



An efficient edge/cloud medical system for rapid detection of level of consciousness in emergency medicine based on explainable machine learning models

Nora El-Rashidy¹ · Ahmed Sedik^{2,3} · Ali I. Siam⁴ · Zainab H. Ali⁴

Received: 20 January 2022 / Accepted: 6 January 2023 / Published online: 15 March 2023
© The Author(s) 2023

Abstract

Emergency medicine (EM) is one of the attractive research fields in which researchers investigate their efforts to diagnose and treat unforeseen illnesses or injuries. There are many tests and observations are involved in EM. Detection of the level of consciousness is one of these observations, which can be detected using several methods. Among these methods, the automatic estimation of the Glasgow coma scale (GCS) is studied in this paper. The GCS is a medical score used to describe a patient's level of consciousness. This type of scoring system requires medical examination that may not be available with the shortage of the medical expert. Therefore, the automatic medical calculation for a patient's level of consciousness is highly needed. Artificial intelligence has been deployed in several applications and appears to have a high performance regarding providing automatic solutions. The main objective of this work is to introduce the edge/cloud system to improve the efficiency of the consciousness measurement through efficient local data processing. Moreover, an efficient machine learning (ML) model to predict the level of consciousness of a certain patient based on the patient's demographic, vital signs, and laboratory tests is proposed, as well as maintaining the explainability issue using Shapley additive explanations (SHAP) that provides natural language explanation in a form that helps the medical expert to understand the final prediction. The developed ML model is validated using vital signs and laboratory tests extracted from the MIMIC III dataset, and it achieves superior performance (mean absolute error (MAE) = 0.269, mean square error (MSE) = 0.625, R^2 score = 0.964). The resulting model is accurate, medically intuitive, and trustworthy.

Keywords Emergency medicine · Explainability · GCS · Machine learning · Edge/cloud medical system

1 Introduction

Emergency medicine (EM) is a rapid-growing specialty which is critical and important for the society. The patients are received in urgent cases in which rapid tests and evaluation of vital signs are very important to obtain an accurate diagnosis and make decisions. Therefore, this field attracts researchers to investigate solutions in it. Artificial intelligence is strongly involved in these investigations including machine learning (ML) and deep learning (DL) [1]. One of the important medical issues is the detection of level of consciousness which is considered as one of the important observations of the patient. For this reason, the care givers should rapidly handle the patients in this case to survive them [2].

Level of consciousness can be obtained by evaluation of Glasgow coma scale (GCS) using several methods such as

✉ Nora El-Rashidy
Noura.alrashidy@ai.kfs.edu.eg

¹ Department of Machine Learning and Information Retrieval, Faculty of Artificial Intelligence, Kafrelsheikh University, Kafr El-Sheikh, Egypt

² Smart Systems Engineering Laboratory, College of Engineering, Prince Sultan University, 11586 Riyadh, Saudi Arabia

³ Department of the Robotics and Intelligent Machines, Faculty of Artificial Intelligence, Kafrelsheikh University, Kafr El-Sheikh, Egypt

⁴ Department of Embedded Network Systems Technology, Faculty of Artificial Intelligence, Kafrelsheikh University, Kafr El-Sheikh, Egypt

electroencephalography (EEG) and vital signs [3, 4]. Traditionally, the consciousness of a certain patient can be determined based on his eye opening, verbal and motor responses which are the factors of GCS. It is a dominant method which is scaled from 0 to 15. This method needs medical examination from the medical expert which is not available all the time. Therefore, there is a need to investigate an automatic method for patients based on vital signs such as hypotensive, heart-rate-disordered and hyperthermal patients and laboratory tests such as albumin and hemoglobin in addition to collecting some medical records such as EEG and EPG signals [5].

The field of artificial intelligence is involved in Big Data and Analytics [6], Cloud/Edge Computing-based Big Computing [7] and the Internet of Things (IoT)/Cyber-Physical Systems (CPS) [8, 9] applications. These applications dominate industry and research for the development of various smart-world systems [10, 11]. Large, complicated datasets may now be approximated and reduced into extremely accurate predictions and transformative output using artificial intelligence, making human-centered smart systems much easier to implement [12–14]. Machine learning techniques can be used to any types of data such as visual, auditory, numerical, text or some combination [15]. Therefore, engineers can build their edge-based platforms based on machine learning techniques due to its high performance and low time consumption. Furthermore, the deployment of machine learning algorithms is involved to providing a security environment to ensure the privacy of the data of the patients which is transmitted through the network [16, 17]. Therefore, this study comprises the issue of security and privacy and its importance in the edge communication system [18].

The utilization of ML in medicine has witnessed an explosion in numerous medical applications [19–22], including automated diagnosis, classification of disease severity, development of new therapies [23], analysis of medical record [24] and improving the quality of medical data [25]. The use of ML methods in automated diagnosis has bifurcated in diverse disease types, including corona virus [26–30], kidney disease [31, 32], heart disease [33, 34], cancer [35, 36], diabetes and retinopathy [37, 38], skin lesion [39] and other diseases [40–42].

