Sensor Data Management and Stream Data Management: A Conceptual Framework and Its Application in Weather Data Monitoring

[1]Mr. Mathew Daniel E., [2]Dr. Okoronkwo Mathew C., [3]Mrs.Ukeoma Pamela E., [4]Prof. Bakpo Francis, [5]Dr. Udanor Collins N., [6]Dr (Mrs.) Ezema Modesta

[1,3]Students, [2,6]Doctors, [5]Ass. Professor, [4]Professor
[1,2,3,4,5,6]Department of Computer Science, University of Nigeria, Nsukka (UNN), Nigeria.
[1]Centre for Atmospheric Research, National Space Research and Development Agency, Anyigba, Nigeria.

Abstract - **With the development of highly intelligent technology and computing, sensors have generated a stream of data that must be effectively collected, processed, and stored to provide an analytical foundation for crucial decisions that affect the global well-being of humans. This study introduces a comprehensive conceptual framework designed to address the challenges associated with these high-velocity data streams and provides an understanding of sensor data management operations like data ingestion, retrieval, queries, storage, and analytics. The focus was on applying atmospheric weather data using the Campbell Scientific weather station to monitor weather parameters. The Atmospheric Weather Station (AWS) used for this study has sensor data and is managed by the Centre for Atmospheric Research, National Space Research and Development Agency, Anyigba. The database designed for this atmospheric data has a twelve-column record where the sensor data has nine columns for the nine weather parameters (rainfall, air temperature, relative humidity, solar radiation, soil volumetric water, soil temperature, wind speed, wind direction, and barometric pressure) and the equipment data has three records. This work also demonstrated that machine learning models are used to train, test, and predict the sensor data collected thereafter, visualize the predicted weather parameters in real time and the results of this prediction would assist both government policymakers and individual decision-makers in planning and socio-economic growth in Nigeria.**

Keywords: **Sensor data, stream data management, atmospheric weather station, weather prediction, machine learning, Internet of Things**.

I. INTRODUCTION

In recent times, data have been considered the new 'gold' because most business and management decisions depend mostly on the information gathered from the data that were analyzed. Therefore, exploiting, protecting, and managing this asset effectively, efficiently, and promptly are the main concerns of most decision makers. In our era of interconnected technologies and data-driven decision-making, sensor data and stream data have emerged as critical elements in understanding and responding to a rapidly changing world. These data streams, often characterized by their real-time, high-velocity nature, offer an invaluable resource for applications ranging from environmental monitoring to industrial automation. To harness the potential of these data streams, efficient and systematic management is paramount. This research delves into the realm of "Sensor Data Management and Stream Data Management" and introduces a comprehensive conceptual framework designed to tackle the inherent complexities of real-time data. Beyond the conceptual framework, the study demonstrates its practical application in weather data monitoring, highlighting the transformational impact of such management in the field of meteorology and environmental science.

Weather data monitoring, a domain where timely and accurate information is mission-critical, serves as an illuminating backdrop for this study's exploration. The ability to integrate real-time sensor data with historical weather data streams empowers meteorologists, climate researchers, and disaster response teams with enhanced predictive capabilities. However, this study transcends the meteorological context, offering a versatile framework for data management that can be applied across a spectrum of domains where the real-time flow of information holds the key to informed decision-making. It is in this intersection of data management and environmental monitoring that this research unfolds, providing a structured approach to address the challenges and seize the opportunities presented by the dynamic world of sensor and stream data.

[1], presented that most physical components can be observed using sensors to gather data whereas the Internet of Things (IoT) and edge computing heavily rely on sensor data to achieve settings and projects. The majority of IoT components may be equipped with their distinctive identification and the capability to transport data across a network. A significant portion of sensor data provided in large volumes provides a wealth of information that is frequently crucial

