IoT-Based Environmental Monitoring and Smart Agriculture Prediction System Using Machine Learning and Mobile Integration

R. M. N. Rathnayaka and A. Mohamed Aslam Sujah

Abstract Environmental monitoring and prediction are critical for fostering sustainability and ensuring effective resource management. However, existing systems often face challenges such as fragmented approaches and limited detailed solutions. Also, traditional monitoring methods often fail to provide timely and accurate insights into key environmental factors. This research aims to integrate IoT, machine learning and mobile application integration to develop an advanced system for environmental intelligence. The proposed system employs a robust IoT infrastructure comprising sensors to monitor parameters such as temperature, soil moisture, smoke levels, and humidity, all integrated with an ESP32 microcontroller. Data collected by these sensors is streamed in real-time to Firebase for storage and processing. A supervised Artificial Neural Network (ANN) model was trained to predict three key environmental conditions with a high accuracy of 99.32%, showcasing the model's reliability and effectiveness. A mobile application was also developed to facilitate real-time monitoring, predictive insights, and enhanced user interaction. The findings highlight the system's ability to deliver accurate and timely environmental insights, enabling proactive decision-making and mitigating potential risks. By bridging the gap between fragmented methodologies and the need for timely insights, this research contributes significantly to advancing environmental intelligence and promoting ecological resilience.

Index Terms **IoT Development, Environmental Monitoring, Machine Learning, Mobile Application, Predictive Analytics**

I. INTRODUCTION

A. Background

THE *Internet of Things (IoT) is the key that changes the* way people think and interact with the surrounding world in the rapidly evolving technological world. Monitoring and prediction of the environment are one of the most important aspects of the IoT in terms of ensuring a sustainable future. The reason is that individuals are realizing the need to raise their standard of living to a particular standard whilst on the other hand, good environmental management is seen as one of the conditions that human beings need to live a better life. The smart systems of environmental monitoring and predictions are gradually gaining recognition as an effective contributor to physical and mental health, in the meantime, the existing ones are still experiencing significant shortcomings and limitations, particularly in the area of real-time monitoring and dependable prediction. In addition, these environmental factors are unpredictable, which presents a lot of challenges. As an example, the un-regular checking of the moisture level of soil may cause over-irrigation or under-irrigation of the agricultural land, which translate to wastage of resources and reduced agricultural produce. Likewise, insufficient air quality, which is worsened by industry emissions and automobile pollution, may cause serious health problems to society.

R.M. Nayanajith Rathnayaka is a Lecturer (Probationary) at the Department of ICT, University of Ruhuna, Sri Lanka. (Email: mailtonayanajith@gmail.com)

A.M.A. Sujah is a Lecturer (Probationary) at the Department of ICT, South Eastern University of Sri Lanka, Sri Lanka. (Email: ameersujah@seu.ac.lk)

Such factors lack proper and timely predictability thus these sectors are inefficient [1].

There are several benefits of predicting the environmental factors especially in enhancing proactive decision and sustainability. Proper predictions allow those with vested interests to foresee the changes and implement changes in their strategies. This helps in improving the management of resources and also the impact on the environment is minimized thus contributing to the sustainability objectives. Also, forecasting the environment conditions may be instrumental in protecting the health of the people by giving a warning to the people. Nevertheless, environmental prediction has some problems. One of the principal problems that were discussed in previous research is the inconsistent work of the IoT-based environmental monitoring systems. This has been mainly due to the dissimilarity in the technological implementations that have been applied in diverse researches. These differences can be in the kind and specifications of sensors used, the system of communication or data-transfer chosen, the machine learning algorithms that are evolved into the analysis process, the different environmental conditions to which these systems are mounted. All these distinctions lead to the differences in accuracy, responsiveness, and reliability of the entire system [2].

In order to solve these problems, recent studies have concentrated on the more integrated architectures of monitoring that involve the combination of the cloud computing, intelligent sensing modules, and the advanced machine learning algorithms. When combined with other technologies, the information derived could be used to make predictions and provide constant data collection when it would be unable to provide the same and predictive data with traditional monitoring methods. This integration helps with

ISSN: 2961-5992

real-time monitoring and prompt identification of irregularities or changes in the environment. These features can be particularly useful in areas of priority, like climate change analysis, pollution control, agriculture optimization, and disaster forecasting where timely and reliable information has a direct impact on decision-making efficacy [3].

Such advanced surveillance tools are based on machine learning. Out of the numerous approaches that were trying, the Artificial Neural Networks (ANNs) and Random Forest Regression always exhibit high predictive power and stability. ANNs especially can find non-linear trends and fine correlations in noisy or incomplete conditions of data that usually prevail in the environment. As an illustration, inter-sensor relationships and environmental trends are examined by ANNs to identify anomalies or malfunctions of sensors. This is particularly critical in uses such as air quality monitoring where any minor deviation in the readings could indicate major environmental variations. With the development of the research, more complex models are likely to enhance predictive effectiveness and the overall reliability of the system which will allow to obtain even more accurate and actionable information concerning the environment [4][5].

The usefulness of the combination of IoT, mobile technologies, and machine learning is evident due to the available real-life applications. As an example, a thermal monitoring of underground cable tunnel using IoT has shown how the continuous analysis of temperatures could be beneficial in enhancing infrastructure safety and efficiency. In the agricultural sector, greenhouse monitoring systems based on IoT have the benefit of delivering real-time data on various important environmental factors, including temperature, humidity, and soil moisture, and enable farmers to make decisions that have a direct positive effect on crop performance. The mobile applications enhance these advantages by providing the end user with remote access to live sensor data, notifications and operational controls, making environmental monitoring more approachable and interactive. Moreover, current smart-agriculture solutions are becoming more and more focused on predictive analytics and data-fusion capabilities to provide more specific recommendations and help to optimize resources management [6].

In general, the paper discusses the important intersection of IoT, advanced machine learning, and mobile app development through environmental intelligence. The proposed system will be an important step towards a sustainable and informed future as sensor data, cloud computing, and predictive analytics are seamlessly integrated. The versatile use of this system makes it a possible substitute of the current solutions in the market in addition to offering a framework that various industries can avoid succeeding in the age of data-driven decision-making.