- Motivation and contributions

The GCS has been extensively used to objectively describe the extent of impaired consciousness within all types of acute medical and trauma patients. GCS assesses patients based on the following aspects: (i) eye-opening verbal, (ii) motor and (iii) verbal responses. These scoring factors are not obtained automatically, which may cause less detailed description. In addition, there is a need to

overcome the difficulty with early detection and diagnosis. Therefore, the objective is to provide a numerical method to evaluate the GCS accurately has become indispensable. The main objective of this work is to introduce a machine learning-based system that is performed through IoT and edge/cloud system to enable automatically measuring the level of consciousness. The proposed system consists of three main phases: (i) vital signs acquisition, (ii) Fog-Assisted Consciousness Management (FACM) and (iii) cloud server and clinical service delivery model. In addition, the proposed system is carried out on the MIMIC dataset. Furthermore, the proposed model maintains the interoperability issue by providing natural language explanation for the developed decisions, in order to provide answers for the medical straightforward inquiries. Contributions of this paper can be illustrated in the following points:

- (1) Investigate a new method for GCS evaluation for automatic estimation of the resulting score.
- (2) Build an internet of medical things (IoMT) system through edge/cloud technology.
- (3) Deploy machine learning techniques for detection of level of consciousness.
- (4) Maintain the interoperability of the ML model in order to provide explanation to the outcome of the model.
- (5) Compare the deployed techniques to obtain an optimal one in terms of evaluation metrics.
- (6) Recommend an optimal system and discuss the limitation of its application.
- (7) Evaluate the effectiveness of adopting the fog technology on the proposed system.

This paper is organized as follows: In Sect. 2, background and related work is reviewed. In Sect. 3, an edge/cloud system for consciousness detection is proposed. In Sect. 4, experimental results are detailed. Section 5 presents a discussion for the results and comparison with the works in the literature. In Sect. 6, the paper is concluded.

2 Background and related work

This work proposes an edge/cloud medical system whose objective is to estimate the level of consciousness. For this purpose, we deployed a set of ML techniques to predict the value of GCS automatically. This section discusses the works in the literature which are relevant to the proposed system. Firstly, we discuss the GCS. GCS is one of the most utilized scores for responsiveness assessment of inpatients. It was introduced in 1974 to standardize the clinical assessment of level of consciousness in patients

with head injuries [43]. GCS is an effective means to compare responses of patients in different coma states and to compare effectiveness of treatments [44]. GCS is a realization of three components: motor response, verbal response and eye opening. The scale originally consisted of fourteen points, four for eye opening and five for each motor response and verbal response. A sixth point was added 2 years later for motor response [45]. The GCS and its score points are shown in Table 1.

The manual calculation of GCS score involves summing the scores corresponding to the best response for each individual behavior. Hence, the total score has values between 3, being in deep coma or death, and 15, being fully alert.

To the best of our knowledge, there are no relevant contributions in the literature that exploit laboratory tests and vital signals to automatically provide a numeric estimate of GCS level using machine learning techniques. Consequently, this study is the first to conduct a similar method and results. However, early studies have shown the feasibility of using GCS levels to automatically determine the functional state of the autonomic nervous system (ANS) in coma patients. However, these studies classify the consciousness level into subgroups: two subgroups (with GCS from 3 to 5 and from 6 to 8) [46], or three subgroups (low, mild and high consciousness) [4].

Estévez et al. [46] presented an approach to classify coma patients into two subgroups according to their GCS based on the heart rate variability (HRV). The experiments were conducted on 47 patients in coma. All patients were in ICU and mechanically ventilated. In this approach, ECG signals have been extracted, resampled into 1000 Hz and then processed using Hilbert–Huang transform (HHT) to

extract a number of key spectral features. A logistic regression model was implemented to classify the consciousness level of patients into two categories: deep and mild coma, based on their HRV. They reported that their model achieved an overall efficiency of 95.74%.

Latifoğlu et al. [4] proposed an approach for automatic evaluation of the state of consciousness of coma patients based on EEG signals. The state of consciousness is classified into either low, mild or high based on GCS levels. They obtained EEG signals from 34 coma patients in ICU. Features are extracted using power spectral density (PSD) method. The authors adopted various machine learning classifiers to classify the consciousness level, and they obtained an accuracy of 92.5%.

Furthermore, machine learning algorithms are widely adopted in health care, relying on medical data to predict or classify various health states [47, 48]. Also, ML and DL have wide applications in emergency medicine [49–53]. An ML-based model was presented in [54] to predict the outcome of patients after traumatic brain injury (TBI). GCS level, besides the other thirteen parameters, was involved in predicting the patients' outcome. Authors have conducted a performance comparison of different nine ML algorithms and reported that the random forest algorithm had achieved the best performance in outcome prediction with an accuracy of 91.7%. They also concluded that GCS score, besides age, fibrin/fibrinogen degradation products and glucose are the most important factors for outcome prediction. Tsiklidis et al. [55] implemented an ML model based on a gradient boosting classifier to predict the mortality rate of trauma patients at admission. They relied on the GCS and other seven health parameters to train the model. The accuracy of the model was 92.4%. The authors remarked that GCS, age and systolic blood pressure had the highest impact on the final decision of the model.