for business decision-making. Many firms are now using sensor data analytics to solve this big-data dilemma. [2], claimed that the progress in larger and heterogeneous sensor networks technology assists sensor nodes in monitoring, processing, storing, and transmitting physical parameters to the network's core nodes that enable users to process and analyze sensor data to gain high-level information. Because of their inherent characteristics, sensors are fundamentally constrained by resource constraints such as low storage capacity, low processing power, low battery power, and limited communication bandwidth, which may lower the capacities and performance of sensor networks and cause message loss, disruptions, delays, missed readings, etc., as well as contribute to the low quality of information from sensor data and quality of service, which can eventually have a negative impact [3].
This work covered the limitation on the in-depth understanding of sensor and stream data management by providing a holistic framework for managing sensor data in real-time, which is critical for various applications, particularly in weather data monitoring and environmental sciences. This paper contributes to a more comprehensive understanding of sensor data management in the IoT and environmental monitoring contexts with focus on the area of atmospheric parameter data as sensor data used in monitoring and predicting the weather but was not adequately addressed in the literature reviewed in section two of this paper then gave a background of sensor and stream data management, types and characteristics, principles of operations, components, and challenges of sensor data, and sensor data management operations, such as data retrieval, data ingestion, and sensor data queries, while section three handles areas concerning current trends and section four concentrates on the application of sensor data management in weather data monitoring in atmospheric weather parameter sensor data collection, database design for sensor data, and the adoption of a machine learning model in atmospheric sensor data. Section five covers the knowledge gained, concluding part of the work and future work suggested.

II. REVIEW OF RELATED LITERATURE

[4], referred to a moving sensor-based smart home system as a stream of sensor data, and the processing channel was broken into five processing stages, where each stage border was identified as the input and output streams as well as their data structures. Streaming systems like sensor data need to be properly managed because of the benefits offered to the data that need to be continuously processed. [5], presented that for quality and context information in data stream processing, there are tools to semi-automatically generate stream processing and sensor models that query the sensor data to simplify the task of data quality monitoring and cleaning. These tools model sensor data quality and integrate quality information into stream data processing. The framework that controls the sensor lifetime, data gathering, processing, and visualization of sensor readings are all factors that must be taken into account while managing sensor data [3].
Sensor data management involves the collection, storage, processing, and analysis of data generated by sensors. In contrast, stream data management (SDM) involves the real-time processing, analysis, and management of data streams, which are continuous flows of data generated by various sources such as sensors, social media, IoT devices, and financial transactions. This research domain is important in environmental monitoring, healthcare, industrial automation, etc. The primary goal of sensor data management is to efficiently handle and make sense of the vast amount of data generated by the sensors. SDM systems are designed to handle constantly evolving data, allowing organizations and researchers to extract valuable insights and make informed decisions as data are generated [6]. In the work of [7], managing sensor data and stream data has several challenges due to the nature of real-time, continuous data streams. These challenges include technical, operational, and privacy-related issues like resource constraints, data quality and consistency, data volume and velocity, data security and privacy, and lots more. Addressing these challenges requires a combination of technical solutions, best practices, and adaptability to rapidly evolving technologies and data management strategies.
[8], introduced an economical system that ensures flexibility, portability, scalability, and user-friendly operations by designing a low-cost weather station that is made up of an outdoor module that measures four weather elements (temperature, atmospheric pressure, relative humidity, and wind speed) through their respective sensors. In the work of [9], the need for real-time stream data management technologies is fuelled by the increasing number of sensors and devices in the cloud computing age and [5] also stated that the Data Stream Management System (DSMS) module receives data from sensors and performs queries on the data to compute quality values for the desired measurements. Thereafter, decisions are made based on the applications used to compute the data in terms of quality values and quality requirements. Successful sensor data and stream data management can lead to valuable insights and improved decision-making across a wide range of domains, from IoT and smart cities to industrial processes and scientific research [4]. Here are some sensor data management operations that are considered:

2.1 Sensor Data

Sensor as an input device gives an output as a signal about a certain physical quantity (input). Based on the input signal, applications, conversion method, and material characteristics of the sensor, such as price, accuracy, or range, sensors are divided into many types [10]. [11.12] classified sensors as Active or Passive. Active sensors require an external excitation signal or power source while Passive sensors don't require any external power signal to provide an output signal directly. Sensors are also categorized according to how they are detected: electrically, biologically, chemically, radioactively, etc, [11]. The input and the output of sensors ultimately gave the classification of sensors as either digital or analog sensors.