B. Problem Statement and Research Questions

The present situation of the environmental monitoring and prediction systems are hard to navigate because of the disjointed schemes and insufficient detailed solutions. Conventional methods of monitoring do not usually give the right and prompt information on key environmental parameters such as temperature, soil moisture, smoke concentration, and humidity [6]. The fact that these environmental factors are unpredictable presents a lot of challenges. As an example, monitoring soil moisture at an irregular cycle may cause either over-irrigation or under-irrigation in farmlands both of which lead to the wastage of resources and reduced production. The same way, low quality of the air, which is worsened by the industrial emissions and the traffic flows, may cause serious problems in the population health. Such factors are inefficiently managed in these sectors in case they are not correctly and timely predicted.

The following questions are directly related to this research and they are regarded as the research gaps. The majority of the research questions will have to be examined on the topic of The Current Existing Environment Monitoring Systems, Its Workflow and Limitations. The identified research questions can be explained as,

RQ1: “What are the fragmented methods currently used in different environmental monitoring systems be combined to develop a better system than existing systems?”

RQ2: “What specific functionalities or features should the predictive framework accomplish to improve existing environment monitoring systems?”

RQ3: “What are the problems or issues that are in the traditional systems to be overcome to implement the introduced system?”

There is a lack of an integrated IoT-based Environment Monitoring and Smart Prediction System, which is quite problematic. The environmental stakeholders, which are farmers, researchers, miners, and tourism operators do not have immediate access to correct data to plan proactively. The existing solutions are unable to manage the dynamics of the environmental changes effectively and thus hamper effective allocation of resources and planning [7]. In addition, the modern environment of environmental data user interfaces is characterized by complexity and inefficiency. This makes the data gathered unintuitive, which impedes the access to and interpretation of this data. This is exacerbated by the fact that differentiated smart systems continue to expand to only provide partial and not a complete and unified experience [8].

Though the current smart systems are promising, they are not integrated and do not have a single platform of collecting, analyzing, and predicting data. Almost all the existing standard environment monitoring systems do not analyze the results and simply present them. This disjointed state of affairs presents a significant disparity in assisting well-informed decision-making in the industries such as agriculture, Bio research, mining and tourism [9]. Also, the fact that one can not reliably forecast the outcomes does not allow active decision-making. The predictive capability in the future developmental changes of the environmental traits is limited by the fact that it relies on the simplistic forecasting methods. Due to such shortsightedness, there are obstacles in implementing efforts to reduce the impact of changes in the environment in terms of planning and adaptive strategies [10].

According to recent research, the use of isolated monitoring solutions is changing to integrated architectures composed of cloud platforms, intelligent sensing modules, and machine learning analytics. These technologies do not work as independent elements, but instead, they work in

ISSN: 2961-5992

unison to facilitate sustained data collection, situational awareness, and real-time prediction of the environmental status. This development has allowed supervisory mechanisms to find patterns in sensor data that can be acted upon, in huge amounts of sensor data, which previously could not be analytically supported due to the rule-based methods. In that regard, the suggested system aims at creating an integrated monitoring and predicting framework that will facilitate evidence-based and timely decision-making. This type of framework is especially applicable in areas where slow or poor environmental data can result in ineffective response, such as assessment of climate variability, reduction of pollution, agricultural control, and disaster preparedness.

II. LITERATURE REVIEW

A. Environment Monitoring Systems

According to recent advances in the field of environmental monitoring, there is an increasing shift to data-driven systems which integrate sensing technology with smart analysis models. Instead of being based purely on traditional monitoring methods, modern solutions are becoming more dependent on the relationship between Internet of Things (IoT) infrastructures and machine learning to enhance the environmental awareness and sustainability performance. This change is also significantly driven by the increasing need to have a continuous environmental evaluation to uphold the health of people and the ecological equilibrium and effective management of resources [11][12].

Nevertheless, empirical studies indicate that a number of the current implementations are unable to sustain reliability during real-time operation and predictive power. To counter this, recent studies have been looking at the improvements on architecture so as to incorporate cloud-based platforms, distributed sensor networks, and learning advanced algorithms. These architectures provide continuous data collection and speedy analytical data processing which allows the monitoring systems to react to the dynamism of the environment. The capabilities are especially important in the fields of application like climate analysis, pollution control, agricultural optimization, and disaster response, in which a slow or inaccurate information source can have serious consequences on the outcomes of decisions [13][14].

B. Predictive Machine Learning Algorithms for Environmental Factors

Modern environmental monitoring systems are also becoming more reliant on capabilities to understand complex and highly variant data streams. In this regard, machine learning approaches have been greatly embraced to increase the degree of analytic as well as predictive accuracy. Other than linear or rule-based models, recent research has shown that learning-based systems, especially Artificial Neural Networks and the ensemble systems like the random forest regression are more effective in modelling complex relationships between sensor generated data [15][16].

Through these models, the monitoring systems can analyze the relationships between various environmental parameters and hence detect the abnormal patterns as well as conduct more accurate forecasting. This feature is particularly useful in the air quality analysis, where slight changes in sensor data can indicate dangerous environmental changes. The current progress in machine learning methodologies is likely to see newer frameworks of algorithms to enhance adaptability, accuracy, and scalability of systems, with the end result being the ability of monitoring platforms to provide more reliable and decision-oriented environmental intelligence [17][18][19].

C. Smart Data Gathering for Environmental Monitoring

Combining IoT-based data collection systems with machine learning analytics have been implemented in a very broad spectrum of application areas, including residential monitoring, agriculture, industrial processes, and urban air quality monitoring. These systems allow monitoring of the environmental conditions around the clock in an indoor environment, which enhances the comfort and management of health. Mobile interfaces further expand the functionality of the system as they enable the visualization of sensor readings remotely, issue alerts, and communicate with monitoring platforms in real time to the user [20][21].

Recent technological advances indicate that such systems can be improved by adding more sensing modalities, e.g. soil chemistry testing and more sophisticated air pollution sensors. Simultaneously, image-based methods of analysis have also received interest due to their capacity to contribute to the early detection of abnormalities, especially in the agricultural setting. These features can result in the early identification of crop stress and deviations in the environment, which allows taking proactive action and optimizing the utilization of the available resources [22][23]. Although these improvements have been made, the literature equally records enduring problems which pertain to system reliability and scalability. Differences in sensor specifications, communication mechanisms, learning algorithms and deployment conditions often cause such inconsistency in performance on implementations. Further, most of the current solutions operate on the environmental parameters individually limiting their capacity to provide an overlay and holistic depiction of environmental conditions [24].

