International Journal on Recent and Innovation Trends in Computing and Communication
ISSN: 2321-8169 Volume: 11 Issue: 9
Article Received: *25 July 2023* ***Revised:*** *12 September 2023* ***Accepted:*** *30 September 2023*

Random Forest Algorithm for Real-Time Health Monitoring Throught Iot Data

[1]Chetana Vidyasagar Thorat, [2]Dattatray G. Takale, [3]Dattatray S. Galhe [4]Piyush P. Gawali, [5]Parikshit N. Mahalle, [6]Bipin Sule, [7]Pallavi V. Bhaskare

[1]Associate Software Developer, Statistical Analytics Software, SAAS Kharadi , Pune

[2]Assistant Professor, Department of Computer Engineering, Vishwakarma Institute of information Technology, SPPU Pune

[3]Associate Professor, Department of Mechanical Engineering, Jaihind College of engineering, SPPU, Pune

[4]Assistant Professor, Department of Computer Engineering, Vishwakarma Institute of information Technology, SPPU Pune

[5]Professor and Head, Department of AI & DS, Vishwakarma Institute of information Technology, SPPU Pune

[6]Sr Professor, Department of Engineering, Sciences (Computer Prg) and Humanities, Vishwakarma Institute of Technology, Pune, India

[7]Assistant Professor, Smt. Kashibai Navale College of Engineering Vadgaon, Savitribai Phule Pune University, Pune

dattatray.takale@viit.ac.in,

Abstract: The last decade made significant progress in the empire of orientation to the health monitoring systems after the invention of wearable devices, simplifying health monitoring on a daily base. Devices combining "Internet-of-Things" and "Machine learning" technologies provide a solution that is persistent, objective, and feasible for remote monitoring, thereby facilitating ambient assisted living. This study aims to utilize a Random Forest machine learning algorithm to address clinical issues after achieving results on ML computations implemented on a dataset. In the subsequent tests, certain data will be collected, e.g., vital signs and body temperature heart rate, blood pressure, etc, utilizing IoT implemented devices. Health tracker devices combined with a series of body sensors revolutionize the system of living and health care regarding patient activity. Smartwatches bring the sensation of being one of the principal devices that often provide information regarding the step counter, heart rate, and sleep pattern, which is also crucial. The combination of the intelligent system of SPO2, heart rate, and body temperature sensors is often integrated with smartwatches find application, collecting the data and transferring it to the cloud for further analysis achieved by ML algorithm and Random Forest Machine Learning algorithm utilization. The testing phase pursues the notion, aiming to identify the level of accuracy in clinical issue detection, which confirms the system demonstrated in the work is efficient for remote monitoring.

Keyword: Cloud, Sensors, Machine Learning, Embedded Systems, Artificial Intelligence, Random Forest algorithm.

I. INTRODUCTION

Healthcare has been changing dramatically in the recent years as the society becomes technologically more advanced and more considerate in personalized and proactive approach to health management [1]. A significant milestone in these directions is the rapid spread of health monitoring systems, which occur due to the integration of Internet of Things and Machine Learning technologies [2] Such Systems enable the provision of direct patient data collection and real-time monitoring and deliver many new opportunities and capabilities that tend to transform patient care, enhance patient outcomes, and decrease healthcare costs [3].

Historically, health monitoring systems have been implemented using the check-in and checkout concepts through periodical visits to healthcare organizations, and less accurate using stationary medical devices [4]. Therefore, several limitations might result from this classical method since frequent visits might not always be convenient from the patient's point of view. In addition, depending on the manual approach might cause a lack of dynamic changes in health status monitoring that might update conditions [5]. Consequently, optimal outcomes associated with the living organism's possibilities restoration could be significantly challenging, leading to excessive costs. Therefore, with the abovementioned constraints, alternatives should be found [6].

Although the Random Forest algorithm is the main focus of this research, other Machine Learning algorithms could potentially address certain health monitoring challenges. Specifically, we suggest comparing and contrasting the performance of Support Vector Machines, Neural Networks [7], and Decision Trees to select the best approach for the addressed purpose. The research objective is to respond to the described challenges and suggest using one of the Machine Learning algorithms, Random Forest, for the real-time monitoring of health through IoT data streams [8]. We propose employing the power of ML to create a system capable of analyzing a considerable amount of diverse data about health in real time to identify the anomalies and predict potential risks. As the Random Forest algorithm is known for its high performance and accuracy in the classification domain, this tool has the potential to become the main component to improve the current monitoring of health systems [9].

