be addressed using data recorded (An et. al. 2019). One of the challenges with IoT Big data is that most systems have difficulty with time data sequence processing in health care applications (Chandy et. al. 2019; Smys, 2019).

Finally, in IoT big data, security and privacy considerations are a significant challenge (Wang et. al. 2018). The data from the sensors to the station base is handled over a wireless connection, which poses security and privacy concerns (Metzger et. al. 2019). A network potential being exposed to attacks that compromise data security. The information is stolen from devices of IoT either actually or whereas the transmission is in transit. Because devices of IoT lack self-defense security procedures, controlling privacy and security in IoT large data is a serious concern. While privacy must be carefully considered, authentication must also be taken into account (Wu et. al. 2019). In the IoT big data world, offering information access is a difficult task. The approach proposed is meant to construct a healthier processing information system in the big data IoT environment, taking into account these constraints (El-Mougy, Ibnkahla & Al-Shiab, 2019). According to the results of the poll, present technologies have challenges with data extraction and processing, as well as security matters. The model proposed is discussed and built in the next part based on the difficulties.

A. Intelligent manufacturing concept

According to Penn et. al., research's intelligent manufacturing will be produced in the future; every product will be entities of some algorithm kind, specifically "pan-robot" period (Preuveneers et al. 2017). It is self-evident that developing intelligent manufacturing is critical for the industry of Chinese manufacturing and even the economy of China. Industry of manufacturing is slowly becoming digitized, and additional data is being collected on similar platform of data due to manufacturing and technology information integration (Penn, Pennerstorfer, & Jungbauer 2018). The industrial business has become fully sophisticated thanks to data analysis (Cardin et al. 2017). The platform system conducts analysis of data to develop value and knowledge centered on manufacturing digitalization and the massive generated data by "Internet+" (Lim et al. 2018). Intelligent manufacturing, according to Guo et al., will encourage the development of

integrated vertically models of business, intelligent platform models of business, and integrated horizontally models of business (Guo, Li, and Pang 2018). Intelligent manufacturing would rebuild future models of business with the platform at their center, which would not help only manufacturers reduce costs and increase productivity, but also allow businesses to rebuild models of business and rethink positioning value (Mohtar 2017).

Learning machine can be used to systems of intelligent manufacturing, according to Ozay et al. (2017). Building an individual system with learning machine roles is one option. Additional option is to create an enterprise-level platform of learning machine that will deliver machine learning capabilities and services to other enterprise systems. As said by Ozay et al. (2017), the data source layer, acquisition data layer, analysis data layer, application layer, and storage data layer constitute the platform system of latter learning machine architecture. As said by Klaine et al. (2017), pattern recognition and expert systems technologies have been widely used and implemented in numerous disciplines, including robotics, language natural understanding, and visual recognition. According to Giusti et al. (2017), the unique system of expert describes the experimental and experience data of professional business in a regular manner system, and then incorporate algorithm programming of mathematics to discover problem optimal answer consistent with the conditions given, for example the scheduling dynamic in programming multi-objective, whereas the recognition pattern is centered on the set features, and model identification is specified by setting parameter technique to attain the selective purpose, concentrating on resolving problem of sensing of small data variation and targets of individual business, for example signal production processing, statistical control process, and image recognition.

Learning machine can apply algorithms standard to study samples history, extract and select characteristics, and continuously optimize and build models, increasing the original system's capability to learn independently, which solves the uncertain business in the production process and improves the system's intelligence level (Ge et al. 2017).

According to Wang et al. (2017), technology of Internet of Things introduction to manufacturing intelligent systems, according to Wang et al., would encourage business models' intelligent manufacturing development. Yang et al. highlighted artificial intelligence's challenges in intelligent manufacturing and provided solutions (Yang et al. 2018). The advantages of small-scale intelligent manufacturing systems are discussed by Day et al., who point out technology of intelligent manufacturing is built on communication information technology and technology manufacturing (Day 2018). According to Lv et al. (2017), full people and machines separation is a trend in technological advancement. Manufacturers need to increase the model competitiveness, flexibility, timely, and sustainability responsiveness of their business intelligent manufacturing via modern technology and creative management techniques to thrive in an ever-changing and highly competitive market (Lv & Lin 2017). The majority of intelligent manufacturing literature now focuses on three aspects: overarching idea description, implementation, and system design, development, and benefits. It's hard to find study on systems of intelligent manufacturing centered on learning machine for pertinent evaluation and analysis, which is precisely what manufacturing companies need right now. As a result, based on international and domestic research, this paper proposes a method of evaluation for business model of intelligent manufacturing so that firms may clearly grasp their current state and deficiencies in order to make next-step judgments.