D. Sustainable Preservation Plans for Multiple Fields

Sensing platforms based on IoT integration along with machine learning analytics have emerged as a crucial enabler in terms of environmental management under sustainability. Ongoing control of such parameters as air and water quality, soil conditions, and ecosystem indicators enables institutions to assess the environmental health more precisely and plan the design of data-driven preservation plans. The availability of the real time environmental information will also be useful in minimizing ecological degradation and long-term conservation planning [25][26].

In the agricultural industry, predictive learning-based smart solutions to monitor the water resource, the soil conditions, and the performances of crops enable the efficient management of such company activities. Data-centric versions of similar techniques have been generalized to the industrial and urban environment setting, where the concept of environmental intelligence is applied to track energy consumption, noise, and atmospheric quality. In conservation-oriented solutions, IoT technologies can be used to conduct systematic monitoring of animal populations and the activities of ecosystems to ensure that the protection system is informed by timely and credible data instead of making periodic periods of evaluation through manual

ISSN: 2961-5992

TABLE I

AN OVERVIEW OF COMPONENTS AND TECHNIQUES OF SELECTED SMART SYSTEMS

System Type	IoT Components	Machine Learning Techniques
Smart Home Monitoring	Smart sensors, actuators, connectivity devices	Pattern recognition, anomaly detection
Industrial Environmental	Industrial sensors, communication infrastructure	Predictive modeling, clustering analysis
Agricultural Monitoring	Smart sensors, actuators, communication devices	Predictive modeling, crop yield prediction
Urban Air Quality Monitoring	Air quality sensors, data integration platforms	Neural networks, regression analysis
Wildlife Habitat Monitoring	Tracking collars, satellite tags, environmental sensors	Species identification, habitat mapping
Water Quality Monitoring	Water quality sensors, communication networks	Data analytics, anomaly detection
Traffic-Related Pollution	Traffic sensors, GIS mapping, communication infrastructure	Regression analysis, traffic flow modeling
Disaster-Prone Area Monitoring	Remote sensors, drones, communication networks	Image recognition, pattern analysis
Energy Consumption Tracking	Smart meters, communication networks	Predictive modeling, optimization algorithms
Noise Pollution Monitoring	Acoustic sensors, communication infrastructure	Noise pattern recognition, data analytics
Green Building Management	Building automation systems, energy meters	Predictive maintenance, energy optimization
Coastal Erosion Monitoring	Coastal sensors, satellite imagery, communication networks	Image analysis, pattern recognition
Air Quality in Public Spaces	Personal IoT devices, mobile sensors, communication networks	AI-based crowd analysis, anomaly detection

analysis [27][28][29]. Outside these sectors, the activities of industries with an environmentally sensitive situation, such as tourism, aquaculture and mining, are becoming more and more dependent on telemetry-based round-the-clock monitoring as a means of ensuring ecological balance. These systems assist in the operations by helping to proactively address changes in the environment before they happen, which leads to sustainable operations. Together, the combination of IoT and machine learning creates a flexible and scalable base of environmental monitoring systems that unite technological progress and the responsible use of resources [30][31].

III. METHODOLOGY

A. Methodology selection

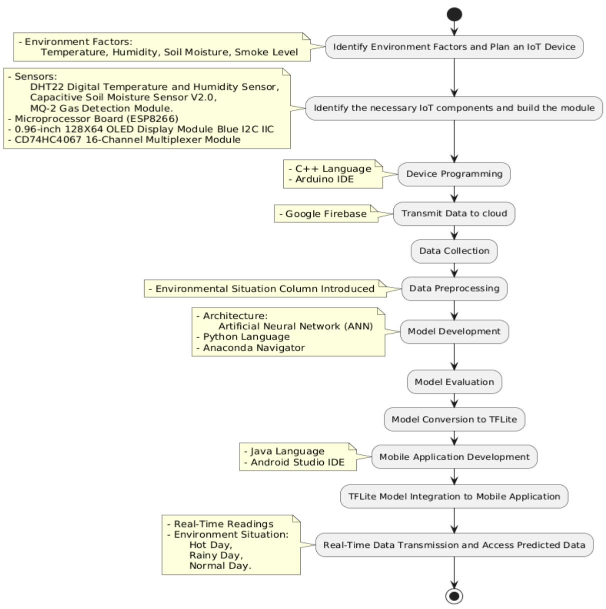

Fig. 1: System Flow Diagram

The initial stage of the research is concerned with the implementation and streamlining of the IoT infrastructure. This will include the proper selection of sensors and calibration so that they can collect information correctly and reliably. The hub is the ESP32 module, which is highly efficient and versatile, and coordinates with the smooth interaction between sensors and the cloud. The dynamism and responsiveness of environmental monitoring system is established with the real-time data that is sent to the cloud. The proposed system is based on the combination of various sensors that can be able to measure various key environmental parameters like the temperature, the level of soil moisture, air pollution, and humidity. These sensors shall be carefully installed, coded and wired to a microcontroller module, which forms the basis of IoT-based framework. The real-time information that will be produced by these sensors will be easily uploaded to a separate IoT cloud services, such as platforms to Firebase, which offers a centralized storage of all-encompassing environmental data.

The second stage of the research explores the world of advanced machine learning methods with a strong IoT infrastructure established. The objective is to exploit the strength of supervised machine learning, deep learning algorithms to process the dataset gathered. Developed model attempts to foresee future predictions of the environment by examining patterns and correlation of the data. Machine learning as a component of the system is a major innovation which allows building a predictive model that transcends monitoring and data reporting. This foresight ability equips the stakeholders with knowledge about the future environmental situation. The system serves as an early warning system, which enables the user to predict the changes and make knowledgeable decisions in order to be sustainable. Through the synergies of the various algorithms, the system seeks to address the limitations of any of the models, which are likely to subject the system to more predictive robust and reliable framework.

ISSN: 2961-5992

The importance of actionable insights can hardly be overestimated as the path towards a sustainable future is taken. The last step in the research process is to provide the users with the predicted environmental data in an easy to use android mobile application. This is an interface that allows users to read, formulate and take action on the information given by the machine learning system. The proposed mobile application will be aimed at a wide range of users who will base on real-time environmental conditions. The applicability of the system can also be applied to the smart agriculture and smart greenhouse monitoring, as well as all in one detection and prediction module system, developing a flexible and adaptable system to different industries.