In addition, Hall et al. [56] implemented a decision tree-based model to identify patients with a potentially modifiable outcome after intracerebral hemorrhage (ICH). They demonstrated that the GCS score is one of the most important predictors to identify the patient outcome. Similar results were concluded in [57] and [58] as the GCS score is the most significant variable in predicting the outcome and mortality of TBI, and subarachnoid hemorrhage (SAH) patients, respectively, using various machine learning models [59, 60].

The term “fog computing” refers to a paradigm that brings cloud computing and its associated services to the edge of a network. In this manner, various issues that are inherently associated with cloud computing, such as latency, lack of mobility support and lack of location awareness, are solved [61]. Fog computing and cloud computing have common and distinct features [62, 63]; Table 2 lists a comparison between the fog and cloud

Table 1 Glasgow coma scale

| Behavior | Response | Score |
|-----------------|----------------------------|-------|
| Eye opening | Spontaneous | 4 |
| | To speech | 3 |
| | To pain | 2 |
| | No response | 1 |
| Verbal response | Oriented | 5 |
| | Confused conversation | 4 |
| | Inappropriate words | 3 |
| | Incomprehensible sounds | 2 |
| | No response | 1 |
| Motor response | Obeys commands | 6 |
| | Localizes pain | 5 |
| | Flex to withdraw from pain | 4 |
| | Abnormal flexion | 3 |
| | Abnormal extension | 2 |
| | No response | 1 |

Table 2 A comparison between the old challenges that caused by cloud computing and the added updates using fog computing and SDN

| Parameter | Cloud computing | Fog & SDN |
|--|-------------------|----------------------------|
| Processing operation [62, 67, 68] | Very high | Moderate |
| Data processing [67, 68, 69] | At cloud server | Locally—at fog and SDN |
| Transmission delay [62, 70, 68] | High | Very low |
| Location-awareness [67, 68, 70] | No | High |
| System reliability [67, 68, 69, 71] | Support | Support |
| System scalability [62, 67, 68] | Support | Support |
| Geographical distribution [62, 67, 68] | No | High |
| Real-time interactions [68] | Not fully support | Support |
| Power consumption [68] | High | Low |
| Ubiquitous services [68] | Support | Support |
| Management model [68] | Centralized | Decentralized |
| Decision-making [68] | Remotely | Locally |
| Security [68] | It's hard | Easy to apply and maintain |

platforms in terms of various technical aspects. With the existence of IoT devices, the requirements for high bandwidth, security and low-latency applications are raised [64]. Therefore, fog computing is fitted here to provide such requirements for IoT networks. Recently, ML has been widely adopted with fog computing to enhance its services. Abdulkareem et al. [61] investigated the different roles of ML in the fog computing paradigm and provided diverse improvements in ML techniques associated with fog computing services, such as security, accuracy and resource management. Kishor et al. [65] presented an ML-based fog computing approach to minimize latency in healthcare applications. They implemented a multimedia data segregation scheme in fog computing to reduce the total latency resulting from the data transmission, computation and network delay. They employed the random forest model to calculate the total latency. The reported results reveal that 95% reduction in latency is achieved compared to another pre-existing model. Khater et al. [66] proposed a lightweight intrusion detection system (IDS) for fog computing using a multilayer perceptron (MLP) model. They evaluated the developed system against two benchmark datasets: Australian Defense Force Academy Linux Dataset (ADFA-LD) and Australian Defense Force Academy Windows Dataset (ADFA-WD), which contain exploits and attacks on various applications. The developed system is implemented using a single hidden layer, and it achieved a 94% accuracy in ADFA-LD and 74% accuracy in ADFA-WD.

3 Proposed edge/cloud system

This work proposes an edge/cloud system for consciousness detection. The deployed scenario is based on fog-assigned consciousness managing system. The main

objective of this proposed system is to enable the scoring system based on vital signs and laboratory tests such as blood pressure, heart rate, respiratory rate and oxygen flow rate to determine the consciousness level of the patient. A theoretical approach is first introduced. Recent research in describing the subsequent level of consciousness has been concerned with determining GCS, which includes several functions such as eye-opening, verbal response and motor response. Then, the collected information is recorded manually by a therapist in a local system computer-based or paper-based. In some cases, this method used may ignore significant factors such as alcohol intoxication, low blood oxygen and drug use histories. In addition, it may suffer a delay time in decision making. All these issues lead to an inappropriate score that can alter and negatively affect a patient's level of consciousness. Therefore, introducing a dynamic data exchange environment with a high ability to deal transparently with a large scale of vital functions should be required.

The structure of the proposed edge/cloud system, as shown in Fig. 1, consists of three main phases: (i) vital signs acquisition, (ii) Fog-Assisted Consciousness Management (FACM), and (iii) cloud server and clinical service delivery model. The edge/cloud system can be implemented in a wide range of intelligent healthcare sectors that are characterized by a massive infusion of data and a high need for careful and rapid decision-making, such as emergency departments and intensive care units.

