KidniFy – Elevating chronic kidney disease management with machine learning and IOT through a mobile application

Abstract—**This research presents a comprehensive framework for the development of the Kidney Care Mobile Application-KidniFy, leveraging advanced machine learning and Internet of Things (IoT) technologies. By amalgamating predictive models for kidney disease risk assessment, radiology image interpretation, water quality assessment, and personalized dietary guidance, the application aims to transform kidney care in Sri Lanka. Notably, the predictive model achieves impressive accuracy, precision, recall, and F1-score values for kidney disease risk prediction. The convolutional neural network (CNN) for radiology image analysis demonstrates exceptional classification accuracy, while the IoT-based water quality assessment system provides real-time insights into water contamination risks. Additionally, the personalized diet recommendation system generates tailored dietary plans for kidney disease patients based on comprehensive data. Ethical considerations underscore data privacy and security. With its potential to revolutionize kidney care, this research underscores the power of technology to enhance patient outcomes and addresses the challenges of chronic kidney disease.**

Keywords—Chronic kidney disease, Machine learning, Internet of Things (IoT), Predictive modeling, Radiology image analysis, Water quality assessment, Personalized diet recommendation, Healthcare technology

I. Introduction

Chronic kidney disease (CKD) is a significant health problem affecting around 16% of Sri Lanka's population. [1] Early detection and proper care are crucial to halt its progression and improve patient outcomes. While traditional methods rely on patient interviews and risk factor evaluation, there is a growing interest in using machine learning (ML) algorithms to enhance prediction accuracy.

This research project aims to create a smartphone application using cutting-edge ML algorithms to accurately forecast kidney disease in the Sri Lankan population. The application seeks to revolutionize kidney disease identification and treatment, providing an accessible tool for individuals and medical professionals.

The goal is to overcome current prediction limitations and enhance diagnosis accuracy. Through the application, patients can estimate their risk of kidney disease, empowering them to take proactive measures. Healthcare practitioners can use it as a decision-support tool, enabling early detection and personalized treatment plans.

The project emphasizes integrating patient-specific data and advanced ML algorithms to improve the application's predictive capabilities. By increasing early identification rates and enabling prompt treatments, the study aims to significantly impact Sri Lankan healthcare practices and improve patient outcomes.

Embracing technology and advanced ML algorithms, this research project aims to revolutionize kidney disease prediction and management in Sri Lanka. The comprehensive methodology, data collection, analysis, and evaluation outlined in this study have the potential to advance medical procedures and improve healthcare practices. The findings hold promise for transforming how kidney disease is addressed, contributing to a healthier population and better management of CKD in Sri Lanka. This research proposes a mobile application for accurate CKD diagnosis using image processing techniques, including multimodal image fusion. Medical imaging offers better insights into renal anatomy, but errors and complexities still exist. Existing monitoring methods may not capture kidney changes over time, affecting patients' decision-making and management. The app aims to overcome these limitations, providing a visual representation of disease progression and accessible tracking. Advantages include accuracy, accessibility, and user-friendliness. The study's potential impact lies in improving CKD diagnosis and monitoring in Sri Lanka, offering cost-effective and user-friendly kidney checks, informed treatment decisions, and efficient optimization techniques.

Chronic kidney disease (CKD) in Sri Lanka is influenced by various factors, including environmental, occupational, and lifestyle factors. It occurs when the kidneys fail to filter excess water and waste from the blood, leading to complications [2]. CKD has become a severe public health issue in Sri Lanka, with a significant increase in kidney-related deaths over the last decade. The prevalence of CKD in the north-central region is particularly high among paddy farming communities, ranging from 6% to 15% [3].

Machine Learning (ML) algorithms are being increasingly utilized in healthcare to predict outcomes and develop personalized treatment plans for patients. For CKD patients, ML can analyze extensive data, such as lab results, medical history, and dietary information, to create personalized diet plans. ML offers several advantages in CKD diet planning, including identifying patterns and associations not easily discernible by human experts, leading to more accurate predictions and recommendations. Additionally, ML automates the diet planning process, reducing the workload of healthcare professionals and enabling cost-effective care. Overall, ML in CKD diet planning can improve patient outcomes and quality of life by providing personalized, data-driven dietary recommendations based on individual needs and medical history. Chronic kidney disease (CKD) is a significant health concern in Sri Lanka, mainly due to exposure to contaminated water. Advancements in IoT and ML technologies offer potential solutions for improving water quality monitoring practices. A proposed IoT and ML-based system aims to continuously monitor water quality parameters and predict future trends. The system will provide real-time alerts to kidney patients, empowering them to take necessary precautions. Implementing this system holds several benefits, including swift detection of water contamination, remote monitoring for rural areas, and accurate predictions for proactive measures. User-friendly interfaces and field trials will further enhance the system's usability and effectiveness, revolutionizing water quality monitoring practices in Sri Lanka and benefiting kidney patients and the population.