B. Contribution

The RAMI 4.0 (German ZVEI Electronic and Electrical Association of Manufacturers, Germany, Frankfurt) is a manufacturing systems' blueprint that was created to generate a single vision and develop a common understanding among all stakeholders. Layers are well-defined and structured (Integration, Asset, Information, Communication, Business, and Functional layers), RAMI 4.0 is architecture reference of service-oriented that spans the entire product life cycle, including all aspects and IT (information technology) constituents (from product to linked world) (from design to maintenance). About 3 sorts of extensions were propose to architecture RAMI 4.0 to overcome the aforementioned issues associated with AI implementation in smart manufacturing. First, every current layer of RAMI 4.0 will be enhanced, replicating the reality that AI is a cross-cutting issue that affects every aspect of systems of IT manufacturing. Certainly, AIassisted manufacturing necessitates information modeling and new data processing techniques, and AI will enable additional automated and autonomous intelligence business. Second, human-in-the-loop layer was designed that offers techniques, tools, and models to aid human decision-making by facilitating collaboration between virtual AI-based entities and human inside a site of manufacturing. Third, a federation layer is established, which integrates unique concepts like AI-ondemand and federated secure learning schemes to facilitate the flow of knowledge across manufacturing locations regarding AI algorithms, training models, results of threat analysis, best practices, and deployment recipes. Finally, the rising interoperability need at many manufacturing ecosystem levels is one of the major obstacles to the introduction and adoption of AI principles in IT production systems (Zeid et. al. 2019). In this context, the use of existing criteria, for example Open Platform Communications United Architecture (OPC-UA) for service-oriented industrial communication infrastructure and Automation Markup Language (AutomationML) for semantic data exchange provide present ways for identifying the necessary interoperability and connectivity for co-operation intelligent in smart factories (Rosendahl et. al. 2018). As a result, it's critical to integrate and reuse with a variety of existing and emerging technologies.

C. Communication and Information Intelligence

The Information and Communication Intelligence layers' primary role is to conduct every necessary activity on datasets factory-wide so that the upper levels' constituents can make results of AI algorithms decisions running on batch datasets or top of such processed streams of data. The lower layer of our architectural approach processes non-labeled (raw) data created by devices for manufacturing accompanied by all additional necessary information (for example, results quality control data, data logs, and so on) and suitably transforms it before forwarding to the upper layers. The goal is to create an expanded metadata library that covers a wide range of data generated during the manufacturing process. This layer also includes key functionality for deploying algorithms of AI closer to the sensor (computing edge), as well as detecting changes in dataset statistics that indicate the need to retrain algorithms. This layer provides both data logistics and shaping as well as information modeling functionalities to enable the two key kinds of AI algorithms (data-driven machine learning (Zhang, Yang, & Wang, 2019) and knowledge-based models (Chen & Zhao, 2006). Finally, analysis threat constituent watches every data flow among services in real time. Individual primary constituents contained in this layer's functionality is described below. Data pipelines AI-enabled orchestrator constituent facilitates the deployment and creation of processing data pipelines with 2 main goals: I the constituent must enable the creation and deployment of pipelines that combine common processing data responsibilities (feature reduction, feature conversion, data fusion and anonymization, annotation and labeling, data cleaning, and so on) with models of AI (applied in the upper layers' services). In this regard, the constituent offers mechanisms for quickly constructing these pipelines, in addition to out-of-the-box set and extendable processors of data to make handling of data simpler and more effective for the most data prevalent in the factory; (ii) the constituent must be able to deploy pipelines and orchestrate the many frameworks and components that are applied, comprising the models of AI that are made available and built through containers. To accomplish so, the constituent must be capable to orchestrate the various modules and frameworks required to perform the operations on infrastructure distributed (public cloud, edge cloud, edge device). The long-term edge-based learning component aids in models of deep-learning development that should be used on devices edge as data-processing pipelines part (Ferrari et. al. 2019). For latency-sensitive circumstances or/and once bandwidth upstream is limited, such as processing video and audio from light processing ranging and detection (LIDAR) data or AR (augmented reality) headset on a robot mobile, edge-based learning is necessary. The component will provide resource-efficient neural network topologies that can be trained minus a substantial amount of input labelled (Hubara et. al. 2017).

Furthermore, because future factories will operate in changing constantly contexts, the constituent will back algorithms of AI that are not learned when on an individual huge data batch nonetheless are often re-trained whereas in use. The knowledge graph of intra-manufacturing is the platform's chief hub for management of knowledge. To enable knowledge representation and linkage, this module expands and combines already models accessible, both domain-independent and domain-specific. Furthermore, it employs and expands graph analytics and cutting-edge reasoning techniques for entity consolidation and link prediction in order to identify correlations between data from various modules and layers. The intelligence threat manager uses the curated and collected datasets, as well as algorithms of AI, to do analysis of threat to only not forecast possible cybersecurity issues, nonetheless also to mitigate and manage them in a suitable means. This constituent offers an answer to the problems with present signature-based techniques "(that can efficiently detect existing cyber-attacks, but are inherently incapable of discovering zeroday attacks where there is no predefined rule)" and methods centered on anomalies "(that can detect known and zero-day attacks with some limitations of false-positive rates, but cannot detect attack types such as distributed denial of service)".