3.1 Vital signs acquisition

The operation in this face divides into two directions are: (i) collecting information of vital signs from the real sensors that directly connected to the patient; and (ii) getting information from the medical report prepared by specialists. Lately, IoT sensor devices play an essential role in

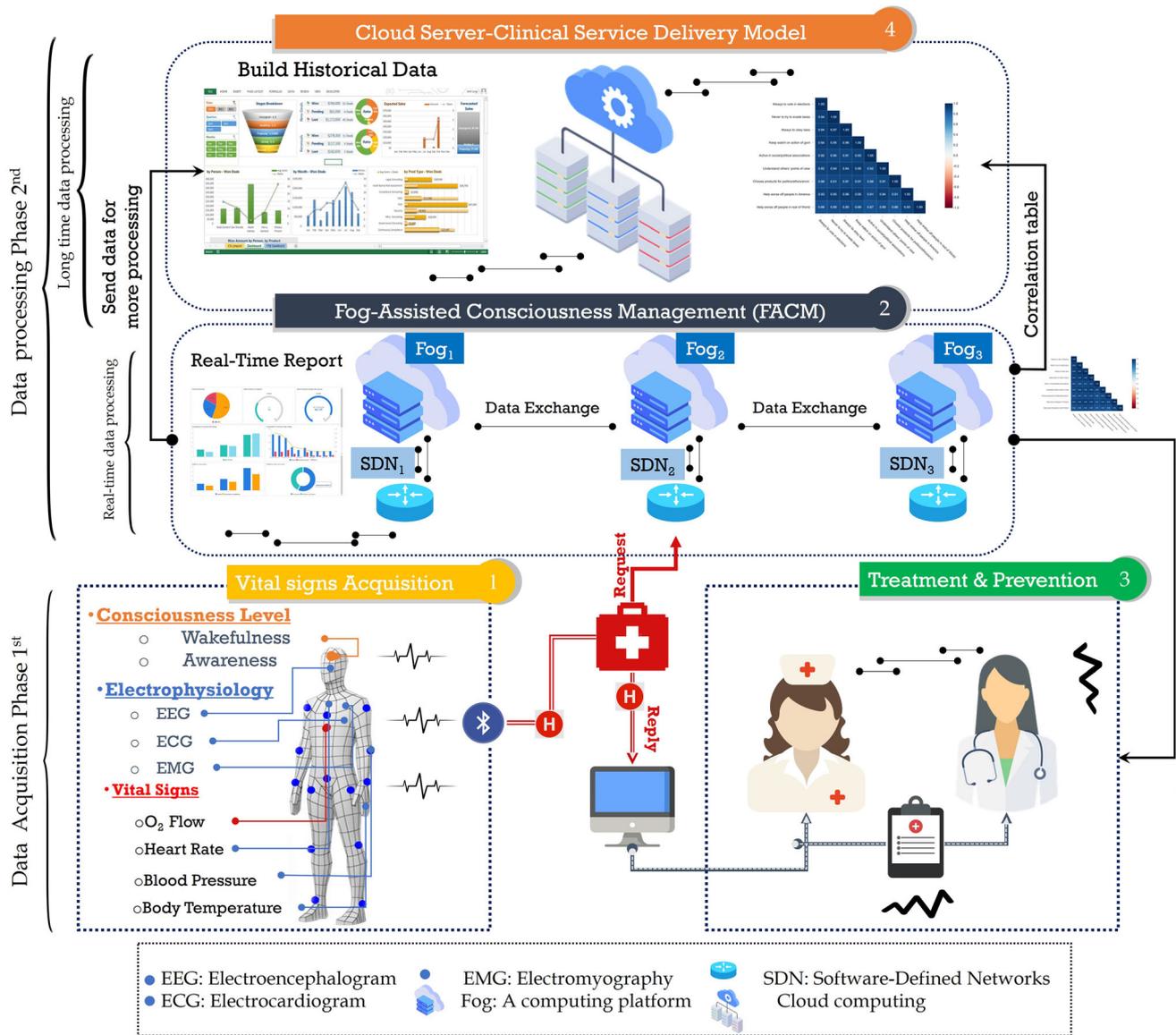


Fig. 1 Conceptual design of the edge/cloud system

medical construction in which it provides interactive real-time network connection to the users and medical devices through various communication technologies. Wearable devices are a part of IoT devices that allow sensor devices to collect real information from patients about their vital signs from everywhere at any time; this technique is called ubiquitous technology. The cope of these collected vital signs will be through the software-defined network (SDN) and fog computing installed in FACM layer. In the traditional architecture proposed in several types of research, the data will transfer to the cloud server; this technique affects not only network bandwidth but also response time. Table 2 describes old challenges that caused by cloud computing and the added updates using fog computing and SDN.