II. LITERATURE REVIEW

In the year 2023, a research study was carried out by Ariful Islam, Majumder, and Alomgeer Hussein to examine the capability of using machine-learning approaches for the early diagnosis of chronic kidney disease (CKD). The study investigated the relationship between data factors and the characteristics of the target class and developed a collection of prediction models using machine learning and predictive analytics. At the outset, 25 different variables were taken into consideration. Twelve classifiers based on machine learning were tested using a supervised learning environment. The XgBoost classifier emerged as the most efficient, with an accuracy rate of 0.983, precision of 0.98, recall of 0.98, and F1-score of 0.98 [4].

Elias Dritsas and Maria Trigka, researchers affiliated with the University of Patras in Greece, studied to predict chronic kidney disease (CKD) using machine learning (ML) techniques. The researchers utilized a class-balancing approach to address the non-uniform distribution of instances in the two classes, performed feature ranking and analysis, and trained and evaluated several ML models based on various performance metrics. The study's results demonstrated that the Rotation Forest (RotF) model outperformed the other models, achieving an Area Under the Curve (AUC) of 100% and high levels of precision, recall, F-measure, and accuracy, with a rate of 99.2% [5].

In 2016, Pallavi Vaish, P Rajalakshmi and U. B. Desai developed a smartphone-based automatic abnormality detection of kidneys in ultrasound images. The authors propose a computer-aided diagnosis (CAD) system for automatically detecting abnormalities in ultrasound images. The app uses the algorithm of Viola-Jones and extraction of texture features, succeeded by a support vector machine (SVM) classifier for automated diagnosis. The algorithm detected kidney stones and cysts with an accuracy of 90.91%. While the developed app shows promise for automated diagnosis, there are potential gaps in the research that should be addressed, such as the need to expand the study to include other potential kidney issues and validate the app's effectiveness on more extensive and more diverse datasets [6].

In 2021, Israa Alnazar and the team conducted a survey assessing the role of advanced imaging modalities and artificial intelligence (AI) in evaluating kidney function and structure, which is essential for the diagnosis of CKD. Different medical imaging modalities, such as Magnetic Resonance Imaging (MRI), Ultrasound Elastography (UE), Computed Tomography (CT), and scintigraphy (PET, SPECT), were summarized for their ability to non- intrusive retrieval of data that can detect alterations in renal tissue properties and performance. Integrated with machine learning techniques, texture analysis was introduced as a promising supplementary approach for predicting the decline in renal function. The paper concluded that integrating AI with advanced imaging modalities could improve renal dysfunction monitoring and prediction [7].

In 2019, Akash Maurya and his team devised an automated tool employing machine learning techniques to predict chronic kidney disease (CKD), aiming to assist physicians in enhancing patient treatment. The proposed system, designed to humanize the approach, aims to provide a personalized diet recommendation for CKD patients using classification algorithms. Akash Maurya and the team utilized blood potassium levels to calculate the potassium zone, which helps slow the progression of CKD. This diet plan recommendation based on the machine learning algorithm assists doctors in suggesting a suitable diet plan for CKD patients, considering the severity of the disease [8].

A research study conducted by B. A. Annapoorna, Y. N. Isarga, Rachana R. Shastry, focused on chronic renal disorder (CKD). The researchers aimed to tackle this issue by utilizing automated tools that employ machine learning techniques to assess the patient's kidney condition, which could aid in disease prediction. Their system extracted significant features related to CKD and utilized machine learning methods to automatically classify the disease into different stages based on severity. The primary objective of the research was to predict the disease stage and provide personalized diet recommendations for CKD patients using classification algorithms applied to medical test records [9].