B. IoT Device Development

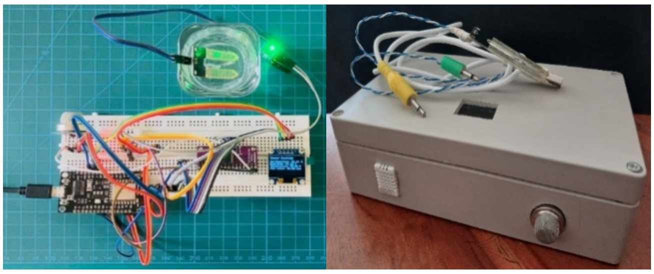

Fig. 2: IoT Device

The IoT-based environmental monitoring system is developed considering the close selection of the components and the orderly workflow to guarantee the lack of interruptions and a smooth-flowing system. A NodeMCU ESP8266 microcontroller is the main part of the system because of the in-built Wi-Fi interface, low prices, and ability to implement various sensor integration. It is the main data aggregation node and it can collect, process and send environmental data to the cloud. This Wi-Fi based microcontroller offers the infrastructure needed to monitor and transmit real-time data thus it is an appropriate selection to the IoT based project.

To record the changes in the ambient environmental conditions, the system will have a series of sensors that will be chosen on the basis of measuring the environmental conditions and their ability to work over a long period of time. To achieve the temperature and humidity, it was decided to use DHT22 sensor because of its stable range of work over a long range. Instead of being a measurement unit, the sensor is a component of a sensing system based on the NodeMCU, which allows acquiring and transmitting data in real time. The obtained readings are periodically sent to the cloud system, which will store them and provide them to further analytical processing.

The Capacitive Soil Moisture Sensor V2.0 is used to monitor the soil moisture. This sensor is also preferred to resistive sensors since it does not corrode thus giving a long-term stability and accuracy. Monitoring of soil moisture is of vital importance to applications like precision agriculture where crops require optimal soil moisture to stay healthy. The sensor is placed in the soil at the required depth, and it measures the moisture content at periodic intervals which is relayed as a voltage to the NodeMCU that in turn sends the data to the cloud.

To monitor the air quality, it is fitted with the MQ-2 gas sensor. This sensor provides the detection of flammable gases like LPG, methane, carbon monoxide and hydrogen. It is very sensitive and offers fast response time hence it is a good option to use in detecting the dangerous gases in the environment. This sensor is employed in the project to detect the air pollution level particularly the identification of the smoke and dangerous gases that may impact on the environment and human health.

It has a 0.96-inch OLED display which is provided to enable real-time sensors visualization on the device. The display is small but efficient to show real-time readings of the sensors associated with it, and one can use it to check the conditions in the environment without the necessity to connect to the data in the cloud. It is also compatible with I2C communication and therefore can easily interface with the NodeMCU. The CD74HC4067 16-channel multiplexer is needed to address several sensors using a small number of GPS pins on the NodeMCU. The device enables the connection of several analog and digital sensors to one microcontroller input, which is optimal in pin utilization. It is also important in integrating a number of sensors into the system, as this would enable and guarantee effective data collection without clogging the input capabilities of the microcontroller. The enclosure is very small; this guarantees protection against the environment risks and an orderly design so that it is easy to use. The power supply is based on 5V 1A power source, stability and safe performance of the system.

Arduino IDE is used to program the NodeMCU ESP8266 which is a popular platform that provides a range of libraries that can be used to interface with the chosen sensors. The development process is started with the establishment of the Arduino IDE with ESP8266 board support. This includes the addition of the ESP8266 board manager URL in the Arduino IDE preferences followed by the installation of the necessary board support packages. The code is coded in the Arduino programming language (written in C++) and it performs the task of initializing the sensors, Wi-Fi connectivity and sending data to Google Firebase Realtime Database. The first step involves the initialization of libraries of each sensor and sensor connections. After connecting to the sensors, the NodeMCU is trying to get connected to the Wi-Fi network using the credentials (SSID and password) specified in the code. After the Wi-Fi connection is made, the device is connected to the Firebase Realtime Database with the help of the Firebase credentials. These credentials comprise the Firebase host URL and the database API key that are sure to keep the communication with the cloud platform safe.

Once connected to Firebase, the NodeMCU will routinely read the sensors, work with the raw data, and package it to send to Firebase. The data of every sensor is stored in a separate node of the database and therefore can be readily accessed and can be monitored real-time. The workflow of the system will consist of the following steps: activation of the sensors and modules, Wi-Fi connection, verification of and Firebase connection, sensor data collection, sensor data processing, and the forwarding of the processed sensor data to Firebase. The OLED display periodically shows the current sensor values, which means that one can monitor the current environmental state on the device. Firebase is selected because of its capability to offer real-time synchronization of data on an array of devices and platforms. It is free, scalable,

ISSN: 2961-5992

as well as user friendly and provides a data storage and retrieval platform. Firebase would also be used in this project to handle sensor data in real-time so that the most up-to-date readings are readily available to monitor and analyze them. The information is constantly updated, and people can get real-time environmental information via a mobile application or other connected equipment.

C. Machine Learning Model Development

Data Gathering

Data collection is a very important phase in the machine learning process because the quality and completeness of the data set greatly influences the accuracy of the predictive models. The IoT device network that has been created before is efficient in managing the data gathering process in this project.

These IoT devices comprise a set of different sensors with the ESP8266, which constantly observe the conditions of the environment in terms of temperature, humidity, and soil moisture, as well as smoke concentration. These sensors gather real-time information at 10 seconds. After the data is stored in firebase, it must be exported to be analysed through additional machine learning. Firebase can be used to export the data stored in the format of JSON. In order to transform the JSON file into CSV, an online tool, which is, "https://www.convertcsv.com/json-to-csv.htm" is utilized. The data is transformed into a CSV file as fast as possible and in the most accurate way by uploading the JSON file to this site. The resulting CSV file has all the collected data in a table format, which was easy to manipulate and analyze. This is a structured format that is ideally fed to machine learning algorithms and allows prediction models to be created.

Data Preprocessing

In order to make the dataset rich, a new column was added to the dataset called Environment Situation. The data is categorized into three under changes in weather conditions in the present surrounding as rainy day, hot day and normal day in this column. The classification was executed by exploiting the gathered data and the information of the external daily weather forecasting to identify the environmental context. The introduction of the column of "Environment Situation is important to the feature engineering aspect of model building. With the inclusion of such contextual information, the machine learning model will be able to understand and predict the said environmental conditions.