3.2 Fog-Assisted Consciousness Management (FACM)

This second phase includes several operations related to managing the process of data transmission from and to the cloud server and end-user within the treatment and prevention phase. The integration between SDN and fog technology initially enables data communication systems to be more dynamic, secure, and reliable. The security and data reliability achieved by the SDN will establish a private channel with the fog server via its controller in OpenFlow to ensure the level of data privacy. Inside this channel, the fog server applies an access control policy that is predefined by the fog setting. For more data protection, the fog server uses the integrity check process via adding octets/

bytes with every data packet sent [62]. This technique is a common Internet technique that is used to notify the fog server about any change in the bytes of data. The operations in this phase are listed as follows:

- Providing a decentralized data transmission strategy through distributed SDN nodes integrated with other deployed fog nodes. This strategy supports local data processing and avoids unnecessary data traveling to the cloud server, thus reducing network bandwidth usage.
- Establishing a secure channel for boosting data protection during the operation of data exchange between SDN and IoT sensors and SDN and fog nodes.
- According to vital signs that sent from the lower layer, the fog computing starts to build a predictive model of consciousness level based on an unsupervised learning technique called the principal component analysis (PCA).

As shown in Fig. 2, the vital signs are sent from sensors in an event shape. Each event includes sensor_ID, patient_ID, patient’s location, flow entries that tell the SDN what to do with an incoming packet, as well as the patient’s vital signs. Using OpenFlow (OF) protocol, the events will convert to flow entries for feeding SDN. This technique increases SDN capability in network monitoring and management.

From the network point of view, achieving better network connectivity relies on avoiding communication range violation during the process of data transmission [62]. Therefore, the available communication range for each SDN is determining according to the available predefined range assigned by the fog node. For instance, as seen in

Fig. 2, the communication range in coverage area I is determined by Fog₁ and thus the SDN₁ can only direct connected with sensors deployed in this area. Each event sent from this area has been addressed by a labeled information about sensors and patients connected in this area only. In addition, SDN1 has equipped with cache memory for saving requests and adding security rules for them. Table 3 illustrates a sample of data stored in the SDN₁ caching table.

The next operation in SDN is between SDN and Fog node. Each SDN generates its own private and public key for adding them to the transferred data to the fog node. For example, a single fog node can manage multiple SDNs in the same area or even different areas to provide a large scale of data processing and analysis and overcome high response time. Herein, the data in the fog node have become accessible from multiple users, and it needs to protect.

The transferred data from the SDN to the fog node will be addressed by some flow entries and SDN_ID for labelling. Moreover, the SDN adds a public key with each packet for protection. Only authorized fog node has the private key to decode this packet. It is worth to point out that only data related to the packet remain confidential, while vital signs are sent without encryption for ensuring the acceleration of fog performance. Table 4 discusses the steps of data transmission between SDN and fog node.

According to the proposed edge/cloud system, the deployment of fog nodes in the 2nd layer between the cloud server and IoT sensor devices plays a significant role in carrying out many fundamental operations on the received data from the SDN before passing it to the cloud

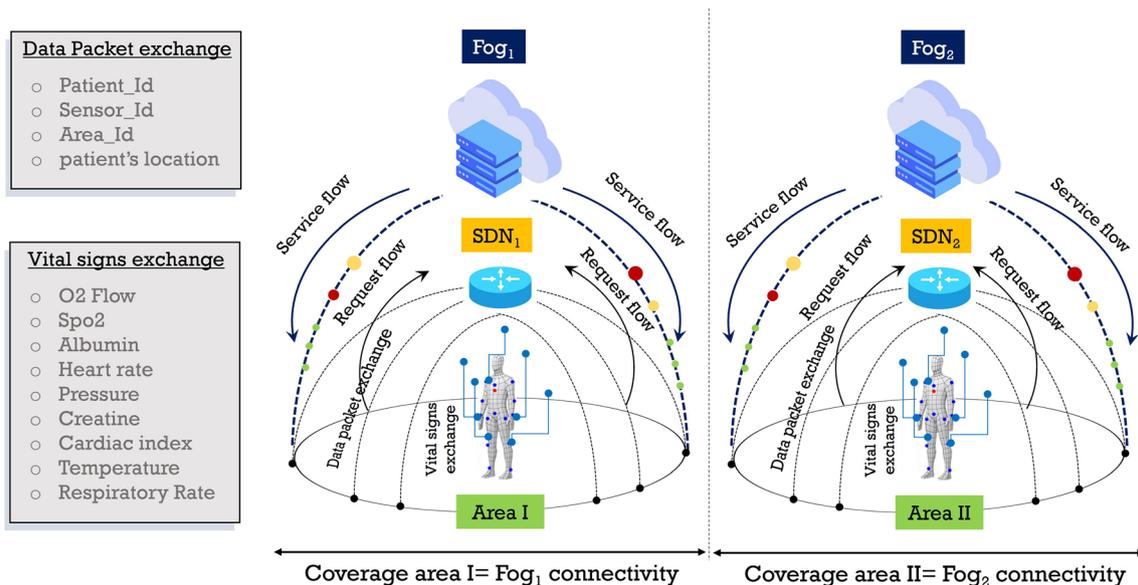


Fig. 2 The process of data and vital signs exchange from patient to SDN and from SDN to Fog

Table 3 A sample of data stored in the SDN₁ caching table

| Header of data packet | | | | Vital signs | | | | | | | | |
|-----------------------|-----------|-------------|------------|-------------|----------|---------------------|---------------------|--------|--------|--------|-------|------|
| Area_ID | Sen_ID | Flow entry# | Patient_ID | loc | Pub. Key | Priv. Key | O ₂ Flow | HR | PR | RR | BUN | |
| | | | | X | Y | | | | | | | |
| I-2 | Srl 186 | 123 | PHU-14708 | 52 | 12 | AABCID234781cy2E48v | EDOLWIFI2eMYdf5421T | 80–100 | 60–100 | 70/120 | 12–16 | 6–24 |

server layer. One of these operations is related to reducing the consumption of network bandwidth [1, 2]. Fog node performs the processing on received data locally. Therefore, the data does not need to travel to the cloud server for processing. On the other hand, the fog nodes use a tree-based approach for building an intelligent model calculating the scoring conscious based on the vital signs received by SDN. Figure 3 depicts the main steps for building an intelligent model on the fog node.

