Application of XGBoost Algorithm for the Analysis of Healthcare Data at the Fog Layer

Syed Mujib Rahaman
Research Scholar Faculty of Computer Engineering
Pacific Academy of Higher Education and Research University
Udaipur, Rajasthan, India

Dilendra Hiran, PhD
Professor
Faculty of Computer Engineering
Pacific Academy of Higher Education and Research University
Udaipur, Rajasthan, India

Priyanka Kumari Bhansali, PhD
Assistant Professor Department of CS & SE
Andhra University College of Engineering(A), Visakhapatnam

ABSTRACT
The present paper focuses on the analysis of healthcare data using XGBoost algorithm at the Fog Layer. The present paper is based on an implementation of the Edge-Fog Layer layered architecture by considering the ESP8266 as the Edge node and Raspberry Pi is the Fog Node. The data has been captured, encrypted based on a custom-built encryption algorithm named SPADE and forwarded to the Raspberry Pi using the MQTT publish subscribe model. The Raspberry Pi acts as the data aggregator from multiple other Edge Nodes. The data within Raspberry Pi then analyzed based on the XGBoost algorithm after decryption which is done using Reverse SPADE algorithm. The results of analysis are then communicated to relevant application programs using FastAPI while also being stored on the Firebase database. The present framework provides a cost-effective implementation mechanism for analysis of healthcare data received from Edge nodes in a secure manner. The results of the analysis are presented in the form of prediction whether the health condition of the patient is critical. This results in the healthcare providers being able to initiate necessary healthcare procedures required to improve the healthcare condition of the patient.

General Terms
Internet of Things, Healthcare, Machine Learning Algorithms.

Keywords
Edge Computing, Fog Computing, Cloud Computing, XGBoost algorithm.

1. INTRODUCTION
Intelligent capture, storage, automated processing of healthcare data followed by real time analysis and consequent intelligent decision making is the major goal of IoT based healthcare systems. The easy availability of cost-effective components, reliable communication protocols and sufficient storage and processing capabilities to execute machine learning algorithms makes the possibility of creating an IoT based healthcare architecture economically and technically feasible. The present paper assumes that it is applied in the context of a healthcare provider mainly focusing on nursing homes or elderly care centers where there is the necessity of continuous but remote monitoring of healthcare parameters, including capture at the Edge Node, their transmission after capture to the Fog Node, and subsequent storage in cloud and analysis based on the chosen machine learning algorithm, XGBoost.

Although many variations of layered architectures have been proposed in earlier studies, the present study relies on the Sensor Layer-Edge Layer-MQTT-FastAPI-Fog Layer-Firebase Cloud Database architectural framework. The following section presents the SEMF3 Architectural Framework followed by details of the suitability of various architectural components. It is followed by a Literature Review of existing work in this area, Experimental Setup, Results and Discussion and Conclusion.

2. SEMF3 ARCHITECTURE
The Sensor Layer-Edge Layer-MQTT-Fast API-Fog Layer-Firebase Cloud Data base architectural framework is presented in Figure 1.

Communication Protocol: **MQTT** (Message Queuing Telemetry Transport) has been chosen as the communication protocol because:

1. It is a Light weight, Event-driven protocol that is based on the publish-subscribe architecture.
2. It has a small codebase and a header size of 2-5 bytes making it ideally suited to run on ESP8266.
3. As it is based on Publish-Subscribe model, it effectively decouples the Broker logic from the clients. This makes it possible for multiple ESP8266 nodes acting as MQTT Clients to talk to one Raspberry Pi Broker. This feature forms the foundation of future scalability of this model.
4. When the Raspberry Pi is acting as Client or Broker it supports Mosquitto.
5. With ESP8266 acting as Client and Raspberry Pi acting as Broker with Mosquitto, it handles communication between ESP8266 (Edge Node) and Raspberry Pi (Fog Node) locally on an offline basis.
6. With Raspberry Pi acting as a Client, it forwards messages to Cloud or Firebase Database cloud platforms.

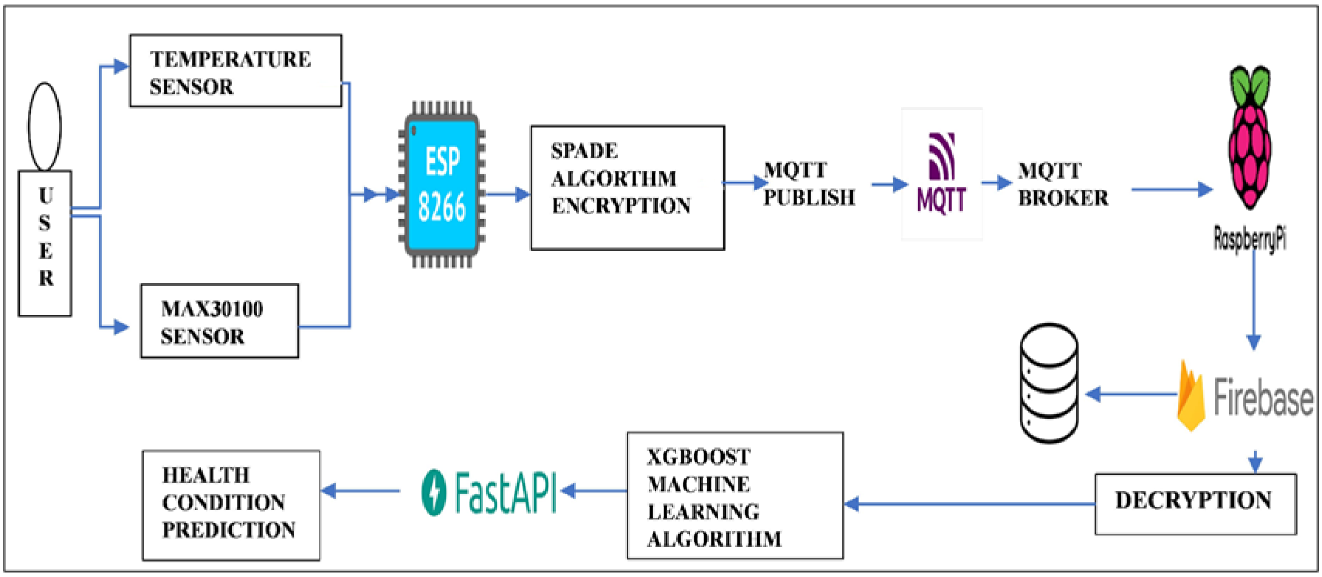

Figure 1 Semf3 Architecture

Fog Layer: **Raspberry Pi** has been chosen to act as the Fog Node because:

1. Raspberry Pi supports moderate processing power which makes it ideal for aggregating data from various ESP8266 Nodes.
2. It also provides support for lightweight Machine Learning Models which can be used to analyze the healthcare data from the sensors and then make predictions regarding the possible critical healthcare condition of the patient. These predictions can then be used to assist healthcare service providers in making informed therapeutic decision making.
3. It provides storage for sufficient RAM and storage necessary for logging patient data locally. This provides a facility for temporarily buffering healthcare data during unforeseen circumstances such as power outages.
4. For the purpose of communication, Raspberry Pi supports protocols such as MQTT, HTTP, HTTPS and CoAP. These can be used for reliably communication with Edge Nodes and Cloud Servers.