Model Development

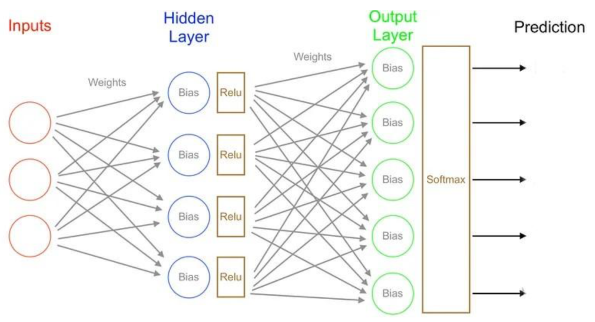

Fig. 3: Prediction Model

In developing a model, Artificial Neural Network (ANN) was used, with a sequential neural network architecture. ANNs are based on the neural structure of the human brain and are great in the modeling of complex relationships between inputs and outputs. They can represent non-linear relationships, and that is why they are suitable to a wide range of prediction applications. The advantage of ANNs is that they can operate on the available data and become better or specific with the additional data, and this makes it useful in continuous monitoring and prediction in the environmental systems.

A sequential neural network refers to a form of ANN, in which the data moves in one direction (input layer all the way to the output layer). It is a simple and easy-to-train architecture suitable to time series and structured data. Sequential neural networks are especially useful with complex and structured data and therefore are useful in monitoring the environment where the patterns of data can be complicated. These models are characterized by great prediction accuracy, flexibility, and scalability. They are capable of learning features of the input data automatically and extracting the relevant features, eliminating the requirement of manual feature engineering, and are resistant to noisy data, which is typical when it comes to real-world environmental monitoring. Through the merits of ANN having a sequential architecture, the project will create a sound and precise predictive model that can give meaningful information under the environmental conditions and eventually lead to making more informed decisions.

Model saving and Conversion

Then the final model was saved in the Keras format. As a high-level neural networks API, Keras is also built atop TensorFlow and is therefore a perfect option to develop a model. Using Keras makes the development of models simpler due to a simple method of creating and training neural networks, which uses less code. It is very modular and easily experimentable and customizable. Moreover, Keras works well with TensorFlow and is compatible with it, being optimized in terms of performance.

In order to fit Keras model to mobile application, it was translated to TensorFlow Lite (TFLite) format. TFLite is mobile and embedded device-friendly and offers a lightweight implementation that does not compromise the performance and accuracy of the model. TFLite models are faster and efficient, which means that they can be used to make real-time inferences on a mobile device with limited computational capability. The conversion process is also less complicated and will make sure that the model is easily compatible with mobile applications, so that it can be deployed effortlessly and provide environmental monitoring and predictions in real-time to the smartphones of the users.

D. Mobile App Design and Development

Within the workflow of environmental intelligence system, the design and development process of mobile app have critical roles, which guarantee continuous access to insights with value. Overall, the mobile application design and creation process is to achieve a compromise between the accessibility of the real time environment readings and the model prediction developed to the users. Android Studio was selected as the integrated development environment (IDE) in

ISSN: 2961-5992

the case of the mobile application development. Android Studio is the official Android IDE, which offers a powerful and a well-rounded Android-specific set of tools to develop Android applications. Android Studio also supports well the TensorFlow Lite and there is also the built-in firebase database integration options as well which is essential in incorporating the machine learning model into the mobile application which also shows the user the real time monitoring of the IoT device. TFLite is mobile and embedded friendly, such that the model can operate effectively on Android devices that have low computing capabilities.

The user interface of the application will be interactive and easy to use. After opening the app, the user is presented with a splash screen that has the app icon. The app will then take one straight to the main dashboard where real-time sensor readings and model prediction are shown. Besides, there is an about research button on the main dashboard. The key dashboard consists of two parts with the first part showing real-time soil moisture, temperature, smoke, and humidity levels, and the second part will be able to give users up-to-date environment information. The second section, which is labeled as Prediction, shows real-time predictions of machine learning model, which provide insight into future trends using available data. There is an About Research button which gives a short description of the research and project. Besides, there is a back button that takes users to the main dashboard that is contained in the About Research section.

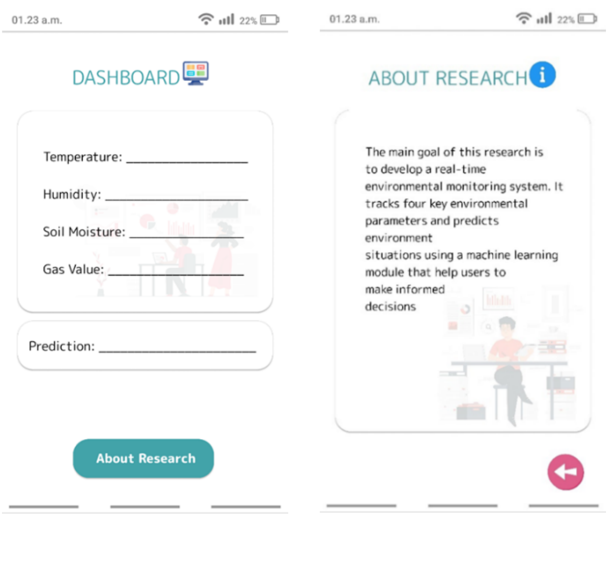

Fig. 4: Dashboard

IV. Result and Discussion

A. Model Architecture and Analysis

The proposed system has a predictive capability that is attained by the use of a multi-layer neural network which is used to model the relationship between the key environmental variables. Instead of concentrating on architectural richness, the model is designed in a way that it balances the representational capacity and the computational efficiency. The environmental parameters such as humidity, smoke concentration, soil moisture, and temperature are also given as the input features to represent a wide range of atmospheric and ground-level conditions. Pattern learning and feature transformation are executed using a series of hideous layers whose neuron count diminishes, this enables the network to refine information in a hierarchical manner. At every stage, non-linear activation is added with the help of the Rectified Linear Unit function, which allows the model to discover intricate interactions among the data without making the gradient flow change. The network makes categorical predictions that are different environmental conditions at the decision stage. The use of a probabilistic formulation of output is used to normalize the prediction scores in order to measure the likelihood of a class and to categorize hot, normal, and rainy environmental states with a quantifiable degree of confidence.