D. Functional and Business Intelligence

This layer contains components and services that model the behavior and status of entire assets and manufacturing process operations (comprising humans). Information models and processing data abilities from the bottom levels are used to create these models, which are produced utilizing AI algorithms and trained models. Other AI-enhanced services will be created on top of these, either in HITL operations or for automatic optimization business goal. The next sections detail the constituents that make up this layer. Behavior models of component-oriented are twins digital with particular constituent's behavioral models, which move away from the conventional system-wide behavioral models. Each component not only digitally records the state of assets and properties through the asset administration shell (thus giving a digital twin), but also models state transitions, such as using a finite state machine. This enables us to reflect the current state of the manufacturing process in digital twins, including the logic that governs the transition to subsequent stages or states. Learning behavior on a component-by-component basis and assembling a complete model from them has the advantage that if one component fails, just the related model needs to be retrained, leaving the system-wide behavior model intact. If a component is replaced, the time it takes to adapt the entire behavior model is reduced. Another benefit of this method is that componentbased models are more likely to be reusable in different contexts, allowing for collaborative learning and component monetization. The basic goal of digital human/context models is to learn and model human processes, strategies, and judgments as part of cooperative tasks in order to produce efficient human-centric intelligent control systems. In order to address complicated circumstances defined by varying levels of uncertainty, environmental and context elements are combined with the derived operational models. There are two types of models created: classification models that can recognize situations and models that can predict human behaviors and decisions in a workflow "(e.g., short-term future movements of human operators in the shop-floor considering current situation, typical decisions for a given event)". In this way, more intuitive user interfaces may be created because the AI system anticipates how the human would interact while maintaining control. This layer also includes application-specific services

for business objective optimization. Section 6 discusses the platform's potential applications.

E. Human in the Loop

This layer contains cutting-edge technologies for efficient and intuitive cooperation among AI systems, machines, and humans, permitting them to leverage each other's strengths for extra cooperative effectiveness and intuitive execution of task and making of decision. By moving past typical mechanisms interaction between IT systems and humans on the floor shop, for example command-line computer screens, buttons, and pendant consoles, the multichannel and context-aware interaction manager promotes innovation. A good interface user is intuitive and does not necessitate the training of human operators in certain structures or actions. Voice commands could be one answer, however due to machine noise, AI systems on the shop floor have a hard time catching and processing voice commands. In its place, this constituent enables for many simultaneous channels of input (speech, facial expressions, and gestures) to offer input redundancy, overcoming individual channels' diminished robustness, such as that caused by changeable lighting or noise on the manufacturing floor. Context information, such as the operator's location or current production parameters, is used to advance human-machine interaction's spontaneity (Liu et. al. 2018). The decision support intelligent system (DSIS) will permit humans to make logical decisions at the business or strategic level, such as maximizing the performance of a manufacturing system, based on expert knowledge, considerable experience, empirical data, and context information. Traditional decision systems support (DSSs) that are commonly utilized in the manufacturing area will be enhanced with unique abilities, such as threat intelligence models or digital twins, in our suggested methodology. The combination of digital twins and IDSSs holds a lot of promise: the former lack knowledge of business enterprise limits and goals, whereas the final necessitate advanced and holistic simulation models to make references. As a result, enhanced data analysis tools will enable objective and evidence-based insights to be relied on.

F. Federated Intelligence

The Federated Learning constituent seeks to solve data challenge collecting for training or feeding models of AI whereas ensuring that the data is confidential and owned. Since they directly relate to aspects of the manufacturing process, volumes, product qualities, and so on, most (if not all) data and information in manufacturing is proprietary. This component uses private set intersection (PSI) technologies (Pinkas, Schneider, & Zohner, 2018) to enable 2 parties with private information set to recognize their information sets' intersection without enlightening any info other than the intersection, whereas open-source frameworks like TensorFlow Federated support decentralized AI models (Lim et. al. 2020). The intermanufacturing knowledge exchange acts as a conduit for information interchange between different manufacturing sites or processes. Instead of giving access open to the local

repository knowledge, this constituent incorporates a query engine to handle outside demands, as there is a requirement to limit what information is published and exchanged. These query engines also make it possible to implement a federated query-processing method across different sites.