The fog node takes the data preparation file from the previous phase in the 1st step. In the 2nd step, the fog node applies PCA for selecting the most effective features among a huge number of vital signs collected from the patient. This operation decreases the size of data that affects the system performance. In the 3rd step, the fog node builds the consciousness prediction model based on a random forest (RF) algorithm. The use of an RF to predict the GCS using vital signs has returned to several reasons that are: (i) RF gives the ability to measure feature correlation for all features using the Gini index that indicates the impact of each feature in the model [72], (ii) RF compromising the explainability and accuracy issues. Generally, models that have a good performance in terms of classification accuracy as SVM and LDA are not able to provide a clear explanation about the output decision [73], whereas the most tree-based algorithms are very good explainability, but may not be the best algorithm in terms of performance [74], and (iii) RF is a tree-based algorithm that utilized several trees and then combined the final decision using a majority voting algorithm. Information gain is used to split points in each tree. As a result, outliers are ignored by most trees that make RF a more stable algorithm [75]. In the 4th step, the GCS with correlation table will be stored at the cloud server to build historical data about these patients.

3.3 Cloud server and clinical service delivery model

In the healthcare industry, cloud computing plays a pivotal role in supporting the shift of conventional storage to the digitalization of healthcare data. The revised vital data collected from the FACM phase will be travelled to cloud computing for saving, computing and analyzing. This technique affects network bandwidth usage and accelerates cloud decisions. With cloud computing, a historical healthcare database will build to wrap up patients’ data flowing from FACM. This database intends to create data linkages throughout the healthcare systems, irrespective of where the data originate or are stored. Moreover, cloud-based data analysis can prepare more personalized care reports for patients on an individual level, and thus several healthcare-related functions will be improved in terms of

GCS evaluation and detection of the level of consciousness automatically.

4 Experimental results

4.1 Experiment 1

This section is going to measure the performance of the edge/cloud medical system-based fog technology compared to traditional monitoring system without fog technology. The NS2.35 is used to model the wireless sensor network and fog sensor with varying network range from 100 to 400 sensor, Table 5 shows all network settings. The sensor nodes connect to others via IEEE 802.11p/WAVE. Table 6 depicts the configuration of the IEEE 802.11p interface described as it is in [76], and fog nodes were reconfigured based on the reference guide called “*Fog hierarchical deployment model*” from OpenFog Reference Architecture [77].

Reliable communication has typically relied upon the quality of the packet transmission process. To this end, the performance of the proposed edge/cloud medical system in achieving high reliability was tested. In Fig. 4a, the x -axis denotes the network size, and the y -axis denotes the total speed of sensors, while in Fig. 4b, the x -axis denotes the total number of data packets, and the y -axis denotes the system throughput (kb/s). As observed shown in Fig. 4a b, the proposed system-based fog computing outperforms the traditional monitoring system-based cloud computing because the fog computing technology can increase the number of successful packets sent without affecting the network bandwidth, as is evident from Fig. 4b. Fog computing also provides the process of data processing locally,

and thus, there is no need to transfer data via the Internet; only the data that need more processing and analysis will send to the cloud server.

4.2 Experiment 2

In this experiment, we exploit the proposed edge/cloud system to improve the efficiency of the consciousness measurement through efficient local data processing. Moreover, an efficient machine learning (ML) model to predict the level of consciousness of a certain patient based on the patient’s demographic, vital signs and laboratory tests is proposed, as well as maintaining the explainability issue using Shapley additive explanations (SHAP) that provides natural language explanation in a form that helps the medical expert to understand the final prediction.

4.2.1 Data collection

Medical Information Mart for Intensive Care III (MIMIC-III) is a publicly available dataset for intensive care units (ICU). It comprises electronic health record (EHR) data extracted from the bedside monitors inside ICU units of the Israel Medical center in Boston, USA [78]. MIMIC III was approved and maintained by the Massachusetts Institute of Technology (MIT), and it is freely available on PhysioNet [79]. MIMIC III includes data for 46,520 different patients and 58,976 different admissions, gathered between 2001 and 2012 [80, 81]. Each patient is associated with minute-by-minute vital signs measurements and laboratory tests. The median age of adult patients is 65.8 years, and 55.9% of patients are males. The sampling interval of records ranges from few seconds to hours, according to the acquired physiological measurements.