Database: **Firebase** Database has been chosen because:

1. It enables live synchronization of healthcare data across multiple dashboards. These include the web, mobile and cloud dashboards which are used by healthcare providers for real time monitoring of healthcare data.

As the API: **FastAPI** has been chosen for this purpose because:

1. It provides a secure, scalable and modern interface to manage the incoming data from ESP8266, and then perform the necessary preprocessing or validation.
2. It provides secure endpoints providing support to JWT, OAuth2 Security which can be used for validation and protecting data access.
3. It can be used for Encryption of healthcare data. It can be used to perform Edge-Level Encryption using S-Box. It can handle key indexing, offsetting, XOR with previous values using FastAPI Routes.
4. It can be used to store or forward data to Firebase database. FastAPI subscribes to Mosquitto topics and forwards data using Fast API logic.
5. It can provide real time access of healthcare data to healthcare service providers by providing service to secure APIs or dashboards. If offers REST endpoints to view/query sensor data in real time.
6. It can be used to control access of patient data either only to the caregiver or to any authorized application.

Machine Learning Algorithm applied at Fog Layer: **XGBoost algorithm** has been chosen to be applied at the Fog Layer (Raspberry Pi) for the purpose of analyzing the healthcare data received from the Edge Node (ESP8266). Being an optimized Gradient Boosting Framework, XGBoost implements Gradient Boosting which builds models sequentially where each model corrects the errors made by the previous models.

3. LITERATURE REVIEW

A review of the literature indicates that there have no previous studies covering the ESP8266-MQTT-Raspberry Pi-Firebase-FastAPI framework. Although a few studies covering a few of these components have been carried out, none of them have been applied in the healthcare context.

3.1 Choice of ESP8266 as Edge node

A collection of 62 Research Articles from the years 2017 to 2024 have been retrieved from the Mendeley database using the search key word " ESP8266 and Healthcare". These articles have been analyzed and a few representative samples have been represented in the following tables represented below in Table

Table 1 ESP8266 as Edge Node

Pathology	Sensors Used				Methodology Used	Reference
COVID-19; Telemedicine and e-health record database	Heart Rate & Pulse MAX30100	Temperature Sensor	ESP8266 WiFi	GPS for patient	Web Server and control panel used for remote access. Data base used for Electronic	[1]
IoT-Based E-Health Monitoring System for Pre-Schoolers	Hydration sensor DHT11	Saturation levels MAX30100	ESP8266 WiFi module	Mobile application for visualizing	Medicated pacifier where the respective hydration and oxygen saturation levels are calculated using DHT11 and MAX30100 sensors which are connected to the ESP8266 through which the data would be transferred wirelessly. Data visualization done on Blink mobile application.	[2]

Pathology	Sensors used		Methodology	Reference
Speech Disorders	Intel Edison	Raspberry Pi	The Intel Edison and Raspberry Pi that allows acquisition, computing, storage and communication of the various medical data such as pathological speech data of individuals with speech disorders, Phonocardiogram (PCG) signal for heart rate estimation, and Electrocardiogram (ECG)-based Q, R, S detection.	[3]
Heart Issues	Heart Rate, Respiration rate,	Body Temperature, Body Movements	This study, Monitors patient's heart rate, body temperature, Respiration rate and body movements using Raspberry Pi. After connecting Internet to the Raspberry Pi board, it act as a server. If these parameters are goes to abnormal, it will automatically send alert message to the doctor. The data send by Raspberry pi is stored on a server. The detailed information of patients and doctor registered through website on stored on server. The people to check health status with help of sensors and patients to access the website from anywhere.	[4]
Healthcare monitoring system			IoT-based, intelligent HMS which continuously monitors patient's health parameters like blood pressure, heart beat and ECG. Data from blood pressure sensor, heart rate sensor and ECG sensor automatically monitored by Arduino UNO and Pi-camera attached to raspberry pi for video. Arduino UNO sends sensor data to raspberry pi which fed data to server's database using Wi-Fi, finally server sent data to webpage, which updates every 2 minutes. Doctor access data anywhere using internet and give feedback accordingly using text.	[5]
ESP8266 based web server		AdaFruit IO Cloud providing security and the ability to easily access the data to healthcare providers.	This project uses sensors such as the ESP8266 and max30100 and the Adafruit IO Cloud, which provide security and the ability to view all data, making it useful for clinicians during treatment. If the system deems it necessary, a warning is also generated at the time of any critical conditions.	[6]
Heart Rate and Blood Oxygen Sensor (MAX30100)	Infrared Body Temperature Sensor (MLX90614) ECG sensor (AD8232).	Arduino Uno acts as the central module to collect and process these signals. ESP8266 Wi-Fi module used to transmit signals to database for Deep Learning Analysis based on LSTM for temperature, heart rate and oxygen saturation prediction while CNN was used to analyze ECG signals from the Image datasets.	Deep learning tools used were the powerful python language, python-based Anaconda, Google' s TensorFlow and open-source neural network library Keras. The algorithm was used for evaluation using the available MIT-BIH ECG database from Physionet databases which attained 99.05% accuracy and arrived at only 4.96% loss rate after 30 training steps.	[7]

An analysis of the result of Literature Review indicates that:.

- ESP8266 has been used for the process of aggregating data from the various sensors.
- Most of the IoT applications focused on sensors required for remotely monitoring and analyzing the body condition of the patient. For this purpose, sensors such as Temperature Sensor, Pulse Oximeter and Heart Beat Sensor, Blood Pressure Sensor, ECG Sensor, Blood Glucose Sensor are the most frequently used Sensors.
- ESP8266 is being used for the purpose of analysis of the sensor information at the Edge Node while also allowing the sensor data to be transmitted via its Wi-Fi Module to storage databases or to access points of applications responsible for forwarding this data to healthcare providers.
- Healthcare applications using ESP8266 also utilized the service of Cloud services such as AdaFruit in order to help with providing to secure healthcare data logging, storage after analysis.
- Most healthcare applications incorporated mechanism to not only monitor the healthcare parameters of the patient remotely but also provided a functionality of informing and if necessary escalating the healthcare condition to the healthcare providers in case of any abnormalities in the observed sensor values.
- Few IoT applications performed extensive analysis of healthcare data by using Deep Learning Algorithms which were implemented on Arduino Uno as the computing capabilities of ESP8266 are not sufficient to perform such computationally intensive tasks. In this particular research work, [7], the Arduino Uno seems to be used as the Fog Node which definitely requires better computing, storage and analysis capabilities when compared to the Edge Node. This resulted in the present work focusing on trying to find a Fog Node which was both technically and economically feasible to suit the needs of a healthcare provider primarily focusing on nursing homes and elderly care facilities where both the budget and the technical capabilities are a major challenge.
- It is in this context that Raspberry Pi was found to be the best possible option for a Fog Node which has the computational capabilities required to execute the XGBoost Machine Learning Algorithm for the purpose of analyzing the healthcare data of the patients.