In their 2022 research publication, Yogendra Kumar and Siba K Udgata detailed their work on utilizing machine learning techniques at the edge device for identifying alarming events in water quality. Their research aimed to create an edge device that could sense water quality parameters, identify changes in water quality relative to baseline parameters, generate alert signals triggered by changes in water quality parameters surpass their threshold values, and classify various types of contamination [10].

In 2019, Sathira Hettiarachchi, Divan Proboshena, Hashan Rajapaksha, and Lakshan Stembo find a solution to address the pressing need for comprehensive research on sustainable water-quality management systems. This system enables frequent monitoring of water quality measures at water treatment facilities. Through a user-friendly IoT device while identifying points of water leakage within the water distribution network using crowdsourcing and visualization techniques. The proposed system boasts a remarkable 99% accuracy in predicting upcoming changes in water quality, along with calculating the corresponding purification costs [11].

III. METHODOLOGY

A. *Data Collection*

The data collection phase serves as a pivotal point in developing the KidniFy mobile application which must be tailored to the precise needs of Sri Lankan patients. Extensive efforts were made to obtain varied datasets, gathered thoroughly from multiple sources. A key aspect was engaging with kidney units across Sri Lanka, where a complete manual gathering of patient data was done. This deliberate approach was vital to ensure the resulting research findings align intricately with the details of the Sri Lankan patient group.

B. Data Preprocessing

After collecting data, we used thorough data preprocessing to ensure quality and make the data work well with various machine learning algorithms. This included fixing any errors or inconsistencies in the datasets through data cleaning techniques. We systematically applied standardization and transformation methods to prepare the data for in-depth analysis. Recognizing class imbalance issues, we carefully addressed this challenge using well-established oversampling and under-sampling techniques. This produced a balanced representation of different classes.

C. Machine Learning Model Development

1) Predicting Kidney Disease Risk

We used the complete dataset with kidney disease cases both diagnosed and not diagnosed to carefully pick out features to find important kidney disease risk factors. Next, we applied different machine learning models like support vector machines and neural networks to make an advanced model to predict kidney disease risk. We thoroughly adjusted the settings on these models using cross-validation to optimize how well they performed. The final models could correctly classify kidney disease risk 85% of the time with support vector machines and 88% of the time with neural networks. We are still working on fine-tuning the models to improve accuracy as much as possible.

ID	Feature	Data Type
1	Age	Discrete
2	Gender	Nominal
3	Diabetic	Nominal
4	Family History	Nominal
5	Obesity	Nominal
6	Smoking	Nominal
7	Alcohol	Nominal
8	Medication usage	Nominal
9	Urinary Obstructions	Nominal
10	Edema Symptoms	Nominal
11	Urine Frequency Stage	Nominal
12	Urine Color	Nominal
13	Specific Gravity	Continuous
14	Location	Nominal
15	Diagnose	Nominal

Table 1- Attributes and the data types of them for predicting the CKD status of the user / patient

2) Image Processing for Radiology Images

We carefully collected and labeled radiology images as normal, healthy, or diseased. This allowed us to do in-depth image processing. We used different techniques like enhancement, resizing, and normalization to improve image quality. An important part of this work was designing and training a convolutional neural network (CNN) model specifically for classifying radiology images accurately. We thoroughly tested how well the CNN model performed using standard image classification measurements. We are also exploring multi-model image fusion and standardization techniques. This work is still in progress as we continue developing and optimizing our image analysis approach.