G. Authorization and Security

The platform proposed likewise handles data and information sharing security and permission requirements. The signcryption schemes component, which is centered on novel primitives cryptographic such policy ciphertext encryption of attribute-based (CP-ABE) systems, provides an effective and scalable approach to this aim. The technique encrypts the data using a policy of access control centered on attributes set and ensures that the keys of user are linked to their descriptive qualities. Consequently, the data owner has thorough control over the control access rules used for data encryption, knowing that a user can only decrypt the data only if their secret key equals the policy access used for data encryption. As a result of this unique technique (Taha, Talhi, & Ould-Slimane, 2019), this system significantly minimizes the effort administrative required for main administration and distribution whereas assuring end-to-end data security. CP-ABE can be integrated with encryption of symmetric techniques (for instance, AES, Advanced Encryption Standard) to meet the needs of shop floor devices regarding resources processing and speed to encryption. This ensures the essential trade-off between performance and information granular protection.

H. Cybersecurity for Artificial Intelligence (AI)

Numerous constituents of manufacturing system of IoTbased are integrated with AI in the framework provided. However, this raises additional worries about the security and reliability of AI systems in general. Input assaults and poisoning attacks are two types of artificial intelligence attacks that target the AI algorithm. The former entails tampering with the AI system's input in the operation stage so that it produces incorrect results. Because input attacks do not require a controlled AI system, they are relatively simple to launch and succeed. Poisoning attacks, on the other hand, are caused by the corruption of the AI model's construction process. In this situation, the model is fed erroneous or mislabeled data in the training stage in order to control the process of learning. This attack type can likewise be used against learning federated; in this situation, algorithm or modified data from a federation member can cause the global model to be corrupted. The key feature of trustworthy AI is the protection against adversarial attacks, which has recently gained a lot of attention (Comiter, 2019). Traditional cybersecurity policies and mechanisms can be utilized as a beginning point to defend AI systems. In this regard, the suggested architecture's security-related components, which provide confidentiality, integrity, and threat detection, can provide a first protection degree. Though, delivering total AI cybersecurity, particularly in the case of IIoT, will necessitate additional modifications to address these AI systems' inherent vulnerabilities, which will be left to future study.

4. Proposed Work

By constructing an efficient industrial monitoring system for effective information processing, the approach proposed for big data IoT is examined. The process is separated in 3 stages: decision making, processing information for example analysis and classification, and data aggregation and collection.

IoT is used widely in a variety of science and engineering applications, and it has recently been used to improve industrial systems. Cloud computing supports a variety of services and applications, which are further enhanced by combining IoT with cloud. Significant elements in applications of manufacturing, for example data remote processing, necessitate a dependability high level. When using computing cloud for this setting, the cloud provides improved pre-processing data, ensuring a clean forwarding process that enhances the productivity of health care apps in a novel way. In production monitoring applications that collect data and information, a variety of sensors are typically used. The data is collected and sent to the network layer of IoT sensor, which performs encryption, data collecting, compression, and aggregation. This data collection procedure gathers information about a person's physical status and their surroundings. Data is acquired from sensor devices located in manufacturing firms, the environment, and the neighborhood through wireless networks based on the environment, manufacturing center, and location. It's a complicated process to collect and provide sensitive data in the accumulation model. For IoT-based manufacturing applications that share data across a wireless network, an efficient data protection operation is required. Through data hierarchical compression and encryption, the suggested solution reduces data confidentiality concerns. All of the information that has to be communicated, as well as the secret key, is shared among the device nodes and user in this procedure. If the nodes sensing satisfy the established enquiry conditions, the data is encrypted as text cipher. Sensing nodes employ encryption of data to share data with the base station using a secret key, and message compressed is sent to the node aggregate using a similar encryption key.

If the node aggregate is the way only for data sharing with the station base, the transfer data procedure is started; otherwise, nodes aggregate share the data collected with other nodes aggregate, and that node is regarded a mediator. The mediator facilitates information exchange and data aggregation with its data so that the suggested process can achieve the data difference. Normalization, information analysis, extraction, classification, and filtration are all part of data pre-processing. Information is exchanged as aggregated and encrypted messages from every node aggregate from the base station to the cloud layer. Data normalization is a technique for standardizing aggregate data. For normalization efficiency in an environment of IoT, the approach proposed utilizes the maxmax normalization technique. The model proposed flow process is in detail depicted in Figure 2.

Data filtration is a technique for removing noise and undesired elements from gathered data. The proposed solution makes use of a kalman filer to improve the noise reduction process. It separates the significant from the unimportant noises while also increasing the system's data processing speed. The final step in a manufacturing application is data analysis. Once the filtered data has been received, data analysis requires an indepth core analysis. The nodes master partition the data highspeed in diverse data fixed-size packets to manage huge data in applications of manufacturing at high speeds. As a master node, these packets data are separated in data fixed-size, which is then processed in slave nodes at the same time. The proposed approach makes use of Hadoop's distributed file system to disseminate data packets to all slave nodes (DFS). If above 1 packet needs processing by the nodes slave, the function map is applied to process them. However, Hadoop has limits when dealing with data real-time, thus the approach proposed relies on Apache Kafka, a distributed message of high-throughput system. The relevant parameters are retrieved using Apache Kafka, and monitoring of real-time is advanced in the approach suggested based on the extracted features and data categorization findings. The purpose of such a manufacturing system is to limit risks of production and to detect anomalous changes quickly. As a result, the IoT-based factory monitoring system will increase the use of right processes, accurate resources, quality products, and efficient manufacturing time.