3.2 Choice of Raspberry Pi as Fog Node

A collection of 26 Research Articles have been accessed from the Mendeley Database with the search key "Raspberry Pi in Healthcare". These articles were studied with the aim of trying to identify the suitability of Raspberry Pi to be used in healthcare context. This would be evident in case the research articles provide case studies of IoT-based healthcare implementations based on Raspberry Pi as the Fog Node. As a representative sample a few of these studies are being summarized in the following tables.

Table 2. Raspberry Pi in Healthcare

Pathology	Sensors Used			Methodology Used	Reference
IoT-Based Control System to Measure, Analyze, and Track Basic Vital Indicators in	Arduino Mega 2560	ESP8266 Wi-Fi Modules in an IoT ecosystem	Body temperature and Respiration rate sensors	The aims of the project are to create a healthcare monitoring system which can detect physiological parameters levels, analyze vital sign levels depending on the patient age, offer alerts for problematic conditions, and remotely show data using Android applications.	[8]
IOT Based Patient Monitoring System for Stroke Affected Patients	Blood Pressure sensor	Heart Beat sensor	Blood Glucose, Temperature	Informing the doctor and caretaker about variation in risk factors such as Blood pressure, Heart beat, Blood glucose, Temperature etc. The current health parameters are monitored using sensors and the values from sensors are stored in the cloud	[9]
Real-Time Analysis of Wearable Sensor Data Using IoT and Machine Learning in Healthcare	LM35 Temperature sensor, Pulse Oximeter Sensor	Arduino UNO Controller board communicates with each of these sensors.	Transmissions of ESP8266 Wi- Fi module are tracked by use of Augmented Reality Glasses. AdaFruit, an IoT Platform stores the sensor data reserved by ESP8266 module making the patient medical history accessible on demand.	Abnormal sensor values result in alarms to the healthcare providers facilitating remote monitoring of patient health condition.	[10]

Pathology	Sensors used	Methodology	References
Heart disease	Public database containing ECG images was used.	study aims to develop algorithmic models to analyze ECG tracings to predict cardiovascular diseases. This study conducted numerous experiments to optimize deep-learning parameters of both MobileNetV2 and VGG16 algorithms implementation on Raspberry Pi for cardiovascular disease diagnosis and prediction.	[11]
Heart Disease but can also be extended	ECG Graphs, SpO2 Sensor, Temperature Sensors, Heart Rate Sensors	Arduino Uno acts as the Edge Node and Raspberry Pi is used as the Fog Node although not specifically mentioned in this way. Arduino Uno acquires data from sensors and sends it to Raspberry Pi which in turn forwards it a server. The healthcare information is made accessible to the healthcare providers through smartphones.	[12]
Heart Disease	ECG Signals connected to Raspberry Pi and GSM Module for communication with healthcare provider	ECG machines are connected to Raspberry Pi in which a program monitors the parameters continuously. If any abnormal values are observed, a message is sent to the healthcare providers through a GSM module. MySQL dB has been used to update the website database continuously.	[13]
Low-cost Remote Primary Healthcare Services through Telemedicine	Raspberry Pi 3 Model B Blood Pressure, ECG, Height, Weight, SpO2, Airflow, Body position, temperature, glucometer sensor.	This paper demonstrates how a low-cost Remote Primary Healthcare Services can be established in Bangladesh. Here Raspberry Pi 2 B Model is used for the client-module to connect with the health system. Portable Telemedicine Tool kit has been developed using Arduino Uno, e-health shield and Bluetooth.	[14]

Analysis of a few representative samples of the use of Raspberry Pi in healthcare indicates that:

- Raspberry Pi owing to its low cost but sufficient computational capacity provided by the ARM processor makes it the ideal component to build any health care application whether it is a highly computation intensive Deep learning-based application as well as a simple primary healthcare-based application.
- Its low cost and low power capabilities make it to be the right choice for being used as the Fog Node.

3.3 Choice of ESP8266 and Raspberry Pi in healthcare context

[42] presents a study involving Heart Beat Detection and Monitoring. In this case the responsibility of detecting the heartbeat using the Pulse Sensor lies with the Raspberry Pi. ESP8266 is responsible for transmitting the sensor values to the ThingSpeak server from which they can be accessed by healthcare providers. This study indicates how computationally complex tasks can be assigned to the Raspberry Pi while computationally light tasks are usually assigned to ESP8266. This study also indicates that most healthcare applications use the Wi-Fi ESP8266 for transmission of healthcare data to storage servers or cloud. [15] is a study that indicates the use of MQTT protocol for the purpose of secure communication in IoT based healthcare applications.

3.4 Choice of XGBoost Algorithm as Machine Learning Algorithm applied at Fog Node

References have been gathered from Mendeley Cite database to try to analyze preexisting trends where the XGBoost Algorithm was being applied in addition to understanding any possible algorithmic variants.

A collection of 63 Research Articles was retrieved from the Mendeley Database using the key words "XGBoost and Healthcare Prediction" in order to identify the recently published articles during the years 2020 to 2024 covering the areas of machine learning algorithms specifically XGBoost algorithms and their applications in Healthcare. Among the results displayed the top 65 results were chosen based on suitability and relevance. Out of these around three Research Articles were not considered for the analysis as they were suitable for the problem being considered which is the Appropriateness of XGBoost Machine Learning Algorithm for analyzing the healthcare data.

Analysis of the resulting research articles indicates that:

1. Machine Learning Algorithms specifically XGBoost Algorithms have been used extensively in a wide range of pathologies including COVID-19, Type-2 Diabetes Mellitus, Heart Disease, Kidney Disease, apart from a whole range of issues covering Healthcare provider administration, and other pathologies.
2. In articles covering COVID-19, focus was seen on

the application of XGBoost machine learning algorithm used to predict severity of symptoms in addition to symptom analysis [16] ,[17]. These machine learning algorithms have also been used to predict stress in the healthcare industry during COVID-19, [17] as well as predict mortality due to COVID-19 [18].