Digital sensors work with discrete or digital data, whereas analog sensors create an analog output that gives a continuous output signal about the quantity being measured [13]. Digital sensors that are utilized for the transmission and manipulation of data contain digital data. Depending on the kind of physical quantity it is measuring, each sensor operates according to a particular principle where a few sensor properties about the physical quantity are being measured. In the architectural design of sensors, there are four main components to be considered and they are; a sensing unit, a processing unit, a transceiver unit, and a power unit. On the network, sensor data management can be found in the following three stages:

- During creation, the sensor signals are gathered and converted into data.
- Generated data is delivered to computers via network protocols but security, loss-tolerance, and timeliness requirements determine the best transmission technique.
- In storage, sensor data is accessible for use, analysis, and prediction in many ways.

Occasionally, sensor data is delivered in real-time as soon as it is created, while other times, it is held for a while before being sent in batches to its next location. The amount of sensor data communicated and the method by which it is sent can be limited by factors like storage capacity and bandwidth. It's crucial to be aware that large amounts of sensor data are stored in the cloud.

2.2 Sensor Data Management

Data that are generated by sensors are efficiently and securely managed to ensure data integrity, data accessibility, and data analysis at any time. The major sources of these sensor data are IoT devices, industrial sensors, environmental sensors, and lots more. An overview of sensor data management issues discussed in the work of [14] noted that data management includes data collection and on-site management, data communication and transfer to off-site facilities, and data storage while [15], identifies major requirements and solutions for real-time data management because the warehousing method stores data in a central database and since sensor devices are considered as local databases, queries may be performed on the data locally since data collected by sensors must characterize the current state of the environment which are subject to logic and time limitations.

Some of the techniques considered in sensor data management are time series databases, cloud storage, data compression, data format and encoding, data lifecycle management, database management systems, access control and security, backup and disaster recovery, metadata management, scalability planning, monitoring and logging, data retention policies, etc. The Aurora DBMS is a monitoring application for data management where data streams are passed through a data-flow system and then output a stream that can be used by applications [16].

2.3 Data Retrieval

Sensor data retrieval is a process of accessing and extracting data collected by sensors for analysis, visualization, and decision-making which usually involves querying the data storage system based on the readings or datasets collected from the sensors. Data retrieval is an important step adopted in harnessing a clear understanding and value of data generated from sensors. The tools and steps used for data retrieval vary depending on the volume of data, the complexity of the infrastructure used for the sensor data, and the analytical needs of the processes. Data retrieval is vital for monitoring real-time processes, conducting research, and making decisions in data-driven projects.

[17], stated that data retrieval processes such as searching, identifying, and extracting vital data from a database are used in writing data extraction queries or commands. Data retrieved can be stored in a file, printed, or viewed using the appropriate applications or formats. The major steps involved in sensor data retrieval are data source identification, query formulation, data storage system access, query execution, data transformation, data retrieval, data integration, data validation, and quality checks, data storage, data analysis and visualization, data presentation, data security, monitoring and logging, and automation and scheduling.

2.4 Data Ingestion

Data ingestion is the process of collecting, receiving, and importing data from sensors and other sources into a data store or a database or processing system for analysis, storage, and utilization. Sensor data may be transmitted in batches or in real-time, and whenever real-time data is ingested, the required record is updated and the sensor data is released. If data is ingested in batches, a buffer location is created for the records and then transmitted in discrete amounts into

the database at periodic intervals [17]. The necessary steps in a successful data ingestion process are forming information sources, validating each information, and sending the information in the right direction. Making sure that data is digested effectively and with confidence is the main concern when storing a lot of sensor data.
However, once out-of-order or overloaded data is ingested, it causes inaccurate and incomplete sensor data to be pushed into the storage, therefore causing incorrect analysis. [17], further presented that sensor data intake may be insufficient for two reasons: First, it is not considered as important as data science or data analytics. Second, it's believed to be simple. The major steps involved in sensor data ingestion are data source identification, data source connection, data format, and protocol conversion, data validation, data pre-processing, data routing, data storage, data indexing, data enrichment, real-time processing, data logging and auditing, monitoring and alerts, data security, data backup and recovery, and scaling and optimization.