5. Result and Discussion

The model proposed is verified experimentally in simulation by arraying many manufacturing sensors over a 1000x1000m area, with the experimentation being done in Python. Sinks and aggregate nodes that perform encryption, compression, transmission, decryption, and decompression operations using the described algorithms in the work proposed section make up the network. Parameters including cost functions regarding storage and transmission, accuracy, f-measure, specificity, and sensitivity are likened to current clustering hierarchical and neural network of back propagation models to validate the proposed work's performance. The simulation parameters utilized for the suggested work experimentation are listed in Table 1.

Figures 3 and 4 provide a comparison of the suggested model's specificity and sensitivity to the standard model. It has been discovered that the proposed information processing system outperforms other algorithms regarding specificity and sensitivity. The performance is graded on a scale of 5, 10, 15, 20, 25, 30 centered on the actions and requests. The approach

proposed achieves well performance on ninety-six percent average specificity, which is two percent higher than the neural network of back propagation model and four percent higher than the clustering hierarchical method, thanks to the online and offline operations enabled by Hadoop and Apache. The proposed approach achieves an average of 95 percent sensitivity, as shown in figure 3.









Figure 6 shows how the suggested model compares to the hierarchical and back propagation models in terms of accuracy. The accuracy parameter is determined by the amount of information extracted from aggregate nodes' performance and raw data. When compared to other models, the suggested model has a classification accuracy of 97 percent on average.

6. Conclusion

For optimal handling of massive data in an IoT environment, the study work provided a new processing information in IoT centered factory system of monitoring. Data management is a critical and necessary activity in IoT systems, and present solutions of big data-centered are sufficient to satisfy every requirement. In the IoT Big data context, it's critical to increase data handling performance because most systems are solely designed for real-time data collecting. The proposed strategy utilizes Hadoop and Apache Kafka to meet the need for realtime data collection as well as offline processing. In comparison to the clustering model of traditional hierarchical and the neural network of back propagation model, the approach proposed performs well in data management and information extraction. The suggested model achieves a 97 percent accuracy, which is a significant increase in data processing in the Big data IoT context. The research's next focus could be on employing optimization models to improve performance.