3. In the studies focusing on Diabetes Mellitus, XGBoost machine learning algorithm have been used for Prediction, [19], [20], [21], [22] , [23] Studies also focus on the risk of progressing from Diabetes Mellitus to comorbidities such as renal disease and fracture [24] , [25].
4. In the case of heart disease, few studies indicate the use of XGBoost algorithm for Identification of Risk factors, [26], Prediction and analysis of disease [27], [28]. Studies also propose frameworks for early detection of cardiac arrest condition and risk [29]. Studies also indicate that these machine learning algorithms have been used for prediction of Heart Failure after a medical procedure, [30].
5. Further studies indicate application of XGBoost Algorithm based frameworks for Health Prediction, [31], Risk Prediction, [32], Improving diagnosis of disease [33] as well as for Disease Detection and Classification, [34].
6. Few studies indicate the use of Machine Learning enriched with the interpretative insights provided by Explainable Artificial Intelligence [35], [36].
7. Studies indicate the use of Machine Learning to predict emergency department admissions [37], ambulatory non- arrivals, [38], late arrivals at the adult outpatient department, [39], length of stay in Intensive Care Unit, [40] and detection of Health Insurance Fraud [41]. Studies have been carried out focusing on COVID-19 Risk Stratification and Mortality Prediction in Hospitalized Indian Patients, [42]. Studies have been done covering COVID-19 Predicting maternal and new born healthcare providers perception of safety [43]. Studies focusing on Hospital Administration including Workflow management [44] as well as predicting emergency department disposition are available in the literature, [45].
8. XGBoost machine learning algorithms have been applied to cancer, cervical cancer and prostate cancer. Studies indicate how a comparative study of machine learning algorithms has been applied for predictive risk assessment of cervical cancer, [46]. Machine learning has been combined with Pattern analysis for early identification of Prostate Cancer, [47]. Studies comprising of a combination of precision Cancer classification using liquid biopsy and advanced machine learning techniques has been done, [48].
9. Studies covering use of XGBoost algorithm to identify risk factors for insomnia, [49], early prediction of blood stream infection, [50] , Chronic Kidney infection [51], and Urine infection, [52], Non-alcoholic Fatty Liver, [53], [54]. Prediction of Childhood Obesity, [55], Life Expectancy, [56], and Population-Level Mortality [57] has also been done as indicated in these studies.
10. XGBoost algorithm provides better performance in terms of less prediction errors, and for being faster and suitable to be used on resource constrained devices. Having the capability to exploit CPU Cache and multithreaded execution, XGBoost can be used for training or for gathering inferences even on devices with resource limitations. For example, the Raspberry Pi having a cost of around 35$ usually supports RAM capacity of up to 8GB and works with a power usage in the range 2.7-7 W. The XGBoost algorithm can be used to train models and draw inferences on such resource constrained devices such as Raspberry Pi. Devices such as Raspberry Pi are most suitable to be used as Fog Nodes in Nursing Homes and Elderly care settings where the primary task is to continuously and economically monitor the health condition of the individuals and report them immediately to the healthcare providers without any delay.

3.5 Previous Studies covering ESP8266-MQTT-Raspberry Pi-Firebase-FastAPI in healthcare context

There exist no prior studies covering all these various architectural components which have been applied in the healthcare context.

4. EXPERIMENTAL SETUP

The main aim of the SEMF3 Architecture is to classify medical conditions as Critical or Non-Critical based on Body Temperature (Celsius), Blood Oxygen Saturation (SpO2 Percentage) and Heart Beats Per Minute (BPM). After the initial set up involving ESP8266-MQTT-FastAPI-Raspberry Pi and Firebase has been completed, the XGBoost Classification Algorithm is applied.

Data set Characteristics

1. Size: 2194 Samples
2. Features: 4 Columns (3 Input Features + 1 Target)
3. Data Types: Integer (3 Columns), Float(1 Column)
4. Memory Usage:15.6 KB
5. Missing Values: None

Table 3 Feature Set

Feature	Mean	Std Dev	Mi n	Ma x	25th%	75th%
HeartBeat_B PM	90.76	29.41	40	149	67	114
Temperature_C	37.37	1.39	34.1	40.7	36.7	38.3
SpO2_percent	93.04	4.45	82	100	90	97

Data Preprocessing

Outlier Handling: Outliers have been eliminated and extreme values were capped to clinically reasonable ranges:

- Heart Rate: 50-140 BPM
- Temperature: 35-40°C
- SpO2: 85-100%

Feature Engineering

Three interaction terms were created to capture relationships between vital signs:

- BPM_Temp: HeartBeat_BPM × Temperature_C
- BPM_SpO2: HeartBeat_BPM × SpO2_percent
- Temp_SpO2: Temperature_C × SpO2_percent

Feature Scaling

StandardScaler was applied to normalize all features to have mean=0 and std=1, ensuring equal contribution from all features during training.

Data Splitting

- Training Set: 80% (2194 samples)
- Test Set: 20% (296 samples)
- Stratification: Applied to maintain class distribution across splits

Feature Statistics:

Data Preprocessing

Outlier Handling: Outliers have been eliminated and extreme values were capped to clinically reasonable ranges:

- Heart Rate: 50-140 BPM
- Temperature: 35-40°C
- SpO2: 85-100%

Feature Engineering

Three interaction terms were created to capture relationships between vital signs:

- BPM_Temp: HeartBeat_BPM × Temperature_C
- BPM_SpO2: HeartBeat_BPM × SpO2_percent
- Temp_SpO2: Temperature_C × SpO2_percent

Feature Scaling

StandardScaler was applied to normalize all features to have mean=0 and std=1, ensuring equal contribution from all features during training.

Data Splitting

- Training Set: 80% (2194 samples)
- Test Set: 20% (296 samples)
- Stratification: Applied to maintain class distribution across splits

4.1 Model Development Cross-Validation Results

5-fold cross-validation was performed on the training set:

- Individual Fold Scores: [98.73%, 100%, 97.47%,
- 93.67%, 98.73%]
- Mean CV Accuracy: 97.72%
- Standard Deviation: 2.26%

4.2 Hyperparameter Tuning

- GridSearchCV was used to optimize model parameters: Search Space
- n_estimators: [100, 200]
- max_depth: [3, 5, 7]
- learning_rate: [0.01, 0.1]
- subsample: [0.8, 1.0] Optimal Parameters
- n_estimators: 100
- max_depth: 5
- learning_rate: 0.1
- subsample: 0.8
- Best CV Accuracy: 97.97%

Model Performance Test Set Results

- Overall Accuracy: 97.98%
- Total Predictions: 99
- Correct Predictions: 97
- Incorrect Predictions: 2

Table 4 Confusion Matrix

		Predicted	
		Non-Critical	Critical
Actual	Non-Critical	40	0
	Critical	2	57

Class-wise Performance Metrics Non-Critical Class (0)

- Precision: 95.24% (40/42)
- Recall: 100% (40/40)
- F1-Score: 97.56%
- Support: 40 samples Critical Class (1)
- Precision: 100% (57/57)
- Recall: 96.61% (57/59)
- F1-Score: 98.28%
- Support: 59 samples Aggregate Metrics
- Macro Average: Precision: 97.62%, Recall: 98.31%, F1-Score: 97.92%
- Weighted Average: Precision: 98.08%, Recall: 97.98%, F1-Score: 97.98%

Feature Importance Analysis

The model identified the following feature importance ranking which has been presented in Table 5.