The Adam optimizer (a combination of AdaGrad and RMSProp) gives each parameter its own learning rate, making it an effective optimization algorithm in large datasets and high-dimensional models. Sparse categorical cross-entropy is a enhancer of multi-class classification with class codes being integers, the difference between actual targets and predicted probabilities is calculated. The model is trained in 50 epochs with different training and validation sets. The data is separated into 80 percent training and 20 percent validation to analyze the results on the unseen samples. During training, the validation outcomes are saved at every epoch to track the progress and ensure that the model is generalizable. The process helps in assessing predictive accuracy and reducing overfitting that provides a consistent estimate of the overall ability of the model. Such assessments are used to create a stable learning behavior and reliable performance in diverse environmental datasets over time.

B. Model Accuracy and Performance

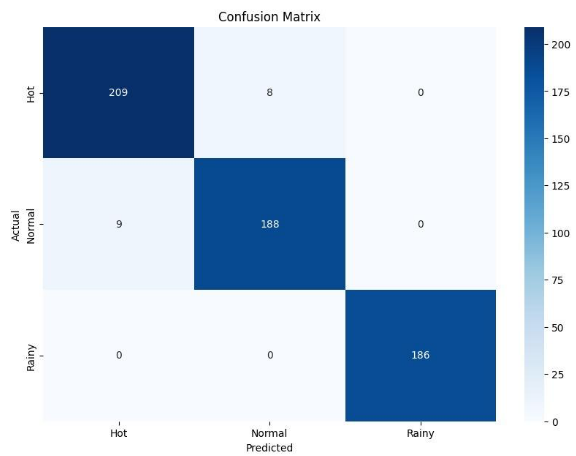

Fig. 5: Confusion Matrix before Remove Misclassified Rows

ISSN: 2961-5992

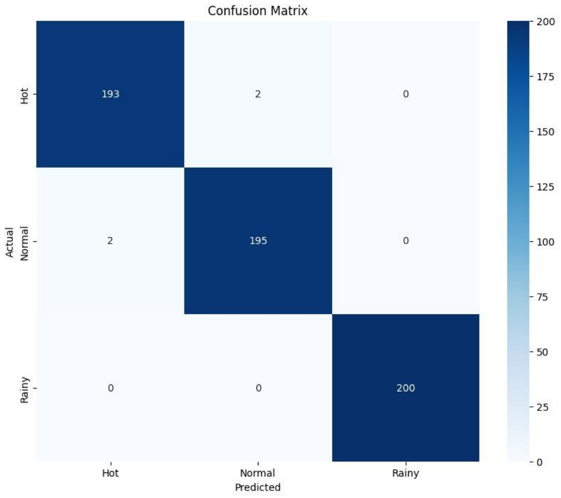

Fig. 6: Confusion Matrix after Removing Misclassified Rows

Confusion matrix is a performance indicator of machine learning classification problem in which the output may be two or more classes. To further improve the use of the model, I deleted the misclassified rows within the training data and re-trained the model. The step is used to make the model more refined by removing the cases that could be called noisy or mislabeled.

Figure 6 represents the confusion matrix after sorting out the misclassified rows and has fewer misclassification values. This means that the model has been made sturdier and more precise following the purification of the data. Figure 7 and Figure 8 represent the training and validation accuracy and the loss after 50 epochs. The training and validation accuracy are 97.09 and 99.32 respectively. In the same manner, the training loss goes down to 0.0629 and the validation loss reduces to 0.0302. The near parallel correlation between training and validation accuracy, as well as the declining loss numbers, is a sign that the training model is well-trained, and the overfitting is not significant.

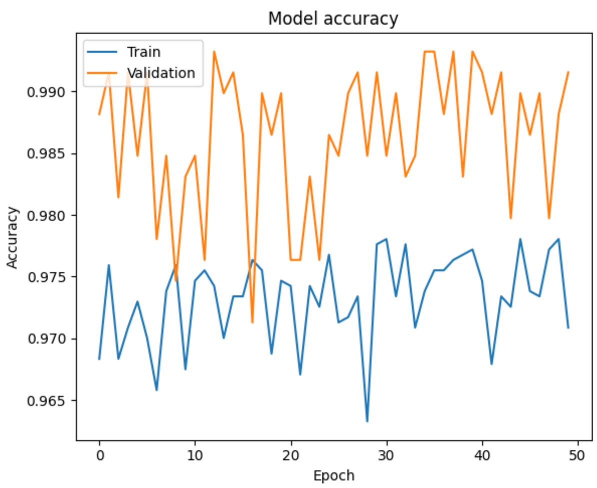

Fig. 7: Training and Validation Accuracy

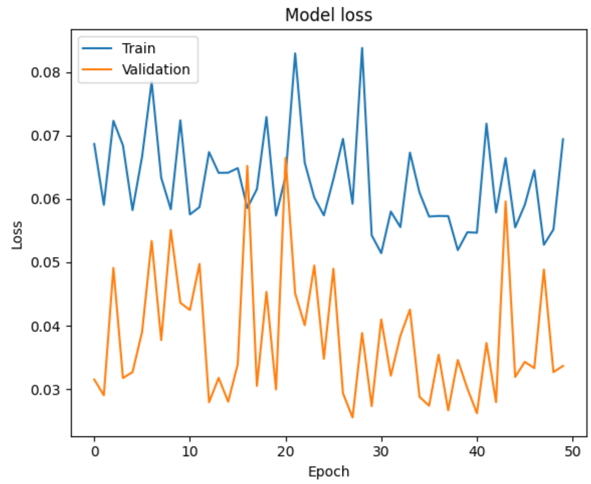

Fig. 8: Training and Validation loss curves

The last aspect of the model development consisted of making a given prediction in relation to a given set of input features. The model was able to give the correct class of the given input and thus they are able to generalize the training data to the new unseen data. In order to estimate the overall model performance, accuracy at the validation set was computed. The overall accuracy of the model was 0.99155 which is a high rate of accuracy and reliability in the prediction of the model. The model is highly trained and works on the validation set and it does not overfit. The fact that the model is robust and reliable in predicting the aforementioned situations in the environment, as the high accuracy, precision, recall, and F1-scores of all classes support, is further confirmed by the fact that the model was trained on the specified input features.