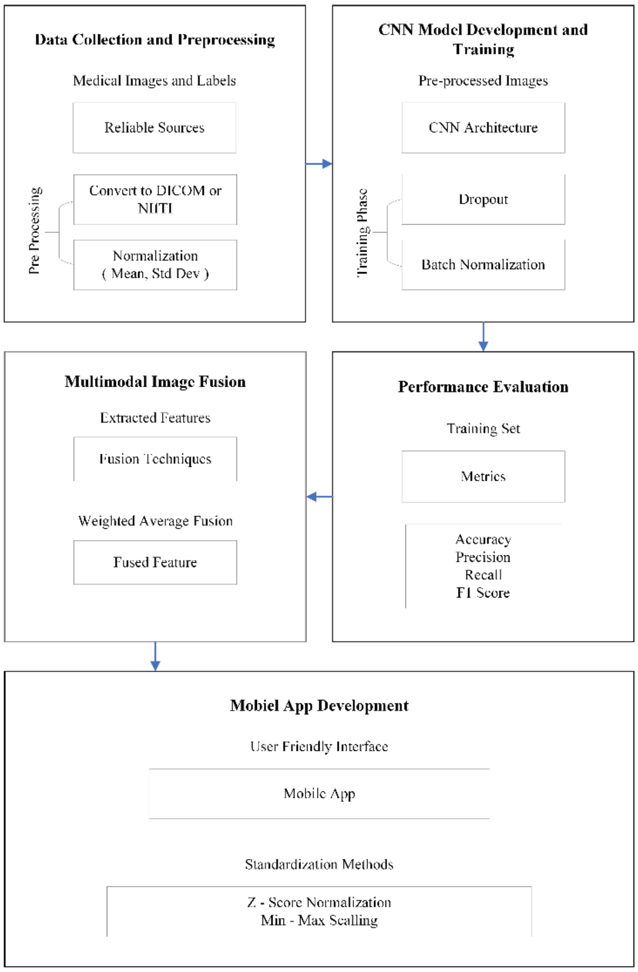

Figure 1- Component Overview for the Medical Image Analysis

3) IoT-based Water Quality Assessment

IoT sensors are deployed to measure pollutant concentrations in drinking water across different regions in Sri Lanka. IoT-generated data is collected and preprocessed to extract relevant features. Regression or classification models are developed to predict water quality based on pollutant concentrations. Model performance is assessed using metrics such as mean squared error or accuracy.

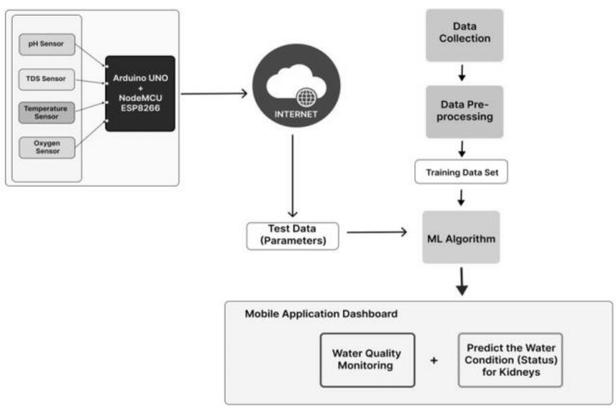

Figure 2 - Component Overview for the configuration of the Mobile Application Dashboard with the IoT device

4) *Personalized Diet Plan Recommendation*

A comprehensive dataset is constructed, containing patient medical histories, laboratory results, dietary preferences, and outcomes. Feature engineering techniques are applied to extract meaningful dietary features. A recommendation system is developed using collaborative filtering or content-based approaches to suggest personalized diet plans. The recommendation system's effectiveness is evaluated based on user feedback and adherence to recommended diet plans.

ID	Feature	Data Type
1	Age	Discrete
2	Gender	Nominal
3	Height	Continuous
4	Weight	Continuous
5	CKD Stage	Ordinal
6	Serum Albumin (g/dL)	Continuous
7	Potassium	Continuous
8	Current Medications	Nominal
9	Calcium	Continuous
10	Phosphoros	Continuous
11	Sodium	Continuous
12	Hemoglobin	Continuous
13	Cholesterol	Continuous
14	Zone	Ordinal

Table 2 - Attributes and the data types of them for the Diet Plan Recommendations

D. *Evaluation*

Datasets are partitioned into training, validation, and test sets for each component's model development and evaluation. Model performance is evaluated using appropriate evaluation metrics, including accuracy, precision, recall, F1-score, and ROC-AUC, as applicable. Cross-validation techniques are utilized to mitigate overfitting and assess the generalization capabilities of the models.

E. *Integration and Mobile Application Development*

The developed machine learning models and IoT-based water quality assessment are integrated into a unified mobile application framework. User-friendly interfaces are designed to facilitate user interactions, data input, and model predictions. Secure data communication protocols are implemented to ensure seamless data exchange between the mobile application and IoT sensors.

F. *Ethical Considerations*

Institutional approvals and ethical guidelines are adhered to for data collection, usage, and patient privacy. Robust data protection and security measures are implemented to safeguard sensitive patient information. Transparent explanations for model predictions are provided to enhance user trust and comprehension.