2.5 Sensor Data Queries

Sensor data querying is the process of retrieving specific information or subsets of data from a sensor data store or database based on defined criteria. This querying process allows users, researchers, or applications to access and extract important sensor data for analysis, visualization, or decision-making. A database query is a request for data management software called a database management system (DBMS) to provide information. Since practically any object that can be imagined may be equipped with a unique identifier and be able to communicate over a network. Sensor data querying plays a critical role in extracting valuable information from large volumes of sensor-generated data and the effectiveness of querying depends on the clarity of the query criteria, the performance of the data storage system, and the analysis techniques applied to the retrieved data. Sensor data querying is done through the following steps; data storage system, query formulation (filter conditions, sorting, aggregation, projection), query language (SQL, NoSQL, etc) query execution, result retrieval, data analysis or visualization, data presentation, iterative queries, security and access control, and query optimization.

III. CURRENT TRENDS IN SENSOR DATA AND STREAM DATA MANAGEMENT

Sensor data and stream data management are evolving fields and so many trends are shaping their development and adoption. Some of the key trends in sensor data management are in machine learning and AI, IoT integration, edge and fog computing, real-time analytics, data fusion, time series databases, data security, environmental monitoring, data visualization, and supply chain optimization, etc. while stream data management covers areas like edge computing and IoT integration, continuous intelligence, event-driven architectures, multi-cloud deployments, open source and vendor solutions, data security and privacy, data governance and compliance, real-time analytics in various industries etc.
AI and machine learning techniques are increasingly integrated into sensor and stream data management and enable advanced analytics, predictive maintenance, anomaly detection, and more, enhancing the value of sensor data. Edge computing continues to gain prominence, with more sensor data being processed at the data source. This trend reduces latency, bandwidth usage, and dependence on centralized data processing, making real-time analysis and decision-making more feasible. IoT applications in sensor and stream data management use open-source frameworks and tools, such as Apache Kafka, Apache Flink, and others, as technology continues to evolve and gain popularity.
Time series databases are gaining popularity due to their efficiency in handling time-stamped sensor data as many organizations are adopting time series databases for better storage and querying capabilities while block-chain technology is one of the current trends in data management being explored for enhancing data security, traceability, and trust in sensor data management, especially in the supply chain. The work of [18] focuses on processing, fusion, and analyses of IoT sensor data and other data sources to produce an expert understanding of hidden data patterns for quick decision-making based on real-time IoT sensor data on a variety of challenges like unclean sensor data and a high consumption cost that can find a solution to data processing techniques. [19] confirmed that the continuous spread of IoT systems helps in the transmission of data from various sensors to a remote Operation and Management (OM) server. Technology trends evolve rapidly, and to stay up-to-date with the latest developments in sensor and stream data management, it's essential to monitor industry news, research, and emerging technologies in the field.

IV. APPLICATION OF SENSOR DATA IN WEATHER DATA MONITORING

Global climate change has effects on agriculture, the environment, human health, the economy, society, and many other areas which makes it one of the biggest concerns of the twenty-first century. Therefore, these climate changes and their impacts need to be reduced and controlled if the available weather data are effectively gathered, processed, and offered for effective weather analysis and forecast. [20], reviewed that to ensure accurate decision-making, data gathered from multiple sources must be free of outliers. Other works, such as [21] focused on time series-based weather forecasting using machine learning algorithms, and [22] used some weather sensor data for three locations in India to predict the quality of air the next day using linear regression model as a machine learning algorithm. Note that the primary goal of sensor data management is to efficiently handle and make sense of the vast amounts of data generated by sensors. There are evolving domains with key attention and interest where sensor data are used, like smart homes, environmental monitoring, healthcare, industrial automation, agriculture, and more.

Our focus is on the management of sensor data from the automatic weather station (AWS), which measures solar radiation, atmospheric pressure, air temperature, wind speed, visibility, and other meteorological parameters as the main tool used to monitor meteorological data [23]. Through a variety of communication channels, the weather station's collected meteorological data can be transferred to the computer at the meteorological center. Multiple parameters are measured by the weather sensors, which are a component of weather stations built around a programmable data logger that processes, saves, and transmits the data after processing the measurements from the sensors.