Table 5 Feature Importance

	Feature	Importance	Interpretation
1	SpO2_percent	66.48%	Blood oxygen saturation is the most critical indicator

2	HeartBeat_BPM	12.29%	Heart rate provides secondary diagnostic value
3	Temp_SpO2	9.59%	Interaction between temperature and oxygen levels
4	Temperature_C	8.17%	Body temperature as standalone feature
5	BPM_Temp	2.26%	Heart rate-temperature interaction
6	BPM_SpO2	1.20%	Heart rate-oxygen interaction

Key Insights:

Experimental results indicate that Blood Oxygen Saturation accounts for nearly two-thirds of the model's decision making thereby dominating the other model features. It also indicates that the combined importance of interaction terms (13.05%) suggests non-linear relationships. Furthermore, Temperature feature although may not be important individually, but it becomes important in combination with other features.

4.3 Model Interpretation

Clinical Relevance

The model is clinically relevant as the Feature importance as observed in the model aligns with medical knowledge. SpO2, which is also called as Blood Oxygen Saturation is responsible for assessing respiratory function and blood oxygen delivery. Heart rate is an indicator of cardiovascular stress and metabolic demand while Temperature reflects inflammatory processes and metabolic state.

Decision Boundaries

The model learned complex, non-linear decision boundaries through Tree-based splits on individual features, Gradient boosting for error correction and Feature interactions capturing physiological relationships.

4.4 Model Deployment

Saved Artifacts

1. xgb_model.pkl: Trained XGBoost classifier
2. scaler.pkl: StandardScaler for feature preprocessing

Prediction Pipeline

```
# Load model and scaler
model = joblib.load('xgb_model.pkl') scaler =
joblib.load('scaler.pkl')
# New patient data
new_patient = [120, 38.5, 89] # [BPM,
Temp, SpO2] # Feature engineering
interactions = [  new_patient[0] *
new_patient[1], # BPM_Temp
new_patient[0] * new_patient[2], #
BPM_SpO2 new_patient[1] *
new_patient[2]  # Temp_SpO2
]

full_features = new_patient +
interactions # Scale and
predict
scaled_features = scaler.transform([full_features])
prediction = model.predict(scaled_features)[0]
```

Technical Specifications

Environment Requirements

- Python 3.7+
- pandas, numpy, scikit-learn
- xgboost
- matplotlib, seaborn (for visualization)
- joblib (for

model persistence) Model

Hyperparameters

XGBClassifier (

n_estimators=100, max_depth=5,

learning_rate=0.1, subsample=0.8,

random_state=42)

Preprocessing Configuration StandardScaler (

with_mean=True, with_std=True

)

5. RESULTS AND DISCUSSION

The experimental results which have been obtained are being presented in Table 6. Here the healthcare parameters including Temperature, Heartrate and Blood Oxygen Saturation Levels have been indicated and the result of prediction of healthcare status of the patient whether it is critical or not is indicated using Normal/ Critical.

5.1 Performance Validation

Strengths: The results indicate that the model performs well as indicated by high accuracy, balanced performance, robust cross-validation and clinical interpretability. The model provides an accuracy of 97.98% which indicates excellent performance. The model provides balanced performance as indicated by Good Precision and Recall for both classes. The model provides robust cross-validation as indicated by consistent performance across folds. The results provided by the model are clinically interpretable as indicated by the fact that the feature importance suggested by the model coincides with medical knowledge

Table 6 Results indicating Prediction of Patient Health Condition based on XGBoost Algorithm

Timestamp	Temperature (C)	Heart Rate (bpm)	SpO2 (%)	Health Status	Condition Message
2:42:20	5.6	71	97	Normal	Vital signs are within normal range
2:42:20	35.6	71	97	Normal	Vital signs are within normal range
2:42:21	35.6	75.8	97	Normal	Vital signs are within normal range

2:42:21	35.6	75.8	97	Normal	Vital signs are within normal range
2:42:21	35.6	75.8	97	Normal	Vital signs are within normal range
2:42:22	35.6	76.4	97	Normal	Vital signs are within normal range
2:42:22	35.6	76.4	97	Normal	Vital signs are within normal range
2:42:22	35.6	76.4	97	Normal	Vital signs are within normal range
2:42:23	35.6	78	97	Normal	Vital signs are within normal range
2:42:23	35.6	78	97	Normal	Vital signs are within normal range
2:42:23	35.6	78	97	Normal	Vital signs are within normal range
2:42:24	35.6	80.6	97	Normal	Vital signs are within normal range

5.2 Model usage recommendations
The experimental results indicate that the model can be recommended to be used for initial triage of Patients i.e, for initial Screening purpose. As the model provides alerts for predicted critical cases it can be used as an Alert System at healthcare providers with resource constraints. The model provides healthcare condition predictions which can be used to supplement clinical judgement rather than replace it. Thus this model acts as a Decision support system in the healthcare context. Regular performance assessments may be done with new data in order to provide continuous monitoring of patient health condition by healthcare providers.

5.3 Limitations
For the purpose of the present work a data set having a size of only 2194 samples have been considered. The limited size of the dataset may limit the extent to which the model results may be generalized. Furthermore, in order to track patient health condition only three vital signs including Temperature, Heart Beat and Oxygen Saturation are being considered. This limits the scope of the number of features that may be responsible for the patient health condition. This model may be extended to incorporate multiple vital signs covering varied health conditions. However, the model suffers from a slight bias towards critical cases as compared to non-critical or healthy cases. Additionally, no consideration has been given to vital sign trends over time.

5.4 Risk Assessment
The present model misclassified 3.39% of critical cases i.e., the number of False Negatives is 2, whereas the number of False Positives is 0 which means that no non- critical case was misclassified. It was observed that this model overpredicts critical cases, an aspect representing that this model is Clinically safe.

5.5 Monitoring and Maintenance
The performance of the model may be enhanced by monitoring the prediction accuracy of the model on new data. By updating the model quarterly with new samples, model retraining may be achieved. By monitoring changes in vital sign distributions, Feature Drift may be observed. By regularly reviewing the obtained results with medical professionals, Clinical Validation of the Model can be achieved. This clinical validation would provide the feedback necessary for making the necessary modifications so that model performance perfectly matches their corresponding clinical interpretations.

6. CONCLUSIONS
The present work has demonstrated how the confluence of IoT and Machine Learning can be used cost effectively for the purpose of predicting the health condition of the patient while giving utmost priority to security of the patient data. This has been achieved by the design, development and use of the SEMF3 Architecture. The present work further reinstates that the ESP8266 and Raspberry Pi combination is ideally suited as the Edge Node – Fog Node combination in healthcare contexts facing resource constraints. ESP8266 receives patient data using the temperature, heart rate and oxygen saturation sensors which it encrypts using SPADE algorithm. The present work indirectly demonstrates that SPADE algorithm can be used for healthcare data encryption at the Edge node. MQTT publish and subscribe have been used for the purpose of reliable transfer of encrypted data to the Fog Node. It is at the Fog Node i.e., Raspberry Pi that the XGBoost Machine Learning algorithm is applied after decrypting the encrypted data using the Reverse SPADE algorithm. The present paper limits itself to the use of Machine Learning algorithms for the purpose of enhancing the quality of service provided by healthcare providers in terms of online monitoring and health condition prediction and reporting. The XGBoost Machine learning algorithm has been chosen as a suitable machine learning algorithm in the present context. The XGBoost medical condition classification model demonstrates excellent performance for predicting critical conditions based on vital signs. With 97.98% accuracy and clinically interpretable feature importance, it provides a reliable tool for medical triage and decision support. The model's conservative approach (zero false positives for non-critical cases) makes it suitable for healthcare applications where missing critical cases is more costly than over-alerting. The present paper demonstrates the use of FastAPI framework for providing RESTful endpoints for triggering predictions and receiving and sending responses to the healthcare providers. The present work demonstrates the utility of Firebase cloud database for the purpose of storage and visualization of results. From the Data Security perspective, the present work demonstrates that [a] SPADE algorithm ensures data confidentiality at the Edge Node before transmission, [b] MQTT provides secure data transport between the Edge and Fog Nodes, [c] FastAPI prevents unauthorized access while [d] Firebase Security Rules can be used to limit access based on user rules. Thus the above paper demonstrates data security from the Edge Node to the Fog Node in the healthcare context.