Table 4 The steps of data transmission between SDN and fog node

| Step | Description |
|------|---|
| 1 | Each SDN stores sensing data received from its coverage area into a flow table. This table contains two types of the information: (i) information about forwarded data packet and (ii) information about vital signs |
| 2 | Each SDN generates its own public key and adds it to every data packet as it is shown in Table 3 |
| 3 | The traffics between SDN and Fog node in the same area are occurred in a secure channel using SSH protocol |
| 4 | SDN uses a public key to encrypt only the header of the packet, and therefore the data packet has become anonymous or without personally identifiable information. The vital signs only will send clearly |
| 5 | The traffic sent from SDN is directly connected to the fog node located in the same area. This process helps to mitigate communication overhead |
| 6 | Each fog node provides an intelligent model that can predict the consciousness level according to vital signs sent from SDN. Only fog node that has the private key can decrypt messages sent from a certain SDN |
| 7 | Fog in a certain area can send the data packet to another fog node in another area for processing. This technique helps to achieve network load balancing so that if the fog node is busy the message will immediately redirect to another fog node. In our case, the fog node only can communicate with other fog nodes in different areas, while each SDN can establish the communication between its fog node located in the same area |

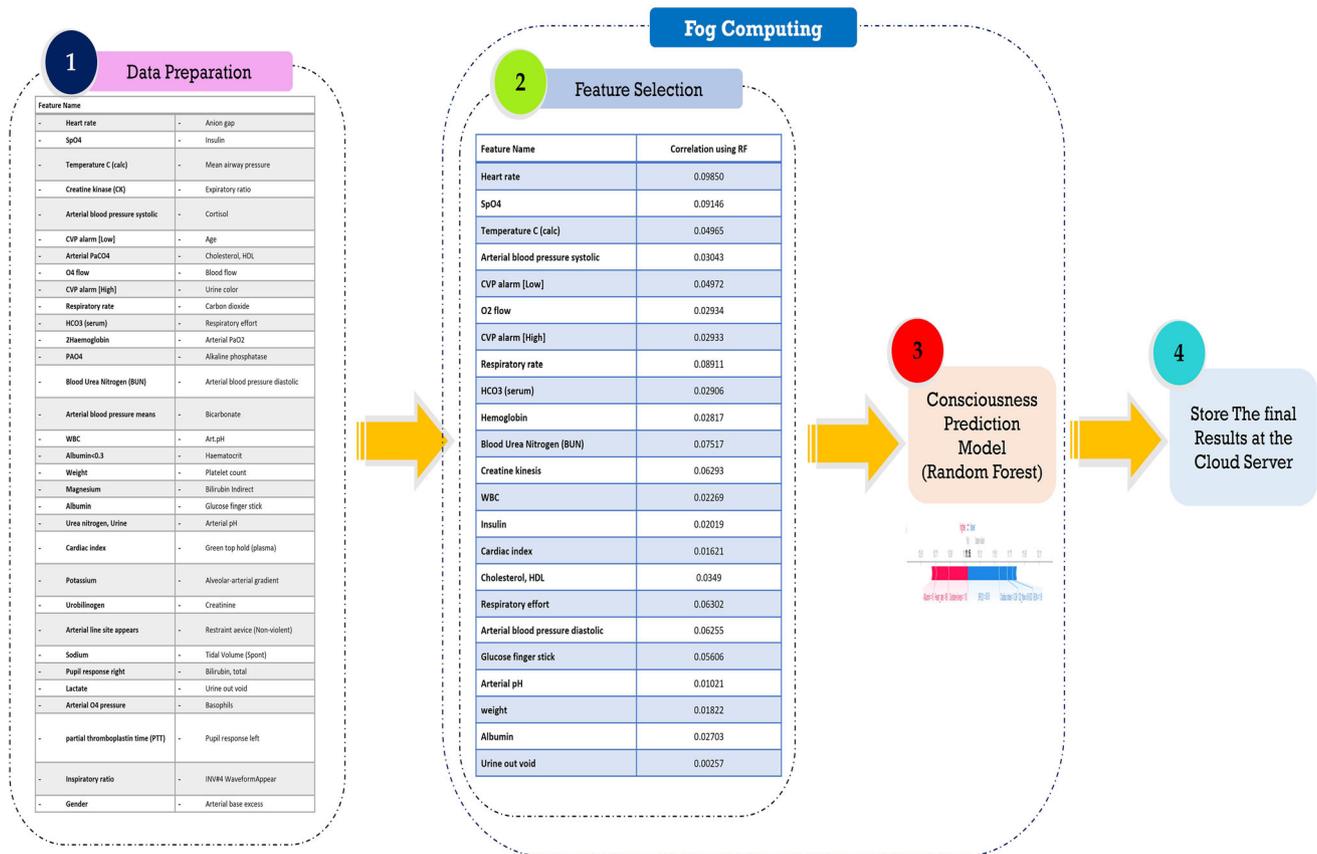


Fig. 3 The main steps for calculating the level of consciousness

In the current study, we extracted 10,349 records from the entire MIMIC dataset and are used to train/test the model to predict the level of consciousness. The data were randomly split into 80/20 ratio for training and testing the ML models, respectively. Appendix A gives some information (normal range, unit of measurement) of physiological measurements and laboratory tests that are adopted in the current study.