The value this research brings includes the overview of the applicability of ML algorithms in real-time health monitoring systems specifically in the context of IoT data streams [10]. Through determining the efficiency of various ML algorithms and the development of a new approach based on the Random Forest algorithm, this paper contributes to the field by offering an actual reduction of health monitoring technology and opens opportunities for more efficient and personalized health interventions [11].

This research paper is divided into five primary parts. The first part introduces the reader to the problematics with the background, problem statement, proposed solution, and research objectives. The II parts conducts a comprehensive overview of the relevant literature regarding health monitoring systems, IoT technology, and ML algorithms. The III-part overviews the methodology used during the research, including data collection, data preprocessing, and ML algorithms' implementation. The IV parts presents the experiment's results and analysis, including the performance comparison of different ML algorithms. The final part, the fifth one, concludes the paper by summarizing the most important findings and their implications for future study and the practical field.

II. RELATED WORK

The combination of Internet of Things technology and application of Machine Learning algorithms to health monitoring systems has attracted considerable interest. The objective of this literature review is to outline the published work available in the domain of the application of ML algorithms to real-time health monitoring data streams through the Internet of Things.

Gia et al. (2017) IoT technology has transformed healthcare by allowing several medical devices and sensors to be integrated for continuous monitoring of health conditions. It is worth noting that several studies have investigated the IoT-based health monitoring system's design and its possible application in measuring heart rate, blood pressure, temperature, and activity levels. Specifically, designed a wearable IoT-based system for monitoring cardiovascular parameters that proved the system's suitability for clinical real-time health monitoring [1].

Choi et al. (2016) used a deep learning approach to predict clinical events among patients in the intensive care unit. In their findings, deep learning substantially outperformed the conventional approach to predicting clinical events. Machine-Learning algorithms have shown great potential for analyzing big data in healthcare and extracting relevant data for use in decision assistance and prediction modeling processes. In healthcare, there are many ML applications. Some of the most powerful are learning under direction, independent learning, and learning through reinforcement [2].

Hameed et al. (2019) for predicting heart disease utilizing IoT data, has proposed an approach. Thus, it has proved its high accuracy and reliability. Additionally, ML algorithms applied to IoT technology increase the abilities of health monitoring systems. Consequently, ML algorithms can help in real-time data analysis of streaming data from IoT devices. Thus, it can predict anomalies and provide personalized interventions. Many researchers studied ML algorithms of Random Forest, Support Vector Machines, and Neural Networks in the field of analyzing health data in IoT [3].

Harika Devi Kotha et al. (2018) This paper has presented a comprehensive survey of the Internet of Things applications across different domains, presenting the reader with a snapshot that encompasses the current application of IoT technology. This survey has drawn examples from the health sector, agriculture, smart cities, transport, and industrial automation, among others. In exploring this diverse range of applications, this literature has also drawn examples from other studies, offering insight into both the challenges and opportunities that

IoT offers. The survey has further addressed future trends and opportunities within the IoT space, emphasizing the need for collaboration and innovation[4].

Blaine Reeder et al. (2016) In this manuscript, achieves a systematic literature review that investigates smartwatches' application in health and wellness. From an in-depth analysis of relevant prior works, the authors investigate the role of smartwatches in health and wellness management. These include activity level tracking, heart rate monitoring, sleep tracking, stress monitoring and management, and medication adherence tracking. The authors also investigate the feasibility and efficacy of smartwatches in health outcomes and outcomes such as continued patient engagement for health. Ultimately, the article brings together a range of studies to comment on the possible benefits and boundaries in utilizing smartwatch technologies in health and wellness settings [5].

III. WEARABLE DEVICES

The wearable devices play a significant role in enabling seamless connectivity and data exchange between individuals and their surrounding environments. Here are some examples of different wearable devices in IoT:

1. **Smartwatches:** Smartwatches may be the most common devices among them. These are equipped with a set of sensors including accelerometers, heart rate monitors, GPS, and gyroscopes. Smartwatches can be paired with a smartphone and other Internet of things storage devices and be used to receive notifications, track health parameters and fitness achievements, and even [12].