The Campbell Scientific Weather Station is the AWS considered in this paper with a data logger with onboard instructions, a wide operating temperature range, programmable execution intervals of 5 minutes, and ample input channels for commonly used sensors. The sensor device collects data for air temperature, solar radiation, rain rate, relative humidity, soil temperature, wind speed, wind direction, barometric pressure, and volumetric water. Programming can be modified at any time depending on the type of sensor data required.

4.1 Weather Sensor Data Management

Most meteorological sensor data are measured and stored on data loggers. Customizing each weather station allows the data on the data logger till the data is transmitted to another platform or storage system. Data retrieval may use a variety of communication channels that may be blended inside the same network. Radiofrequency, satellite, and on-site solutions such as laptop computers and other storage modules are some examples of telecommunications routes. Automating weather monitoring processes and sensor data management like data collection, retrieval, display, and analysis of data involves software and programming. The software can also manage the retrieval of data automatically from a network of sensors or a single AWS sensor. Robust error-checking guarantees data integrity and gives room for this data usage in AI, machine learning, IoT, and cloud computing technologies.

Figure:1 CAR-NASRDA Atmospheric Weather Station

The Campbell Scientific AWS shown in Figure 1 used for this study is operated and maintained by the Centre for Atmospheric Research, National Space Research and Development Agency (CAR-NASRDA) in the Federal Ministry of Science, Technology and Innovation (FMSTI), Nigeria. The standard station consists of a waterproof container housing a highly dependable Campbell data logger, a 12V battery, and a charge controller. It is a fully configured, solar-powered, automated weather station. The weather station is designed for long-term unmanned operations and is perfect for meteorological, weather monitoring, and climate study applications. The data logger is programmed using the control register basic for the supplied sensors; once completely connected, the weather station will automatically start to take measurements through each of the weather parameter sensors.

4.2 Database Design for Weather Sensor Data

The database of these atmospheric weather stations was designed to collect the following data; time and date, records of data logger, battery voltage, and sensor values (rainfall, solar radiation, air temperature, relative humidity, soil temperature, soil volumetric water, wind speed, wind direction and barometric pressure) which are collected in every 5 minutes unto the data loggers thereafter transformed and stored in a comma separated value formats.
[(Date_time): dd/mm/yyyyhhmm, (CR_record): integer, (Battery_volt): float, (Rain_mm_tot): float, (SlrW_Avg): float, (AirTC_Avg): float, (RH): float, (T107_C_Avg): float, (VW_Avg): float, (WS_ms_Avg): float, (WindDir): float, (BarPress_Avg): integer]

	A	B	C	D	E	F	G	H	I	J	K	L
1	Date_time	RECORD	Batt_Volt_Avg	Rain_mm_Tot	SlrW_Avg	AirTC_Avg	RH	T107_C_Avg	VW_Avg	WS_ms_Avg	WindDir	BarPress_Avg
2	TS	RN	Volts	mm	W/m^2	Deg C	%	Deg C		meters/second	Degrees	mV
3			Avg	Tot	Avg	Avg	Smp	Avg	Avg	Avg	Smp	Avg
4	16/05/2023 11:40	0	13.35	0	972	34.97	31.81	30.54	0.003	0.808	251.9	941
5	16/05/2023 11:45	1	13.4	0	970	35.69	29.1	30.6	0.083	0.514	66.72	941
6	16/05/2023 11:50	2	13.32	0	971	35.92	27.92	30.69	0.083	0.55	66.72	941
7	16/05/2023 11:55	3	13.32	0	974	35.83	29.37	30.77	0.083	0.525	20.45	941
8	16/05/2023 12:00	4	13.39	0	973	36.01	28.9	30.86	0.083	0.84	25.15	941
9	16/05/2023 12:05	5	13.42	0	975	35.44	28.39	30.95	0.083	0.965	82.7	940
10	16/05/2023 12:10	6	13.36	0	977	35.4	26.2	31.04	0.083	0.78	75.45	940
11	16/05/2023 12:15	7	13.38	0	974	35.2	26.97	31.13	0.083	0.66	6.24	940
12	16/05/2023 12:20	8	13.54	0	968	35.76	25.96	31.23	0.083	0.539	24.86	940
13	16/05/2023 12:25	9	13.89	0	966	35.87	27.72	31.33	0.083	0.739	159.8	940
14	16/05/2023 12:30	10	13.96	0	967	35.62	25.99	31.43	0.083	1.004	166.6	940
15	16/05/2023 12:35	11	13.98	0	964	36.25	26.3	31.54	0.083	0.662	234.6	940
16	16/05/2023 12:40	12	14.02	0	962	36.51	24.98	31.65	0.083	0.525	280.5	940
17	16/05/2023 12:45	13	14.06	0	950	36.67	25.45	31.76	0.083	0.903	34.08	940
18	16/05/2023 12:50	14	14.19	0	950	35.96	26.16	31.87	0.083	1.343	51.89	940
19	16/05/2023 12:55	15	14.25	0	945	36.34	24.27	31.97	0.084	0.55	35.33	940
20	16/05/2023 13:00	16	14.29	0	673.9	36.59	24.71	32.08	0.084	1.019	178.3	940