6.1 Future Improvements

As discussed earlier the data set considered initially was limited to 2194 samples. Data expansion may be achieved by collecting more samples. This would result in better generalization of obtained results. Extending the feature set to include more features and including patient history would extend the model applicability. Furthermore, time-series analysis may be included in order to provide trend detection. Finally, the model presently utilizes the XGBoost algorithm. This may be extended to include other ensemble machine learning algorithms in order to provide for improved robustness and prediction accuracy.

7. REFERENCES

[1] Palli, G. H., Mirza, G. F., & Chowdhry, B. S. (2022). Novel IoT-Based E-Health System: Hospital Management, Telemedicine and Quarantine Management for COVID-19. *Proceedings of 3rd International Conference on Latest Trends in Electrical Engineering and Computing Technologies, INTELLECT 2022*. https://doi.org/10.1109/INTELLECT55495.2022.9969396

[2] Samal, S., & Mitra, T. (2023). IoT-Based E-Health Monitoring System for Pre-Schoolers. In *Applications of Artificial Intelligence in the Healthcare Sector*.

[3] Dubey, H., Monteiro, A., Constant, N., Abtahi, M., Borthakur, D., Mahler, L., Sun, Y., Yang, Q., Akbar, U., & Mankodiya, K. (2017). *Fog Computing in Medical Internet-of-Things: Architecture, Implementation, and Applications*. https://doi.org/10.1007/978-3-319-58280-1_11

[4] Selvarathi, C., & Sujatha, R. (2018). Frequent updater for health monitoring system using Raspberry Pi. *International Journal of Engineering and Technology (UAE)*, *7*(3.34 Special Issue (34). https://doi.org/10.14419/ijet.v7i3.34.18975

[5] Keserwani, H., Kakade, S. v., Sharma, S. K., Manchanda, M., & Nama, G. F. (2023). Real-Time Analysis of Wearable Sensor Data Using IoT and Machine Learning in Healthcare. International Journal of Intelligent Systems and Applications in Engineering, 11(7s).

[6] Hatem, A. S., Abed, J. K., & Ajel, A. R. (2023). Based on IoT: Design and implementation of health care monitoring system. *AIP Conference Proceedings*, *2804*(1). https://doi.org/10.1063/5.0155061

[7] Panganiban, Edward & Chung, Wen-Yaw & Tai, Wei-Chieh & Paglinawan, Arnold & Lai, Jheng-Siang & Cheng, Ren & Chang, Ming-Kai & Chang, Po-Hsuan. (2020). Real-Time Intelligent Healthcare Monitoring and Diagnosis System Through Deep Learning and Segmented Analysis. 10.1007/978-3-030-30636-6_3.

[8] Upreti, K., Vats, P., Haque, M., Goyal, A. A., Wankhede, S. R., Singh, P., Alam, M. S., & Nasir, M. S. (2023). IoT-Based Control System to Measure, Analyze, and Track Basic Vital Indicators in Patient Healthcare Monitoring System. *Communications in Computer and Information Science*, *1749 CCIS*. https://doi.org/10.1007/978-3-031-25088-0_63

[9] Hatem, A. S., Abed, J. K., & Ajel, A. R. (2023). Based on IoT: Design and implementation of health care monitoring system. *AIP Conference Proceedings*, *2804*(1). https://doi.org/10.1063/5.0155061

[10] [10] Keserwani, H., Kakade, S. v., Sharma, S. K., Manchanda, M., & Nama, G. F. (2023). Real-Time Analysis of Wearable Sensor Data Using IoT and Machine Learning in Healthcare. International Journal of Intelligent Systems and Applications in Engineering, 11(7s).

[11] Mhamdi, L., Dammak, O., Cottin, F., & Dhaou, I. Ben. (2022). Artificial Intelligence for Cardiac Diseases Diagnosis and Prediction Using ECG Images on Embedded Systems. *Biomedicines*, *10*(8). https://doi.org/10.3390/biomedicines10082013

[12] Lakshmi, G. J., Ghonge, M., & Obaid, A. J. (2021). Cloud based iot smart healthcare system for remote patient monitoring. *EAI Endorsed Transactions on Pervasive Health and Technology*, *7*(28). https://doi.org/10.4108/eai.15-7-2021.170296

[13] Gupta, M. S. D., Patchava, V., & Menezes, V. (2016). Healthcare based on IoT using Raspberry Pi. *Proceedings of the 2015 International Conference on Green Computing and Internet of Things, ICGCIoT 2015*. https://doi.org/10.1109/ICGCIoT.2015.7380571

[14] Prodhan, U. K., Saha, T. K., Shaharin, R., Emon, T. A., & Rahman, M. Z. (2020). Implementation of Low-Cost Remote Primary Healthcare Services through Telemedicine: Bangladesh Perspectives. *International Journal of Advanced Computer Science and Applications*, *11*(11). https://doi.org/10.14569/IJACSA.2020.0111118

[15] Segarra, C., Delgado-Gonzalo, R., & Schiavoni, V. (2020). MQT-Tz: Secure MQTT broker for biomedical signal processing on the edge. *Studies in Health Technology and Informatics*, *270*. https://doi.org/10.3233/SHTI200177

[16] Nambiar, A., Harikrishnaa, S., & Sharanprasath, S. (2023). Model-agnostic explainable artificial intelligence tools for severity prediction and symptom analysis on Indian COVID-19 data. *Frontiers in Artificial Intelligence*, *6*. https://doi.org/10.3389/frai.2023.1272506

[17] Gupta, P., Maji, S., & Mehra, R. (2022). Predictive Modeling of Stress in the Healthcare Industry during COVID-19: A Novel Approach Using XGBoost, SHAP Values, and Tree Explainer. *International Journal of Decision Support System Technology*, *15*(1). https://doi.org/10.4018/IJDSST.315758

[18] Sharifi-Kia, A., Nahvijou, A., & Sheikhtaheri, A. (2023). Machine learning-based mortality prediction models for smoker COVID-19 patients. *BMC Medical Informatics and Decision Making*, *23*(1). https://doi.org/10.1186/s12911- 023-02237-w