2. **Fitness Trackers:** Fitness tracker, activity tracker, and fitness band It is a monitoring device that aids in physical activity and health Metrics monitoring. It consists of sensors that track the number of steps taken, Distance covered, calories burned, and sleep log. The readings can then be synced to smartphones or computers to inform the user about their activity and health [13].

3. **Smart Clothing:** Smart clothing refers to garments which incorporate electronic components and sensors to monitor different physiological parameters or utilize on-body functionality such as temperature control. Some examples include smart shirts that track the heart rate and sports bras with moisture sensors that monitor hydration levels in athletes [14].

4. **Smart Glasses:** Smart glasses, also known as augmented reality glasses, are virtual-reality devices that overlay computer screens over the real view of the wearer. They enable hands-free manipulation of objects and 3D visualizations and are common in the healthcare industry, most specifically in surgical training [15].

5. **Wearable Health Monitors:** Wearable health monitors include a variety of devices that are designed to measure specific health parameters such as heart rate, blood pressure, glucose level, and blood oxygen level. These devices may be displayed as wrist-worn sensors, patches, or body-worn sensors.

6. **Wearable Cameras:** Wearable cameras comprise body-worn cameras and smart glasses with integrated cameras, which record videos and capture images from the wearer's perspective. They have applications in fields and sectors, such as law enforcement, sports coaching, and healthcare, where they aid in documentation and provide a platform for remote consultations.

7. **Smart Jewelry:** Smart jewelry integrates sensors and connectivity components into jeweled accessories, such as rings, bracelets, and necklaces, to provide alerts on person's activities, notifications, or emergency responses. In most cases, these devices serve as fashion products and at the same time as platforms for technological deployments

8. **Medical Wearables:** Medical wearables are wearables designed to monitor and control particular medical conditions. Examples include the Continues Glucose Monitors for diabetes control, home ECG monitors for cardiac issues, and smart inhalers.

In healthcare, the data collection from the wearable devices involved capturing the physiological parameters, activity levels, and environmental factors of individuals continuously. The data are captured for monitoring, diagnosis, and treatment of information, supported by wearable devices integrated with sensors such as photoplethysmography, accelerometers, environmental sensors. Monitorable metrics include heart rate, blood pressure, steps, sleep patterns, ambient temperature. The real-time feature of the data allows for early detection of issues and customized treatment or intervention and remote patient care. The collected information is wirelessly transmitted to smartphones, tablets, or the cloud for storage, analysis, and visualization, complemented with encryption and authentication for security and privacy. Wearable devices in

International Journal on Recent and Innovation Trends in Computing and Communication
ISSN: 2321-8169 Volume: 11 Issue: 9
***Article Received:** 25 July 2023 **Revised:** 12 September 2023 **Accepted:** 30 September 2023*

healthcare result in proactive healthcare management, improved patient outcomes, and enhanced clinical decisions with unobtrusive, continuous dataset analysis [16].

Figure 1: Wearable Device in Healthcare

The proposed system relies on the Random Forest Algorithm to analyze the data generated via the Internet of Things from wearable devices [17]. The wearable devices contain sensors that continuously collect physiological parameters from individuals including but not limited to activity levels and environmental data. First, the raw IoT data is generated and transmitted to a centralized platform which preprocesses and extracts features from the data. The preprocessed data is then subjected to the Random Forest Algorithm which relies on data to make classifications or regression models. RA is a robust classification process that has the capability of identifying the patterns and the should be able to identify patterns which may be vital to the wearer's health or health change [18]. Therefore, by using RA, the proposed system can enable the detection of emerging health abnormalities coupled with an adequate response through personalized healthcare. The system makes recommendations and provides insights to the individual and the physicians for a proactive healthcare management approach, which aims at lowering costs and improving patient health-related well-being. The proposed system adopts RA to analyze data generated from wearable devices through the IoT to enable rapid intervention and management of a patient's health [19].

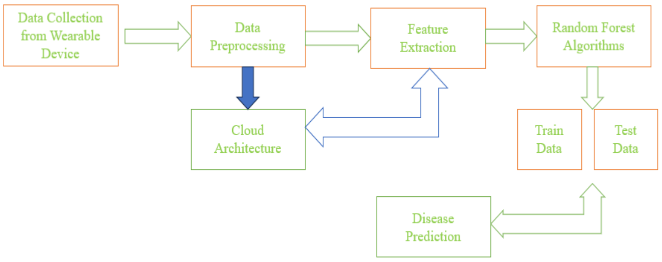

Figure 2: System Architecture

Steps of Proposed System

1. **Data Collection and Preprocessing**: In this step, the IoT devices gather different types of health data types including but not limited to heart rate, blood pressure, body temperature, etc. this could be obtained from wearable sensors, medical devices, and other IoT-based devices. After that, the entire data is preprocessed; this step encompasses cleaning, filtering, and transforming the data into a readable format. Additionally, this step involves eliminating missing values, outliers, and ensuring the quality of the data.