References

- J. An, F. L. Gall, J. Kim, J. Yun, J. Hwang, M. Bauer, M. Zhao, and J. Song, "Toward Global IoT-Enabled Smart Cities Interworking Using Adaptive Semantic Adapter". *IEEE Internet of Things Journal*, 6(3), pp. 5753-5765, 2019.
- [2] A. Bashar, "Intelligent development of big data analytics for manufacturing industry in cloud computing". *Journal of Ubiquitous Computing and Communication Technologies*, 1(01), pp. 13-22, 2019.
- [3] R. Bestak and S. Smys, "Big data analytics for smart cloud-fog based applications". *Journal of trends in Computer Science and Smart technology*, 1(02), pp. 74-83, 2019.
- [4] K. Bresniker, A. Gavrilovska, J. Holt, D. Milojicic and T. Tran, "Grand Challenge: Applying Artificial Intelligence and Machine Learning to Cybersecurity". *Computer*, 52, pp. 45–52, 2019.
- [5] H. Cai, B. Xu, L. Jiang and A. V. Vasilakos, "IoT-Based Big Data Storage Systems in Cloud Computing: Perspectives and Challenges". *IEEE Internet of Things Journal*, 4(1), pp. 75-87, 2017.
- [6] O. Cardin, D. Trentesaux, A. Thomas, P. Castagna, T. Berger, & B. H. El-Haouzi, "Coupling Predictive Scheduling and Reactive Control in Manufacturing Hybrid Control Architectures: State of the Art and Future Challenges [J]". *Journal of Intelligent Manufacturing* 28 (7), pp. 1503– 1517, 2017. https://doi.org/10.1007/s10845-015-1139-0.
- [7] A. Chandy, "A review on IOT based medical imaging technology for healthcare applications". *Journal of Innovative Image Processing*, 1(01), pp. 51-60, 2019.
- [8] I. Chatzigiannakis, L. Maiano, P. Trakadas, A. Anagnostopoulos, F. Bacci, P. Karkazis, P. Spirakis and T. Zahariadis, "Data-Driven Intrusion Detection for Ambient Intelligence". In *Ambient Intelligence, Lecture Notes in Computer Science*; Chatzigiannakis, I., De Ruyter, B., Mavrommati, I., Eds.; Springer: Cham, Switzerland. Volume 11912, 2019.
- [9] B. Chen, J. Wan, L. Shu, P. Li, M. Mukherjee and B. Yin, "Smart Factory of Industry 4.0: Key Technologies, Application Case, and Challenges". *IEEE Access*, 6, 6505–6519, 2017.
- [10] X. Chen and J. Zhao, "Research on the model and application of knowledge-based industrial design". In *Proceedings of the International Technology and Innovation Conference (ITIC 2006), Hangzhou, China*, 6–7; pp. 1369–1374, 2006.
- [11] M. Comiter, "Attacking Artificial Intelligence, AI's Security Vulnerability and What Policymakers Can Do About It". In *Belfer Center Paper*; Harvard Kenedy School: Cambridge, MA, USA, 2019.
- [12] M. W. Condry and C. B. Nelson, "Using Smart Edge IoT Devices for Safer, Rapid Response with Industry IoT Control Operations". *IEEE*, 104(5):938-946, 2016.
- [13] C. P. Day, "Robotics in Industry Their Role in Intelligent Manufacturing [J]". Engineering, 4(4), 440–445, 2018.
- [14] S. El-Hamdi, A. Abouabdellah and M. Oudani, "Industry 4.0: Fundamentals and Main Challenges". In Proceedings of the International Colloquium on Logistics and Supply Chain Management (LOGISTIQUA), Montreuil-Paris, France, pp. 12–14, 2019.
- [15] A. El-Mougy, I. Al-Shiab and M. Ibnkahla, "Scalable Personalized IoT Networks". in *IEEE*, 107(4), pp. 695-710, 2019.
- [16] European Cyber Security Organisation, "Cyber Security for the Industry 4.0 and ICS Sector"; European Cyber Security Organisation: Brussels, Belgium, 2018.
- [17] P. Ferrari, S. Rinaldi, E. Sisinni, F. Colombo, F. Ghelfi, D. Maffei and M. Malara, "Performance evaluation of full-cloud and edge-cloud architectures for Industrial IoT anomaly detection based on deep learning". In Proceedings of the 2019 II Workshop on Metrology for Industry 4.0 and IoT (MetroInd4.0&IoT), Naples, Italy, 4–6; 2019, pp. 420–425.
- [18] K. Fotiadou, T. Velivassaki, A. Voulkidis, K. Railis, P. Trakadas and T. Zahariadis, "Incidents Information Sharing Platform for Distributed Attack Detection". *IEEE Open J. Commun. Soc.*, 1, 593–605, 2020.
- [19] I. García-Magariño, R. Muttukrishnan, and J. Lloret, "Human-Centric AI for Trustworthy IoT Systems with Explainable Multilayer Perceptrons". *IEEE Access*, 7, 125562–125574, 2019.
- [20] [20] Z. Ge, Z. Song, S. X. Ding and B. Huang, "Data Mining and Analytics in the Process Industry: The Role of Machine Learning [J]". *IEEE Access* 5 (99): 20590–20616, 2017.
- [21] A. Giusti, J. Guzzi, D. Ciresan, F. He, J. Rodriguez, F. Fontana, M. Faessler, C. Forster, J. Schmidhuber, G. Caro, D. Scaramuzza and L. Gambardella, "A Machine Learning Approach to Visual Perception of

Forest Trails for Mobile Robots [J]". *IEEE Robotics & Automation Letters* 1 (2): 1, 2017.