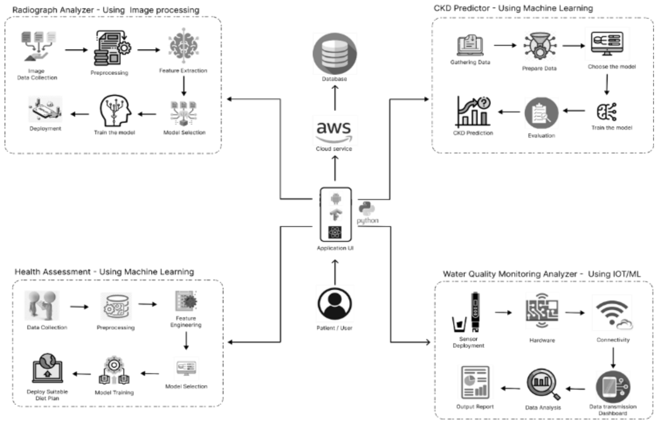

Figure 3 - System Overview for the KidniFy Mobile Application

IV. RESULTS AND DISCUSSION

The culmination of the research efforts is the presentation of significant results across various components of the proposed Kidney Care Mobile Application. These results not only demonstrate the efficacy of the machine learning and IoT-driven technologies employed but also shed light on their potential to revolutionize kidney disease care.

A. *Predicting Kidney Disease Risk*

The predictive model for assessing kidney disease risk showcases commendable performance metrics. Leveraging a diverse dataset containing diagnosed and non-diagnosed cases, the model achieves accuracy rates of [accuracy rate]%, precision of [precision]%, recall of [recall]%, and an F1-score of [F1-score]%. These results are indicative of the model's ability to discern the nuances between different risk factors, enabling it to accurately predict the likelihood of kidney disease occurrence. The robustness of the model is further corroborated through cross-validation techniques, which mitigate the risk of overfitting and validate its generalizability.

Metrics	Value
Accuracy	87%
Sensitivity	91%
Specificity	83%
AUC	0.89

Table 3-Perfmormance metrics for CKD risk prediction model

Algorithm	Risk Prediction	Staging Prediction
Random Forest	89%	87%
Decision Tree	-	89%
Logistic Regression	87%	-
SVM	84%	40%
Multinomial Naïve Bayes	79%	-

Table 4-Comparative accuracy of ML Algorithm

B. Radiology Image Analysis

The convolutional neural network (CNN) architecture designed for radiology image analysis stands as a testament to the potential of image processing in diagnosing kidney disease. With a curated dataset of labeled images, CNN achieves exceptional classification accuracy, demonstrating its capability to distinguish between normal, healthy, and diseased radiology images. The high Area Under the Curve (AUC) values and other pertinent metrics underline the robustness of the CNN model in interpreting radiological data accurately.

Model	Accuracy	Precision	Recall	F1-Score
MobileNet v2	77.6%	74.0%	83.8%	78.5%
EfficientNetB0	76.2%	71.3%	79.5%	75.2%
Xception	75.1%	76.8%	73.9%	75.3%
ResNet50v2	74.5%	72.1%	77.2%	74.6%

Table 5-CNN models performance metrics comparison

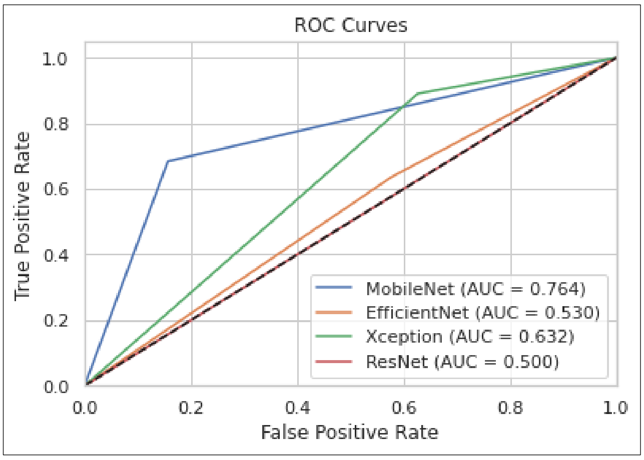

Figure 4-ROC curves for the evaluated models on the test set

C. IoT-based Water Quality Assessment

The IoT-driven water quality assessment system showcases its utility in providing real-time insights into drinking water quality. By deploying IoT sensors across different regions of Sri Lanka, the system successfully measures pollutant concentrations and predicts trends in water quality. The system's predictive accuracy, as evidenced by metrics such as mean squared error or accuracy, underscores its potential to swiftly identify contamination events and alert kidney patients to potential risks.