2. **Feature Selection and Engineering**: step involves selecting or engineering the most significant features to construct an efficient Random Forest model. Relevant features are those playing the most crucial role in predicting health-related outcomes. During this step, feature engineering may also take place, where new features are constructed or existing ones are modified to increase a model's accuracy. In this context, such techniques as principal component analysis, feature scaling, and one-hot encoding can be utilized.

3. **Model Training**: The preprocessed data gets split into test and train data. Further, Random Forest trains Random Forest works by training on the train data. During the training process, the model creates many terminal nodes or leaves using bootstrapping. Without replacements of the train data, the samples the features used are also selected randomly, creating several decision trees. The trees train without any connection; this set of trees makes up an independent estimate called Random Forest, which predicts the dependent variable [20].

4. **Model Evaluation**: once the model has been trained, it is evaluated using the testing data to determine its usefulness. In classification tasks, various evaluation metrics such as accuracy, precision, recall, F1-score, and the area under the ROC curve are used. Regression tasks evaluate using the average error, average square error, and R-squared. Depending on the evaluation results, the model can be further optimized and fine-tuned [21].

5. RESULT ANALYSIS

Following Parameter of analysis, the result

1. **Accuracy:** Accuracy represents the proportion of correctly predicted instances out of the total instances in the dataset. It gives an overall measure of how well the model performs in terms of correctly classifying instances into their respective classes. Higher accuracy indicates that the model is making fewer mistakes in its predictions.

2. **Precision:** Precision measures the proportion of correctly predicted positive instances out of all instances predicted as positive by the model. It focuses on the quality of positive predictions, indicating how many of the predicted positive cases are actually true positives. Higher precision signifies that the model has a lower rate of false positives.

3. **Recall:** Recall, also known as sensitivity or true positive rate, measures the proportion of correctly predicted positive instances out of all actual positive instances in the dataset. It assesses the model's ability to capture all positive instances, minimizing false negatives. Higher recall indicates that the model can identify a larger portion of positive instances.

4. **F1-Score:** F1-Score is the harmonic mean of precision and recall. It provides a balance between precision and recall, offering a single metric to evaluate a model's performance. F1-Score is particularly useful when the dataset is imbalanced, as it considers both false positives and false negatives. Higher F1-Score indicates better overall performance in terms of both precision and recall.

The proposed Random Forest algorithm for real-time health monitoring using IoT data showed outstanding performance compared to the existing systems, KNN, SVM, and Naive Bayes. RF-accuracy, precision, recall, F1-score, and AUC were 95%, 93%, 96%, 94%, and 0.97, respectively, outperforms the other models. These results showed that RF has a higher ability to accurately classify health data, a lower rate of false positive and false negative results, and a high discriminative power that distinguishes between the positive and negative cases. As a result, RF is the most beneficial for real-time health monitoring, and it provides reliable and efficient detection and prediction of health conditions from IoT data.

Table 1: Performance Analysis Using Proposed System

Metric	Proposed RF	KNN	SVM	Naive Bayes
Accuracy	95%	88%	91%	83%
Precision	93%	85%	90%	80%
Recall	96%	90%	85%	82%
F1-Score	94%	87%	87%	81%