- [22] B. Guo, X. Pang and W. Li, "The Role of Top Management Team Diversity in Shaping the Performance of Business Model Innovation: A Threshold Effect [J]". *Technology Analysis & Strategic Management*, no. 4, pp. 1–13, 2018.
- [23] I. Hubara, M. Courbariaux, D. Soudry, R. El-Yaniv and Y. Bengio, "Quantized neural networks: Training neural networks with low precision weights and activations". J. Mach. Learn. Res., 1, 6869–6898, 2017.
- [24] P. V. Klaine, M. A. Imran, O. Onireti and R. D. Souza, "A Survey of Machine Learning Techniques Applied to Self-Organizing Cellular Networks [J]". *IEEE Communications Surveys & Tutorials*, 19 (4), pp. 2392–2431, 2017.
- [25] D. Lagutin, P. Anton, F. Bellesini, T. Bragatto, A. Cavadenti, V. Croce, N. Fotiou, M. Haavala, Y. Kortesniemi, H.C. Leligou, et al. (2020). The SOFIE Approach to Address the Security and Privacy of the IoT using Interledger Technologies. In *Security and Privacy in Internet of Things: Challenges and Solutions*; Ramos, J.L.H., Skarmeta, A., Eds.; IOS Press: Amsterdam, The Netherlands.
- [26] D. Lagutin, F. Bellesini, T. Bragatto, A. Cavadenti, V. Croce, Y. Kortesniemi, H.C. Leligou, Y. Oikonomidis, G.C., Polyzos, G. Raveduto, et al. (2019). Secure Open Federation of IoT Platforms Through Interledger Technologies-The SOFIE Approach. In Proceedings of the European Conference on Networks and Communications (EuCNC), Valencia, Spain, 18–21; pp. 518–522.
- [27] C. H. Lim, M. J. Kim, J. Y. Heo and K. J. Kim, "Design of Informaticsbased Services in Manufacturing Industries: Case Studies Using Large Vehicle-related Databases [J]". *Journal of Intelligent Manufacturing* 29(3), pp. 497–508, 2018.
- [28] W.Y.B. Lim, N.C. Luong, D.T. Hoang, Y. Jiao, Y.-C. Liang, Q. Yang, D. Niyato and C. Miao, "Federated Learning in Mobile Edge Networks: A Comprehensive Survey". *IEEE Commun. Surv. Tutor*, 2020.
- [29] B. Liu, G. Tur, D. Hakkani-Tür, P. Shah and L. Heck, "Dialogue learning with human teaching and feedback in end-to-end trainable task-oriented dialogue systems". *ArXiv*, arXiv:1804.06512, 2018.
- [30] Y. Lv and D. Lin, "Design an Intelligent Real-time Operation Planning System in Distributed Manufacturing Network [J]". *Industrial Management & Data Systems* 117(4), pp. 742–753, 2017.
- [31] F. Metzger, T. Hoßfeld, A. Bauer, S. Kounev and P. E. Heegaard, "Modeling of Aggregated IoT Traffic and Its Application to an IoT Cloud". *IEEE*, 107(4), pp. 679-694, 2019.
- [32] R. H. Mohtar, "A Call for A New Business Model Valuing Water Use and Production: The Water, Energy and Food Nexus Holistic System Approach [J]". *Water International* 42(6), pp. 1–4, 2017.
- [33] R. Moura, L. Ceotto, A. Gonzalez and R. Toledo, "Industrial Internet of Things (IIoT) Platforms-An Evaluation Model". In Proceedings of the International Conference on Computational Science and Computational Intelligence (CSCI), Las Vegas, NV, USA, 12–14; pp. 1002–1009, 2018.
- [34] A. H. Ngu, M. Gutierrez, V. Metsis, S. Nepal and Q. Z. Sheng, "IoT Middleware: A Survey on Issues and Enabling Technologies". *IEEE Internet of Things Journal*, 4(1), pp. 1-20, 2017.
- [35] M. Ozay, I. Esnaola, F. T. Y. Vural, S. R. Kulkarni and H. V. Poor, "Machine Learning Methods for Attack Detection in the Smart Grid [J]". *IEEE Transactions on Neural Networks & Learning Systems* 27(8), pp. 1773–1786, 2017.
- [36] A. P. Pandian, "Enhanced edge model for big data in the internet of things based applications". *Journal of trends in Computer Science and Smart technology*, 1(01), pp. 63-73, 2019.
- [37] [37] J. Penn, P. Pennerstorfer, and A. Jungbauer, "New Generation of Continuous Casting Plants with Intelligent Manufacturing Strategy; Neue Generation Von Stranggie? anlagen Mit Intelligenter Fertigungs strategie; [J]". BHM Berg- und Hüttenmännische Monatshefte 163(1), pp. 11–17, 2018.
- [38] B. Pinkas, T. Schneider and M. Zohner, "Scalable private set intersection based on OT extension". ACM Trans. Priv. Secur., 21, 2018.
- [39] D. Preuveneers and E. Ilie-Zudor, "The Intelligent Industry of the Future: A Survey on Emerging Trends, Research Challenges and Opportunities in Industry 4.0[J]". Journal of Ambient Intelligence and Smart Environments, 9 (3), pp. 287–298, 2017. https://doi.org/10.3233/AIS-170432.
- [40] J. S. Raj, "A Novel Information Processing in IoT Based Real Time Health Care Monitoring System". *Journal of Electronics and Informatics*, 02(03), pp. 188–196, 2020.
- [41] R. Rosendahl, A. Calá, K. Kirchheim, A. Lüder and N. D'Agostino, "Towards Smart Factory: Multi-Agent Integration on Industrial Standards"

for Service-oriented Communication and Semantic Data Exchange". WOA, pp. 124–132, 2018.