C. Comparative Model Analysis

The research paper emphasizes the findings by comparing two predictive modeling methods, the ARIMA and Artificial Neural Networks (ANN) to assess their application in predicting the environmental conditions. ARIMA is a classic statistical time series model, which can be effectively applied to data that has definite linear tendencies and seasonalities. It is also dependent on previous values and thus can only be effective when there is long term and complete historical data. ARIMA however does not work well when there are non-linear trends, abrupt change or even missing or sparse values. In the study, the ARIMA was ineffective because it did not give the correct real-time forecasts since the long-term data were not available.

Conversely, ANN model was found to be better predictive. In contrast to ARIMA, ANN is able to learn complicated relationships within the data such as non-linear and hidden trends that change over time. It is more adaptable to dynamic and real time situations because it can use the entire structure of the dataset to generalize instead of having to rely on a fixed past time sequence. The ANN model performed very well with a training accuracy of 97.09, a validation accuracy of 99.32 and a prediction accuracy of 99.16. It further provided good levels of precision, recall and

ISSN: 2961-5992

F1-scores in all the categories which has affirmed that it is strong and reliable in predicting environmental conditions.

These findings underscore the flexibility of the ANN model in the processing of both linear and non-linear relationships among the data and makes it more applicable on more complex datasets. Conversely, ANN is a more flexible machine learning model that is capable of learning complex and non-linear associations in the data. Having analyzed both models, it was evident that ANN model is more applicable in real-time prediction as it is more accurate, flexible and strong.

D. Representing Output in Mobile User Interfaces

After the completion of mobile application development and the integration of the machine learning model, Users were presented with an interface capable of displaying three distinct prediction results, outlined as follows (Figure 9).

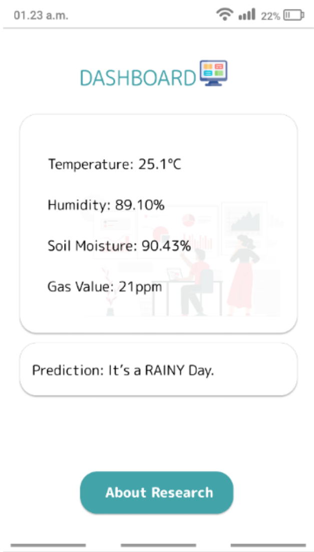

Fig. 9: Output

V. CONCLUSION

The key objective of this study is to construct an environmental monitoring system on the real-time based on four environmental parameters such as temperature, humidity, air pollution, and soil moisture and forecasts the environmental conditions using a machine learning module. Using cheap sensors in the market, an IoT device was developed to keep track of these parameters. The sensors were under control of NodeMCU ESP8266 board, and this guaranteed efficient collection of data. The connection with Google Firebase contributed to smooth storage and dynamic synchronization of the data of the sensor and provided the solid basis to the subsequent analysis. Machine learning was central to the system with a sequential artificial neural network (ANN) being created in Python and its coarse ecosystem of libraries. The ANN was trained to provide environmental parameters, with the information obtained in the IoT device in use. The model was also improved with the addition of a new feature, the Environment Situation which gave the model predictive features, as it allowed the model to give contextual information regarding selective rainy, normal and hot weather. To develop a mobile application, Android Studio was selected and integration with TensorFlow Lite was selected that allowed executing the machine learning model on a mobile device. The android application also gave access to sensor readings and model predictions in real-time to the user, with an interactive interface. The effectiveness of the project implementation was the evidence of the successful work of integrating IoT, machine learning, and mobile technologies to develop a smart environmental monitoring system. The system provided real-time information and insights on future by operating in three distinct environmental situations and empowering the users. This brought easier decision making and proactive responses to the alteration in the environment, and this portrays an exemplar of practicality to administration and comprehension of the environment in real time.

REFERENCES

[1] M. S. Wong, T. Wang, H. C. Ho, C. Y. T. Kwok, K. Lu, and S. Abbas, "Towards a Smart City: Development and application of an improved integrated environmental monitoring system," Sustainability (Switzerland), vol. 10, no. 3, Feb. 2018, doi: 10.3390/su10030623.

[2] S. Zhang et al., "Investigation on environment monitoring system for a combination of hydroponics and aquaculture in greenhouse," Information Processing in Agriculture, vol. 9, no. 1, pp. 123–134, Mar. 2022, doi: 10.1016/j.inpa.2021.06.006.

[3] S. L. Ullo and G. R. Sinha, "Advances in smart environment monitoring systems using iot and sensors," Jun. 01, 2020, MDPI AG. doi: 10.3390/s20113113.

[4] H. B. Fiehn, L. Schiebel, A. F. Avila, B. Miller, and A. Mickelson, "Smart agriculture system based on deep learning," in ACM International Conference Proceeding Series, Association for Computing Machinery, Oct. 2018, pp. 158–165. doi: 10.1145/3289100.3289126.

[5] V. Chang and C. Martin, "An industrial IoT sensor system for high-temperature measurement," Computers and Electrical Engineering, vol. 95, Oct. 2021, doi: 10.1016/j.compeleceng.2021.107439.

[6] N. Novas, J. A. Gázquez, J. MacLennan, R. M. García, M. Fernández-Ros, and F. Manzano-Agugliaro, "A real-time underground environment monitoring system for sustainable tourism of caves," J Clean Prod, vol. 142, pp. 2707–2721, Jan. 2017, doi: 10.1016/j.jclepro.2016.11.005
.

[7] Y. Liu and F. Xiao, "Intelligent Monitoring System of Residential Environment Based on Cloud Computing and Internet of Things," IEEE Access, vol. 9, pp. 58378–58389, 2021, doi: 10.1109/ACCESS.2021.3070344.

[8] S. Qazi, B. A. Khawaja, and Q. U. Farooq, "IoT-Equipped and AI-Enabled Next Generation Smart Agriculture: A Critical Review, Current Challenges and Future Trends," 2022, Institute of Electrical and Electronics Engineers Inc. doi: 10.1109/ACCESS.2022.3152544.

[9] G. Jesus, A. Casimiro, and A. Oliveira, "Using Machine Learning for Dependable Outlier Detection in Environmental Monitoring Systems," ACM Transactions on Cyber-Physical Systems, vol. 5, no. 3, Jul. 2021, doi: 10.1145/3445812.

[10] M. A. Akkaş and R. Sokullu, "An IoT-based greenhouse monitoring system with Micaz motes," in Procedia Computer Science, Elsevier B.V., 2017, pp. 603–608. doi: 10.1016/j.procs.2017.08.300.