Figure 2: Sample Sensor Data from Bauchi Atmospheric Weather Station

From the data sample above, the header text lines contain twelve columns that begin with date/time in the same cell with the format dd/mm/yyyy hh:mm. The second column is the CR1000 record which is the datalogger type used for data collection in this equipment, the third column is the CR1000 Battery Volt and the fourth to the twelfth columns is the meteorological parameters data starting with Rain Rate in (mm), Solar Radiation SLrW in (W/m2), Air Temperature AirTC in Degree Celsius (°C), Relative Humidity RH in Percentage (%), Soil Temperature T107 in (°C), Volumetric Water Content VW *100, Wind Speed WS in (m/sec), Wind Direction in Degrees (°), and Barometric Pressure Barpress in (mbar).

Based on this AWS, a one-year data volume is up to 1,244,160 if it's generated at a cycle of every 5 minutes for the 12 columns of data, and if the data is monthly it has about 103,680 volumes from one weather station. Having multiple weather stations across Nigeria say up to 37 with at least one in each State and the Federal Capital Territory will amount to a very high volume of data for any kind of weather research depending on the number of parameter sensors available on the AWS. The data collected are used for weather prediction and visualization at any time and location which will assist help in the planning and development of Nigeria towards policy and decision-making for socio-economic growth and in sectors like agriculture, security, budget, etc.

4.3 Adoption of Machine Learning in Weather Sensor Data

[24], presented that low-cost sensors can be used for weather data collection, and with the era of highly-performance intelligent computing technology based on the Internet of Things (IoT) and embedded systems to get real-time data, we can achieve accurate weather prediction and visualization using a network of sensors to accumulate atmospheric weather parameter data. This paper focuses on temperature, rainfall, relative humidity, solar radiation, and pressure because they are responsible for global warming and climatic effects on human health, agriculture, and environmental impact in recent times which has become a vital research area and because weather parameter data are time series data, neural network models was proposed to be adopted for the weather prediction thereafter, results should be compared with any other machine learning model to ascertain the best prediction.
The stream data collected from the sensors are aggregated using data mining techniques to harmonize and remove missing data thereafter trained by machine model for predictions. Just like [25] presented, IoT devices collect and generate high quantities of data for various fields and applications. After predicting the weather conditions, it is published through a web service and a web application of the IoT listens for the prediction visualization.

Adopting machine learning in weather sensor data management can significantly enhance the accuracy and efficiency of weather forecasting, climate research, and environmental monitoring. Some of the steps to incorporate are; data collection and integration, data pre-processing, model selection, training and validation, model deployment, etc. By adopting machine learning in weather sensor data management, organizations and research institutions can unlock the potential of more accurate and timely weather predictions, aiding in disaster preparedness, agriculture, renewable energy, and various other applications that depend on weather information.