[19] Nagassou, M., Mwangi, R. W., & Nyarige, E. (2023). A Hybrid Ensemble Learning Approach Utilizing Light Gradient Boosting Machine and Category Boosting Model for Lifestyle-Based Prediction of Type-II Diabetes Mellitus. *Journal of Data Analysis and Information Processing*, *11*(04). https://doi.org/10.4236/jdaip.2023.114025

[20] Tiwari, P., & Singh, V. (2021). Diabetes disease prediction using significant attribute selection and classification approach. *Journal of Physics: Conference Series*, *1714*(1). https://doi.org/10.1088/1742-6596/1714/1/012013

[21] Modak, S. K. S., & Jha, V. K. (2024). Diabetes prediction model using machine learning techniques. *Multimedia Tools and Applications*, *83*(13).

https://doi.org/10.1007/s11042-023-16745-4

[22] Ahamed, B. S., Arya, M. S., & Nancy V, A. O. (2022). Prediction of Type-2 Diabetes Mellitus Disease Using Machine Learning Classifiers and Techniques. In *Frontiers in ComputerScience* (Vol. 4). https://doi.org/10.3389/fcomp.2022.835242

[23] Lu, H., Uddin, S., Hajati, F., Moni, M. A., & Khushi, M. (2022). A patient network-based machine learning model for disease prediction: The case of type 2 diabetes mellitus. *Applied Intelligence*, *52*(3). https://doi.org/10.1007/s10489-021-02533-w

[24] Ou, S. M., Tsai, M. T., Lee, K. H., Tseng, W. C., Yang, C. Y., Chen, T. H., Bin, P. J., Chen, T. J., Lin, Y. P., Sheu, W.H. H., Chu, Y. C., & Tarng, D. C. (2023). Prediction of the risk of developing end-stage renal diseases in newly diagnosed type 2 diabetes mellitus using artificial intelligence algorithms. *BioData Mining*, *16*(1). https://doi.org/10.1186/s13040-023-00324-2

[25] Chen, Y., Yang, T., Gao, X., & Xu, A. (2022). Hybrid deep learning model for risk prediction of fracture in patients with diabetes and osteoporosis. *Frontiers of Medicine*, *16*(3). https://doi.org/10.1007/s11684-021-0828-7

[26] Huang, A. A., & Huang, S. Y. (2023b). Use of machine learning to identify risk factors for insomnia. *PLoS ONE*, *18*(4 April). https://doi.org/10.1371/journal.pone.0282622

[27] Yang, J. C. (2024). The prediction and analysis of heart disease using XGBoost algorithm. *Applied and Computational Engineering*, *41*(1). https://doi.org/10.54254/2755-2721/41/20230711

[28] Ganie, S. M., Pramanik, P. K. D., Malik, M. B., Nayyar, A., & Kwak, K. S. (2023). An Improved Ensemble Learning Approach for Heart Disease Prediction Using Boosting Algorithms. *Computer Systems Science and Engineering*, *46*(3). https://doi.org/10.32604/csse.2023.035244

[29] Shah, A., Ahirrao, S., Pandya, S., Kotecha, K., & Rathod, S. (2021). Smart Cardiac Framework for an Early Detection of Cardiac Arrest Condition and Risk. *Frontiers in Public Health*, *9*. https://doi.org/10.3389/fpubh.2021.762303

[30] Zhang, L., Zhou, X., & Cao, J. (2024). Establishment and validation of a heart failure risk prediction model for elderly patients after coronary rotational atherectomy based on machine learning. *PeerJ*, *12*. https://doi.org/10.7717/peerj.16867

[31] Kumar, D., Sood, S. K., & Rawat, K. S. (2023). Early health prediction framework using XGBoost ensemble algorithm in intelligent environment. *Artificial Intelligence Review*, *56*. https://doi.org/10.1007/s10462-023-10565-6

[32] Li, Y., Cao, W., Chang, W., Zhou, S., & Zhang, R. (2022). An XGBoost risk prediction model of cardiovascular and cerebrovascular diseases with plateau healthcare dataset. *ACM International Conference Proceeding Series*. https://doi.org/10.1145/3573834.3574558

[33] Sharma, A., & Verbeke, W. J. M. I. (2020). Improving Diagnosis of Depression with XGBOOST Machine Learning Model and a Large Biomarkers Dutch Dataset (n=11,081). *Frontiers in Big Data*, *3*. https://doi.org/10.3389/fdata.2020.00015

[34] Kumar, M., Singhal, S., Shekhar, S., Sharma, B., & Srivastava, G. (2022). Optimized Stacking Ensemble Learning Model for Breast Cancer Detection and Classification Using Machine Learning. *Sustainability (Switzerland)*, *14*(21). https://doi.org/10.3390/su142113998

[35] Palkar, A., Dias, C. C., Chadaga, K., & Sampathila, N. (2024). Empowering Glioma Prognosis with Transparent Machine Learning and Interpretative Insights Using Explainable AI. IEEE Access, 12. https://doi.org/10.1109/ACCESS.2024.3370238

[36] Khanna, V. V., Chadaga, K., Sampathila, N., Prabhu, S., & Rajagopala Chadaga, P. (2023). A machine learning and explainable artificial intelligence triage-prediction system for COVID-19. *Decision Analytics Journal*, *7*.

[37] Rocha, C. N., & Rodrigues, F. (2021). Forecasting emergency department admissions. *Journal of Intelligent Information Systems*, *56*(3). https://doi.org/10.1007/s10844-021-00638-9

[38] Coppa, K., Kim, E. J., Oppenheim, M. I., Bock, K. R., Zanos, T. P., & Hirsch, J. S. (2023). Application of a Machine Learning Algorithm to Develop and Validate a Prediction Model for Ambulatory Non-Arrivals. *Journal of General Internal Medicine*, *38*(10). https://doi.org/10.1007/s11606-023-08065-y

[39] Aldhoayan, M. D., & Alobaidi, R. M. (2023). The Use of Machine Learning to Predict Late Arrivals at the Adult Outpatient Department. *Cureus*. https://doi.org/10.7759/cureus.39886

[40] Alabbad, D. A., Almuhaideb, A. M., Alsunaidi, S. J., Alqudaihi, K. S., Alamoudi, F. A., Alhobaishi, M. K., Alaqeel, N. A., & Alshahrani, M. S. (2022). Machine learning model for predicting the length of stay in the intensive care unit for Covid-19 patients in the eastern province of Saudi Arabia. *Informatics in Medicine Unlocked*, *30*. https://doi.org/10.1016/j.imu.2022.100937

[41] Parthasarathy, S., Lakshminarayanan, A. R., Khan, A. A. A., Sathick, K. J., & Jayaraman, V. (2023). Detection of Health Insurance Fraud using Bayesian Optimized XGBoost. *International Journal of Safety and Security Engineering*, *13*(5). https://doi.org/10.18280/ijsse.130509