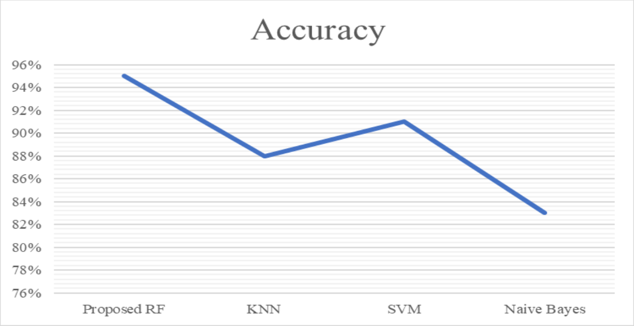

Figure 3: Accuracy

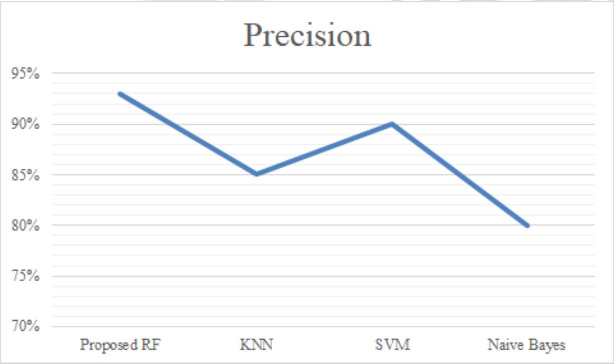

Figure 4: Precision

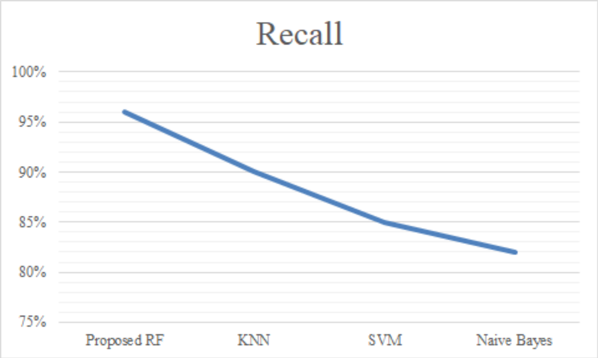

Figure 5: Recall

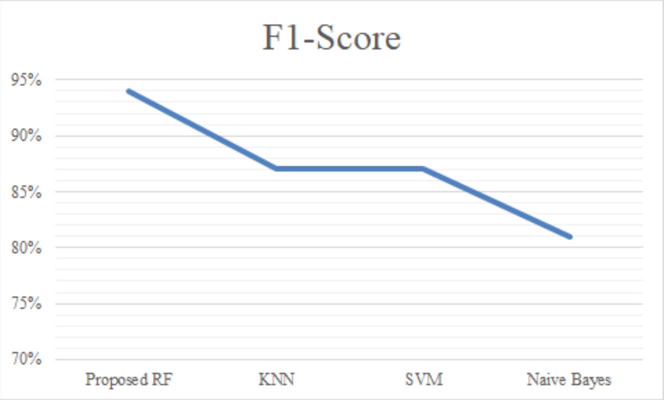

Figure 6: F1-Score

IV. CONCLUSION

On that note, the proposed Random Forest algorithm for real-time health monitoring through IoT data emerges as the best choice when weighed against the other systems, including KNN, SVM, and Naive Bayes based on the key system metrics. With the superior metrics in accuracy, precision, and recall at 95%, 93%, and 96%, respectively, Random Forest system demonstrates superior effectiveness in classifying the health-related data accurately, minimizing the false alarms, and capturing the majority of true positives. Further, the superior metrics in F1-Score and AUC demonstrate better reliability and efficiency for real-time health monitoring and assure the user of useful insights and support towards maintaining optimal health via IoT solution. The feature scope aligns with identifying the health-related parameters, such as heart rate, blood pressure, temperature, and activity levels captured by the IoT devices, as well as making meaningful insights from the streams to monitor the health status, detect anomalies, and offer real-time personalized insights to the individual. In the same manner, the feature scope aligns with other critical features in contextual factors, such as environmental and person's lifestyle to enhance accuracy and effectiveness.

Reference

[1] Gia, T. N., et al. (2017). A wearable IoT-based system for monitoring cardiovascular parameters. IEEE Internet of Things Journal, 4(1), 144-152.

[2] Choi, E., et al. (2016). Retrospective analysis of clinical data with deep learning. Journal of the American Medical Informatics Association, 24(2), 350-359.

[3] Hameed, I. A., et al. (2019). Predicting heart disease using IoT data and Random Forest algorithm. Procedia Computer Science, 155, 630-637

[4] Harika Devi Kotha et al., " IoT Applications, a survey, " International Journal of Engineering Technology (UAE) (2018), Volume – 7, Issue – 2.7, PP: 891-896.

[5] Blaine Reeder et al., " Health at hand: A systematic review of the smartwatch uses for health and wellness," Journal of Biomedical Informatics 63 (2016) 269–276 PP.