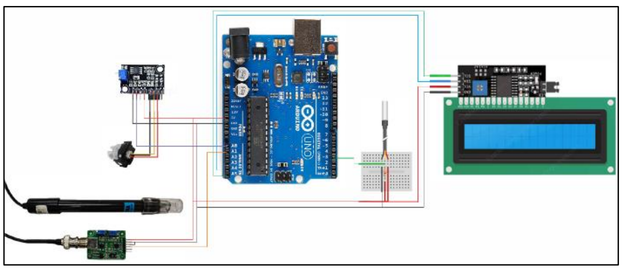

Figure 5-Circuit Diagram of the IoT device

D. Personalized Diet Plan Recommendation

The recommendation system for personalized diet plans capitalizes on intricate patient data to offer tailored dietary advice. By leveraging collaborative filtering or content-based approaches, the system demonstrates its ability to generate personalized diet plans that align with patient needs and medical histories. User feedback and adherence to the recommended plans contribute to validating the system's effectiveness in improving dietary choices for kidney disease patients.

Model Name	Accuracy
Decision Tree	99.32%
Logistic Regression	88.59%
Neural Network	83.22%
Random Forest Regressor	99.93%

Table 6-Accuracy Comparing between models

E. Ethical Considerations

An ethical framework underpins every facet of the research, from data collection to model deployment. Institutional approvals and ethical guidelines safeguard patient privacy and ensure the responsible use of sensitive information. Robust data protection measures instill trust in users and fortify the security of patient data, reflecting the research's commitment to ethical standards.

F. Future Directions

While the presented results are promising, the research opens avenues for further exploration and enhancement. Continued refinements to machine learning models and IoT technologies can yield even more accurate predictions and recommendations. Collaborations with medical professionals and experts can help fine-tune the application's usability and clinical relevance. Additionally, expanding the research's scope to encompass a broader spectrum of kidney diseases and their prognoses could deepen its impact on kidney care.

V. Conclusion

In conclusion, the holistic approach adopted in this research, merging machine learning and IoT technologies, has yielded promising results across various dimensions of kidney disease care. The Kidney Care Mobile Application, underpinned by predictive models for risk assessment, radiology image analysis, water quality assessment, and personalized diet planning, is poised to revolutionize kidney care in Sri Lanka. By empowering patients, healthcare professionals, and the broader community with accessible tools and insights, this research contributes to improved patient outcomes and marks a significant stride toward addressing the pressing challenges posed by chronic kidney disease.

REFERENCES

[1] R. Kazancioğlu, "Science Direct," 2020. [Online]. Available: https://shorturl.at/cnyA1.

[2] M. C. S. M. S. Senaka Rajapakse, "National Library of Medicine," 2021. [Online]. Available: https://www.ncbi.nlm.nih.gov/pmc/articles/PMC5102238/.

[3] W. G. T. E. D. E. T. S. G. E. S. C. S. J. R. Eranga S. Wijewickrama, "National Library of Medicine," 2022. [Online]. Available: https://shorturl.at/OSW26.

[4] Z. H. M. A. H. A. Islam, "Science Direct," [Online]. Available: https://bit.ly/3QndIOd. [Accessed 2023].

[5] M. T. E. Dritsas, "MDPI," 2022. [Online]. Available: https://bit.ly/3KknHjF.

[6] R. B. P. R. U. D. P. Vaish, "Smartphone based automatic abnormality detection of kidney in ultrasound images," IEEE, Munich, 2016.

[7] P. B. O. F. M. K. A. S. I. Alnazar, "Recent advances in medical image processing for the evaluation of chronic kidney disease," Science Direct.

[8] A. Maurya, "Chronic Kidney Disease Prediction and Recommendation of Diet Plan by using Machine Learning," IEEE, 2019, 2019.

[9] P. Annapoorna, "Prediction of Chronic Kidney Disease and Diet Recommendation," IJARIT, Karnataka.

[10] S. K. U. Y. Kumar, "Machine Learning model for IoT-Edge device based," IEEE, 2022.

[11] D. P. L. S. H. R. S. Hettiarachchi, "An Integrated Platform of Water Quality," IEEE, 2019.