[11] S. Dhingra et al., "Internet of Things Mobile-Air Pollution Monitoring System (IoT-Mobair)," IEEE Internet Things J, vol. XX, 2018, [Online].

ISSN: 2961-5992

Available: http://www.ieee.org/publications_standards/publications/rights/index.html

[12] S. Ameer et al., "Comparative Analysis of Machine Learning Techniques for Predicting Air Quality in Smart Cities," IEEE Access, vol. 7, pp. 128325–128338, 2019, doi: 10.1109/ACCESS.2019.2925082.

[13] M. Carminati, O. Kanoun, S. L. Ullo, and S. Marcuccio, "Prospects of Distributed Wireless Sensor Networks for Urban Environmental Monitoring," IEEE Aerospace and Electronic Systems Magazine, vol. 34, no. 6, pp. 44–52, Jun. 2019, doi: 10.1109/MAES.2019.2916294.

[14] F. Deng, P. Zuo, K. Wen, and X. Wu, "Novel soil environment monitoring system based on RFID sensor and LoRa," Comput Electron Agric, vol. 169, Feb. 2020, doi: 10.1016/j.compag.2019.105169.

[15] Institute of Electrical and Electronics Engineers and Manav Rachna International Institute of Research and Studies, Proceedings of the International Conference on Machine Learning, Big Data, Cloud and Parallel Computing : trends, prespectives and prospects : COMITCON-2019 : 14th-16th February, 2019.

[16] W. Li and S. Kara, "Methodology for Monitoring Manufacturing Environment by Using Wireless Sensor Networks (WSN) and the Internet of Things (IoT)," in Procedia CIRP, Elsevier B.V., 2017, pp. 323–328. doi: 10.1016/j.procir.2016.11.182.

[17] H. B. Fiehn, L. Schiebel, A. F. Avila, B. Miller, and A. Mickelson, "Smart agriculture system based on deep learning," in ACM International Conference Proceeding Series, Association for Computing Machinery, Oct. 2018, pp. 158–165. doi: 10.1145/3289100.3289126.

[18] G. Xu, Y. Shi, X. Sun, and W. Shen, "Internet of things in marine environment monitoring: A review," Apr. 01, 2019, MDPI AG. doi: 10.3390/s19071711.

[19] Han'guk Chŏngbo Kwahakhoe, Institute of Electrical and Electronics Engineers, IEEE Computer Society, Denshi Jōhō Tsūshin Gakkai (Japan). Tsūshin Sosaieti, and Han'guk T'ongsin Hakhoe, The 34th International Conference on Information Networking (ICOIN 2020) : January 7 (Tue.)-10 (Fri.), 2020, AC Hotel Barcelona Forum, Barcelona, Spain.

[20] IEEE Power & Energy Society. and Institute of Electrical and Electronics Engineers, IEEE PES GTD 2019 : Grand International Conference & Exposition Asia : conference: March 19-23 ; exposition: March 21-23, Bangkok International Trade & Exhibition Centre (BITEC), Bangkok, Thailand.

[21] O. Bamodu, L. Xia, and L. Tang, "An indoor environment monitoring system using low-cost sensor network," in Energy Procedia, Elsevier Ltd, 2017, pp. 660–666. doi: 10.1016/j.egypro.2017.11.089.

[22] S. R. Shinde, A. H. Karode, and S. R. Suralkar, "Review on-IOT Based Environment Monitoring System," International Journal of Electronics and Communication Engineering and Technology, vol. 8, no. 2, pp. 103–108, [Online]. Available: http://www.iaeme.com/IJECET/index.asp103http://www.iaeme.com/IJECET/issues.asp?JType=IJECET&VType=8&IType=2http://www.iaeme.com/IJECET/issues.asp?JType=IJECET&VType=8&IType=2http://www.iaeme.com/IJECET/index.asp104

[23] G. Mois, S. Folea, and T. Sanislav, "Analysis of Three IoT-Based Wireless Sensors for Environmental Monitoring," IEEE Trans Instrum Meas, vol. 66, no. 8, pp. 2056–2064, Aug. 2017, doi: 10.1109/TIM.2017.2677619.

[24] S. R. Prathibha, A. Hongal, and M. P. Jyothi, "IOT Based Monitoring System in Smart Agriculture," in Proceedings - 2017 International Conference on Recent Advances in Electronics and Communication Technology, ICRAECT 2017, Institute of Electrical and Electronics Engineers Inc., Oct. 2017, pp. 81–84. doi: 10.1109/ICRAECT.2017.52.

[25] W. T. Sung and C. C. Hsu, "Intelligent environment monitoring system based on innovative integration technology via programmable system on chip platform and zigbee network," IET Communications, vol. 7, no. 16, pp. 1789–1801, 2013, doi: 10.1049/iet-com.2013.0266.

[26] I. National Institute of Technology (Punjab, I. D. of C. S. & E. National Institute of Technology (Punjab, Institute of Electrical and Electronics Engineers. Delhi Section, and Institute of Electrical and Electronics Engineers, ICSCCC 2018 : International Conference on Secure Cyber Computing and Communication : December 15-17, 2018.

[27] Mrs. P. P. Kulkarni, "IOT based Smart Agricultural System," Int J Res Appl Sci Eng Technol, vol. 7, no. 4, pp. 2037–2041, Apr. 2019, doi: 10.22214/ijraset.2019.4370.

[28] A. Chehri and R. Saadane, "Zigbee-based remote environmental monitoring for smart industrial mining," in ACM International Conference Proceeding Series, Association for Computing Machinery, Oct. 2019. doi: 10.1145/3368756.3369099.

[29] R. S. AbdulWahhab, "Air quality system using IoT for indoor environmental monitoring," in ACM International Conference Proceeding Series, Association for Computing Machinery, 2019, pp. 184–188. doi: 10.1145/3323933.3324088.

[30] D. Han and M. Rodriguez, "Big Data Analytics, Data Science, ML&AI for Connected, Data-driven Precision Agriculture and Smart Farming Systems: Challenges and Future Directions," in ACM International Conference Proceeding Series, Association for Computing Machinery, May 2023, pp. 378–384. doi: 10.1145/3576914.3588337.

[31] R. M. N. Rathnayaka and A. M. A. Sujah, "IoT-driven environmental intelligence for sustainable tomorrow through advanced machine learning: A systematic literature review," Journal of Information and Communication Technology (JICT), vol. 02, Special Issue, pp. 39–46, Jan. 2025.

ISSN: 2961-5992