[42] Alle, S., Karthikeyan, A., Kanakan, A., Siddiqui, S., Garg, A., Mehta, P., Mishra, N., Chattopadhyay, P., Devi, P., Waghdhare, S., Tyagi, A., Tarai, B., Hazarik, P. P., Das, P., Budhiraja, S., Nangia, V., Dewan, A., Sethuraman, R., Subramanian, C., … Priyakumar, U. D. (2022). COVID-19 Risk Stratification and Mortality Prediction in Hospitalized Indian Patients: Harnessing clinical data for public health benefits. *PLoS ONE*, *17*(3 March). https://doi.org/10.1371/journal.pone.0264785

[43] Hammoud, B., Semaan, A., Elhajj, I., & Benova, L. (2022). Can machine learning models predict maternal and newborn healthcare providers' perception of safety during the COVID-19 pandemic? A cross-sectional study of a global online survey. *Human Resources for Health*, *20*(1). https://doi.org/10.1186/s12960-022-00758-5

[44] Ghosh, S., Das, J., Ghosh, S. K., & Buyya, R. (2020). CLAWER: Context-aware Cloud-Fog based Workflow Management Framework for Health Emergency Services. *Proceedings - 20th IEEE/ACM International Symposium on Cluster, Cloud and Internet Computing, CCGRID 2020*.

https://doi.org/10.1109/CCGrid49817.2020.000-5

[45] Barak-Corren, Y., Chaudhari, P., Perniciaro, J., Waltzman, M., Fine, A. M., & Reis, B. Y. (2021). Prediction across healthcare settings: a case study in predicting emergency department disposition. *Npj Digital Medicine*, *4*(1). https://doi.org/10.1038/s41746-021-00537-x

[46] Chauhan, R., Goel, A., Alankar, B., & Kaur, H. (2024). Predictive modeling and web-based tool for cervical cancer risk assessment: A comparative study of machine learning models. *MethodsX*, *12*. https://doi.org/10.1016/j.mex.2024.102653

[47] Finkelstein, J., Cui, W., Martin, T. C., & Parsons, R. (2022). Machine Learning Approaches for Early Prostate Cancer Prediction Based on Healthcare Utilization Patterns. *Studies in Health Technology and Informatics*, *289*. https://doi.org/10.3233/SHTI210860 Eledkawy, A., Hamza, T., & El-Metwally, S. (2024). Precision cancer classification using liquid biopsy and advanced machine learning techniques. *Scientific Reports*, *14*(1). https://doi.org/10.1038/s41598-024-56419-1

[48] Huang, A. A., & Huang, S. Y. (2023b). Use of machine learning to identify risk factors for insomnia. *PLoS ONE*, *18*(4 April). https://doi.org/10.1371/journal.pone.0282622

[49] Rahmani, K., Garikipati, A., Barnes, G., Hoffman, J., Calvert, J., Mao, Q., & Das, R. (2022). Early prediction of central line associated bloodstream infection using machine learning. *American Journal of Infection Control*, *50*(4). https://doi.org/10.1016/j.ajic.2021.08.017

[50] Hui, M., Ma, J., Yang, H., Gao, B., Wang, F., Wang, J., Lv, J., Zhang, L., Yang, L., & Zhao, M. (2023). ESKD Risk Prediction Model in a Multicenter Chronic Kidney Disease Cohort in China: A Derivation, Validation, and Comparison Study. *Journal of Clinical Medicine*, *12*(4). https://doi.org/10.3390/jcm12041504

[51] Gupta, A., & Singh, A. (2023). Prediction Framework on Early Urine Infection in IoT–Fog Environment Using XGBoost Ensemble Model. *Wireless Personal Communications*, *131*(2). https://doi.org/10.1007/s11277-023-10466-5

[52] Wong, G. L. H., Yuen, P. C., Ma, A. J., Chan, A. W. H., Leung, H. H. W., & Wong, V. W. S. (2021). Artificial intelligence in prediction of non-alcoholic fatty liver disease and fibrosis. *Journal of Gastroenterology and Hepatology (Australia)*, *36*(3). https://doi.org/10.1111/jgh.15385

[53] Ghandian, S., Thapa, R., Garikipati, A., Barnes, G., Green-Saxena, A., Calvert, J., Mao, Q., & Das, R. (2022). Machine learning to predict progression of non-alcoholic fatty liver to non-alcoholic steatohepatitis or fibrosis. *JGH Open*, *6*(3). https://doi.org/10.1002/jgh3.12716

[54] Pang, X., Forrest, C. B., Lê-Scherban, F., & Masino, A. J. (2021). Prediction of early childhood obesity with machine learning and electronic health record data. *International Journal of Medical Informatics*, *150*. https://doi.org/10.1016/j.ijmedinf.2021.104454

[55] Lipesa, B. A., Okango, E., Omolo, B. O., & Omondi, E. O. (2023). An application of a supervised machine learning model for predicting life expectancy. *SN Applied Sciences*, *5*(7). https://doi.org/10.1007/s42452-023-05404-w

[56] Sun, W., Kalmady, S. V., Sepehrvand, N., Salimi, A., Nademi, Y., Bainey, K., Ezekowitz, J. A., Greiner, R., Hindle, A., McAlister, F. A., Sandhu, R. K., & Kaul, P. (2023). Towards artificial intelligence-based learning health system for population-level mortality prediction using electrocardiograms. *Npj Digital Medicine*, *6*(1). https://doi.org/10.1038/s41746-023-00765-3[51]

[52] Wong, G. L. H., Yuen, P. C., Ma, A. J., Chan, A. W. H.,Leung, H. H. W., & Wong, V. W. S. (2021). Artificial intelligence in prediction of non-alcoholic fatty liver disease and fibrosis. *Journal of Gastroenterology and Hepatology (Australia)*, *36*(3). https://doi.org/10.1111/jgh.15385

[53] Yang, J. C. (2024). The prediction and analysis of heart disease using XGBoost algorithm. *Applied and Computational Engineering*, *41*(1). https://doi.org/10.54254/2755-2721/41/20230711

[54] Zhang, L., Zhou, X., & Cao, J. (2024). Establishment and validation of a heart failure risk prediction model for elderly patients after coronary rotational atherectomy based on machine learning. *PeerJ*, *12*. https://doi.org/10.7717/peerj.16867

[55] Segarra, C., Delgado-Gonzalo, R., & Schiavoni, V. (2020). MQT-Tz: Secure MQTT broker for biomedical signal processing on the edge. *Studies in Health Technology and Informatics*, *270*. https://doi.org/10.3233/SHTI200177

[56] Palkar, A., Dias, C. C., Chadaga, K., & Sampathila, N. (2024). Empowering Glioma Prognosis With Transparent Machine Learning and Interpretative Insights Using Explainable AI. *IEEE Access*, *12*. https://doi.org/10.1109/ACCESS.2024.3370238.

[57] Rocha, C. N., & Rodrigues, F. (2021). Forecasting emergency department admissions. *Journal of Intelligent Information Systems*, *56*(3). https://doi.org/10.1007/s10844-021-00638-9