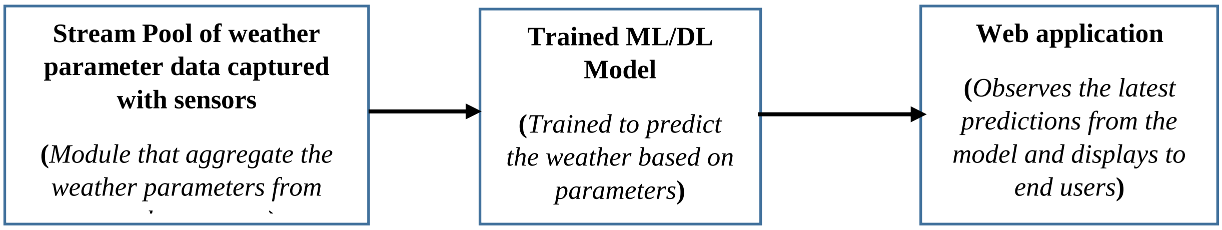

Figure 3: Schematic Flow of the Machine Learning Process

From Figure 3, the weather parameter data would be collected using the sensors on the AWS thereafter the data is trained using a machine learning algorithm then the prediction or deserved outcomes are displayed via the cloud using IoT technologies for access to end users.

Before we arrive at this structure, the machine learning model trains the data. The diagram below shows the training process and the various neural network layers to be used for training the model.

Figure 4: Data Training Layers

From Fig. 4, after pre-processing and standardizing the data from the sensor, the data is inputted into the input layers. The input layer then sends the data to the machine/deep learning algorithm. Then the machine learning model will predict a result in the output layer where the result will be captured.

4.4 Contribution to Knowledge

This research makes significant contributions to the field of data management and environmental monitoring. Firstly, it introduces a comprehensive conceptual framework that systematically organizes the management of sensor data and stream data by providing a structured approach for handling real-time and high-velocity data streams. This framework serves as a valuable resource for researchers and practitioners seeking to design efficient data management systems for a wide range of applications beyond weather monitoring. The paper has also given the basis for applying a machine-learning model to make predictions on weather data.

Secondly, the research applies this conceptual framework to the specific context of weather data monitoring in Nigeria which can be applied to other parts of West Africa by demonstrating the practical utility of the framework in a critical domain. By integrating real-time sensor data with historical weather data streams, the study enhances the accuracy of weather forecasting, disaster preparedness, and climate research. This research not only advances our understanding of sensor and stream data management but also directly benefits the field of meteorology and environmental science, highlighting the broader impact of the conceptual framework in solving real-world challenges.

V. Conclusion and Future Work

Sensor and stream data management play a very important role in socioeconomic development, security, the environment, and many other areas of human endeavours. These data offer opportunities for managing and developing applications that can be applied in many areas of life since a huge volume of data can be collected through sensors and used for adequate information. In this work, sensor data and stream data management have been considered to cover operations like data storage, data ingestion, data retrieval, data queries, and database design for sensor high data. However, there are challenges to sensor and stream data management like infrastructure, storage, query, and data retrieval of information because of the large volumes.

In future works, we suggest designing and implementing a real-time weather data prediction based on sensor data collected from AWS using a machine or deep learning model thereafter, developing a web platform accessible for real-time prediction and visualization of weather data using IoT.

REFERENCES:

[1] Wan, T., Shao, B., Ma, S., Zhou, Y., Li, Q., & Chai, Y. (2022). In‐Sensor Computing: Materials, Devices, and Integration Technologies. *Advanced Materials*, 2203830.

[2] Diallo, O., Rodrigues, J. J., & Sene, M. (2012). Real-time data management on wireless sensor networks: A survey. *Journal of Network and Computer Applications*, *35*(3), 1013-1021.

[3] D'Aniello, G., Gaeta, M., & Hong, T. P. (2017). Effective quality-aware sensor data management. *IEEE Transactions on Emerging Topics in Computational Intelligence*, *2*(1), 65-77.

[4] Jansson, J., & Hakala, I. (2020). Managing sensor data streams in a smart home application. *International Journal of Sensor Networks*, *32*(4), 247-258.

[5] Benabbas, A., & Nicklas, D. (2019, March). Quality-aware sensor data stream management in a living lab environment. In *2019 IEEE International Conference on Pervasive Computing and Communications Workshops (PerCom Workshops)* (pp. 445-446). IEEE.

[6] Wang, D. H. (2020). IoT based clinical sensor data management and transfer using blockchain technology. *Journal of IoT in Social, Mobile, Analytics, and Cloud*, *2*(3), 154-159.