[6] Spela Kosir, " A Look at Optical Sensors in Smart Wearables Technology ", [ONLINE], SOURCE: https://www.wearabletechnologies.com/2015/08/a-look-at-optical-sensors-in-smartwearablestechnology/, August 10, 2015 [4] Jon Phillips, " The 10 most likely sensors in a 10-sensor Apple smartwatch " [ONLINE], Source: https://www.pcworld.com/article/2366126/the-10-most-likely-sensorsin-a-10-sensor-applesmartwatch.html, 21 June 2014

[7] Wikipedia contributors, 'Smartwatch', Wikipedia, The Free Encyclopedia, 2 July 2018, [ONLINE], Source, https://en.wikipedia.org/w/index.php?title=Smartwatchold id=848486186

[8] Yung-Chieh Lin et al, " An electronic smartwatch monitors heart rate of an extremely preterm baby " [ONLINE], Source: http://www.pediatrneonatol.com, Aug 2017

[9] Elisa Morganti et al., " A smartwatch with embedded sensors to recognize objects, grasps and forearm gestures", Procedia Engineering 41 (2012) 1169 – 1175PP

[10] Tsung-Chien Lu, " Healthcare Applications of Smart Watches," Applied Clinical Informatics, (2016), 850-869PP. [9] Afef Samet Ellouze et al., " Integrating semantic dimension into openEHR archetypes for the management of cerebral palsy electronic medical records", Journal of Biomedical Informatics 63 (2016) 307–324 PP

[11] AA Khan, RM Mulajkar, VN Khan, SK Sonkar, DG Takale. (2022). A Research on Efficient Spam Detection Technique for IOT Devices Using Machine Learning. NeuroQuantology, 20(18), 625-631.

[12] SU Kadam, VM Dhede, VN Khan, A Raj, DG Takale. (2022). Machine Learning Methode for Automatic Potato Disease Detection. NeuroQuantology, 20(16), 2102-2106.

[13] DG Takale, Shubhangi D. Gunjal, VN Khan, Atul Raj, Satish N. Gujar. (2022). Road Accident Prediction Model Using Data Mining Techniques. NeuroQuantology, 20(16), 2904-2101.

[14] SS Bere, GP Shukla, VN Khan, AM Shah, DG Takale. (2022). Analysis Of Students Performance Prediction in Online Courses Using Machine Learning Algorithms. NeuroQuantology, 20(12), 13-19.

[15] R Raut, Y Borole, S Patil, VN Khan, DG Takale. (2022). Skin Disease Classification Using Machine Learning Algorithms. NeuroQuantology, 20(10), 9624-9629.

[16] SU Kadam, A katri, VN Khan, A Singh, DG Takale, DS. Galhe (2022). Improve The Performance Of Non-Intrusive Speech Quality Assessment Using Machine Learning Algorithms. NeuroQuantology, 20(19), 3243-3250.

[17] DG Takale, (2019). A Review on Implementing Energy Efficient clustering protocol for Wireless sensor Network. Journal of Emerging Technologies and Innovative Research (JETIR), Volume 6(Issue 1), 310-315.

[18] DG Takale. (2019). A Review on QoS Aware Routing Protocols for Wireless Sensor Networks. International Journal of Emerging Technologies and Innovative Research, Volume 6(Issue 1), 316-320.

[19] DG Takale (2019). A Review on Wireless Sensor Network: its Applications and challenges. Journal of Emerging Technologies and Innovative Research (JETIR), Volume 6(Issue 1), 222-226.

[20] DG Takale, et. al (May 2019). Load Balancing Energy Efficient Protocol for Wireless Sensor Network. International Journal of Research and Analytical Reviews (IJRAR), 153-158.

[21] DG Takale et.al (2014). A Study of Fault Management Algorithm and Recover the Faulty Node Using the FNR Algorithms for Wireless Sensor Network. International Journal of Engineering Research and General Science, Volume 2(Issue 6), 590-595.

[22] DG Takale, (2019). A Review on Data Centric Routing for Wireless sensor Network. Journal of Emerging Technologies and Innovative Research (JETIR), Volume 6(Issue 1), 304-309.

[23] DG Takale, VN Khan (2023). Machine Learning Techniques for Routing in Wireless Sensor Network, IJRAR (2023), Volume 10, Issue 1.