- [42] K. Shin and H. Park, "Smart Manufacturing Systems Engineering for Designing Smart Product-Quality Monitoring System in the Industry 4.0". In Proceedings of the 19th International Conference on Control, Automation and Systems (ICCAS), Jeju, Korea, 15–18; pp. 1693–1698, 2019.
- [43] F. Shrouf, J. Ordieres and G. Miragliotta, "Smart factories in Industry 4.0: A review of the concept and of energy management approached in production based on the Internet of Things paradigm". *In Proceedings of the IEEE International Conference on Industrial Engineering and Engineering Management*, Bandar Sunway, Malaysia, 9–12; pp. 697– 701, 2014.
- [44] S. Smys, "Big data business analytics as a strategic asset for health care industry". *Journal of ISMAC*, 1(02), pp. 92-100, 2019.
- [45] M.B. Taha, C. Talhi and H. Ould-Slimane, "Performance Evaluation of CP-ABE Schemes under Constrained Devices". *Procedia Comput. Sci.*, 155, pp. 425–432, 2019.
- [46] P. Trakadas, N. Nomikos, E.T. Michailidis, T.V. Zahariadis, F.M. Facca, D. Breitgand, S. Rizou, X. Masip-Bruin and P. Gkonis, "Hybrid clouds for data-intensive, 5G-enabled IoT applications: An overview, key issues and relevant architecture". *Sensors*, 19, 3591, 2019.
- [47] A. Varghese and D. Tandur, "Wireless requirements and challenges in Industry 4.0". In Proceedings of the International Conference on Contemporary Computing and Informatics (IC31), Mysore, India, 27–29; pp. 634–638, 2014.
- [48] J. Wan, H. Cai and K. Zhou, "Industrie 4.0: Enabling technologies". In Proceedings of the International Conference on Intelligent Computing and Internet of Things, Harbin, China, 17–18, 2015; pp. 135–140.
- [49] J. Wan, J. Yang, Z. Wang and Q. Hua, "Artificial Intelligence for Cloud-Assisted Smart Factory". *IEEE Access*, 6, 55419–55430, 2018.
- [50] H. Wang, C. Yu, L. Wang and Q. Yu, "Effective Big Data-Space Service Selection over Trust and Heterogeneous QoS Preferences". *IEEE Transactions on Services Computing*, 11(4), pp. 644-657, 2018.
- [51] J. Wang, L. Zhang, L. Duan and R. X. Gao, "A New Paradigm of Cloudbased Predictive Maintenance for Intelligent Manufacturing [J]". *Journal* of Intelligent Manufacturing 28(5), pp. 1125–1137, 2017.
- [52] F. Wu, B. Zhang, W. Fan, X. Tian, S. Huang, C. Yu and Y. Liu, "An Enhanced Random Access Algorithm Based on the Clustering-Reuse

Preamble Allocation in NB-IoT System". *IEEE Access*, 7, 183847-183859, 2019.

- [53] X. Xu and Q. Hua, "Industrial Big Data Analysis in Smart Factory: Current Status and Research Strategies". *IEEE Access*, 5, pp. 17543– 17551, 2017.
- [54] S. Yang, J. Wang, L. Shi, Y. Tan and F. Qiao, "Engineering Management for High-end Equipment Intelligent Manufacturing [J]". *Frontiers of Engineering Management*, 5(4), pp. 10–40, 2018.
- [55] X. Yao, J. Zhou, J. Zhang and C.R. Boër, "From Intelligent Manufacturing to Smart Manufacturing for Industry 4.0 Driven by Next Generation Artificial Intelligence and Further On". *In Proceedings of the* 5th International Conference on Enterprise Systems (ES), Beijing, China, 22–24; pp. 311–318, 2017.
- [56] M. Yi, X. Xu and L. Xu, "An Intelligent Communication Warning Vulnerability Detection Algorithm Based on IoT Technology". *IEEE Access*, 7:164803-164814, 2019.
- [57] A. Zafeiropoulos, E. Fotopoulou, M. Peuster, S. Schneider, P. Gouvas, D. Behnke, M. Muller, P.-B. Bok, P. Trakadas, P. Karkazis, et al. "Benchmarking and Profifiling 5G Verticals' Applications: An Industrial IoT Use Case". In Proceedings of the 6th IEEE Conference on Network Softwarization (NetSoft), Ghent, Belgium, 29; pp. 310–318, 2020.
- [58] T. Zahariadis, A. Voulkidis, P. Karkazis and P. Trakadas, "Preventive maintenance of critical infrastructures using 5G networks & drones". In Proceedings of the 14th IEEE International Conference on Advanced Video and Signal Based Surveillance (AVSS), Lecce, Italy, 29, 2017.
- [59] S. Zeadally, E. Adi, Z. Baig and I.A. Khan, "Harnessing Artificial Intelligence Capabilities to Improve Cybersecurity". *IEEE Access*, 8, 23817–23837, 2020.
- [60] A. Zeid, S. Sundaram, M. Moghaddam, S. Kamarthi and T. Marion, "Interoperability in Smart Manufacturing: Research Challenges". *Machines*, 7, 21, 2019.
- [61] W. Zhang, D. Yang and H. Wang, "Data-Driven Methods for Predictive Maintenance of Industrial Equipment: A Survey". *IEEE Syst. J.*, 13, 2213–2227, 2019.
- [62] K. Zhou, T. Liu and L. Zhou, "Industry 4.0: Towards future industrial opportunities and challenges". In Proceedings of the 12th International Conference on Fuzzy Systems and Knowledge Discovery (FSKD), Zhangjiajie, China, 15–17; pp. 2147–2152, 2015.