[7] Pike, M., Mustafa, N. M., Towey, D., & Brusic, V. (2019, July). Sensor networks and data management in healthcare: Emerging technologies and new challenges. In *2019 IEEE 43rd Annual Computer Software and Applications Conference (COMPSAC)* (Vol. 1, pp. 834-839). IEEE.

[8] Oaihimiiiire, A.S & Ozigi, B. (2023). *Design and construction of a low-cost weather station, International Journal of Science & Engineering Development Research (www.ijsdr.org), ISSN:2455-2631, Vol.8, Issue 1, page no.361 – 371.*

[9] Ng, W. S., Kirchberg, M., Bressan, S., & Tan, K. L. (2011). Towards a privacy-aware stream data management system for cloud applications. *International Journal of Web and Grid Services*, *7*(3), 246-267.

[10] Patel, B. C., Sinha, G. R., & Goel, N. (2020). Introduction to sensors. In *Advances in Modern Sensors: Physics, design, simulation and applications* (pp. 1-1). Bristol, UK: IOP Publishing.

[11] Fraden, J., & Fraden, J. (2016). Humidity and moisture sensors. *Handbook of Modern Sensors: Physics, Designs, and Applications*, 507-523.

[12] Roh, S., Chung, T., & Lee, B. (2011). Overview of the characteristics of micro-and nano-structured surface plasmon resonance sensors. *Sensors*, *11*(2), 1565-1588.

[13] Konwar, S., & Sahoo, B. D. (2023). Johnson Counter Based Multiphase Generation for VCO Based ADC for Direct Digitization of Low Amplitude Sensor Signals. *IEEE Transactions on Instrumentation and Measurement*.

[14] Law, K. H., Smarsly, K., & Wang, Y. (2014). Sensor data management technologies for infrastructure asset management. In *Sensor Technologies for Civil Infrastructures* (pp. 3-32). Woodhead Publishing.

[15] Bao, Y., Ren, L., Zhang, L., Zhang, X., & Luo, Y. (2012, July). Massive sensor data management framework in cloud manufacturing based on Hadoop. In *IEEE 10th International Conference on Industrial Informatics* (pp. 397-401). IEEE.

[16] Silva, D., Ghanem, M., & Guo, Y. (2012). WikiSensing: an online collaborative approach for sensor data management. *Sensors*, *12*(10), 13295-13332.

[17] Sangat, P., Indrawan‐Santiago, M., & Taniar, D. (2018). Sensor data management in the cloud: Data storage, data ingestion, and data retrieval. *Concurrency and Computation: Practice and Experience*, *30*(1), e4354.

[18] Krishnamurthi, R., Kumar, A., Gopinathan, D., Nayyar, A., & Qureshi, B. (2020). An overview of IoT sensor data processing, fusion, and analysis techniques. *Sensors*, *20*(21), 6076.

[19] Yoon, G., Choi, D., Lee, J., & Choi, H. (2019). Management of IoT sensor data using a fog computing node. *Journal of Sensors*, *2019*.

[20] GS Campos, N., Rocha, A. R., Gondim, R., Coelho da Silva, T. L., & Gomes, D. G. (2019). Smart & green: An internet-of-things framework for smart irrigation. *Sensors*, *20*(1), 190.

[21] Chavan, G., & Momin, B. (2017, February). An integrated approach for weather forecasting over the Internet of Things: A brief review. In *2017 international conference on I-SMAC (IoT in social, mobile, analytics and cloud) (I-SMAC)* (pp. 83-88). IEEE.

[22] Kumar, R., Kumar, P., & Kumar, Y. (2020). Time series data prediction using IoT and machine learning techniques. *Procedia computer science*, *167*, 373-381.

[23] Li, C., & Sun, X. (2018). A novel meteorological sensor data acquisition approach based on unmanned aerial vehicles. *International Journal of Sensor Networks*, *28*(2), 80-88.

[24] Ioannou, K., Karampatzakis, D., Amanatidis, P., Aggelopoulos, V., & Karmiris, I. (2021). Low-cost automatic weather stations in the Internet of Things. *Information*, *12*(4), 146.

[25] Garg, D., & Alam, M. (2020). Deep learning and IoT for agricultural applications. *Internet of Things (IoT) Concepts and Applications*, 273-284.