4.2.2 Data pre-processing

These steps aim to enhance the quality of the chosen dataset. MIMIC III dataset includes several challenges, including missing values, outliers, etc. This may occur due to sensor or transmission failure, error in saving data, etc. Training a model in such noisy and incomplete data is considered the main reason for a model with poor performance. The following subsections discuss the steps taken to handle such data challenges.

Table 5 Network setting

| Parameter | Value |
|-----------------|--------------|
| Sensor range | 100–400 |
| Simulation time | 1500 S |
| MAC protocol | IEEE 802.11p |
| Channel type | WirelessPhy |
| Energy model | Battery |
| Antenna model | OmniAntenna |
| Packet size | 512 bytes |
| Traffic source | CBR |

Table 6 IEEE80211.p interface configuration [76]

| Parameter | Value |
|----------------------|------------|
| Channel | 175 |
| Bandwidth | 20 MHz |
| Frequency | 5.875 GHz |
| Antenna gain | 2dBi |
| Setup TxPower | 23/18 dBm |
| Receiver sensitivity | – 95.2 dBm |

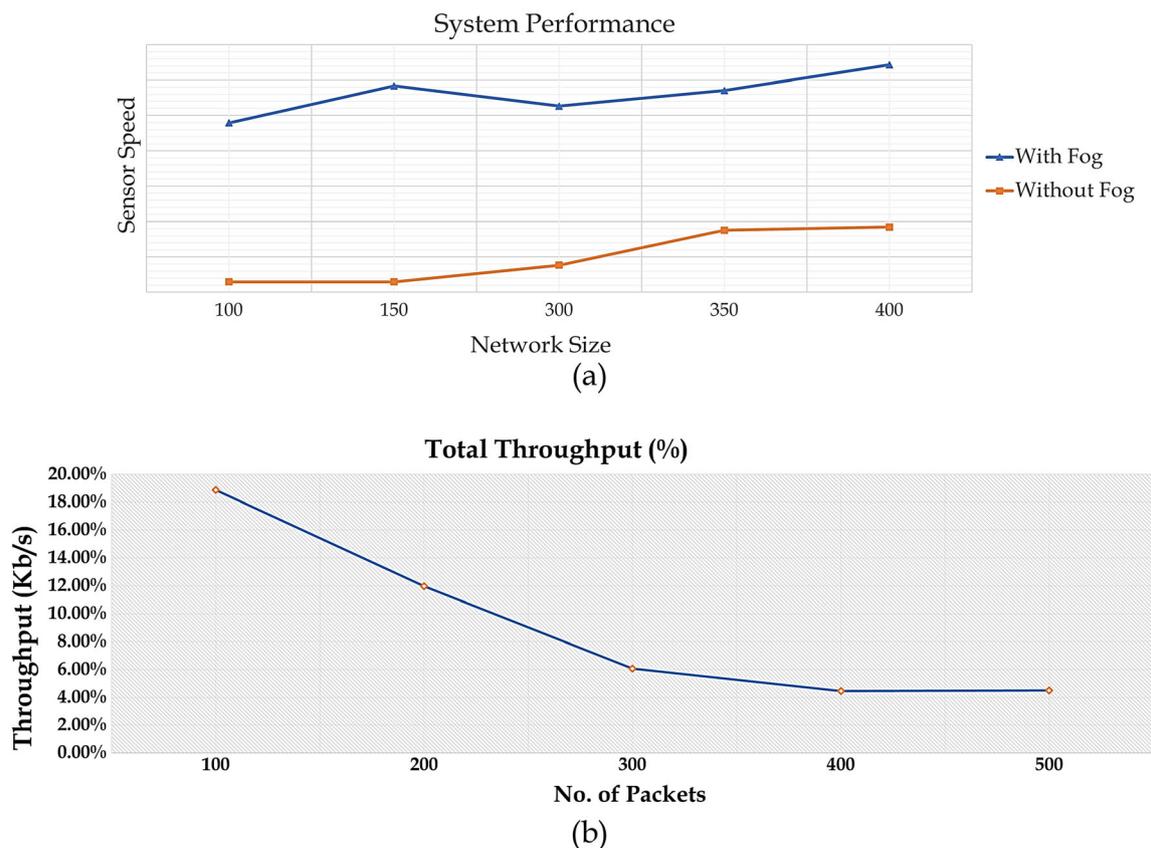


Fig. 4 The performance of edge/cloud medical system-based on fog computing versus the traditional monitoring system-based on cloud computing is shown in figure (a), and the overall system throughput is shown in Figure (b)

4.2.3 Irregular time interval

In MIMIC III, vital signs are measured at irregular time. Some of them measured every couple of minutes and other measured every few seconds [82, 83]. Unless most ML techniques are not prepared to deal with time series data, some of them could handle it when sampled with the same interval. To solve this problem, we aggregated patient's vital signs observations to provide a single record every 1 h by taking the average of all measurements over that hour. As a result, each record includes consistent values.

4.2.4 Removing outliers

Outliers are values that are too far from the normal range [75, 84]. Normal range specified according to medical expert opinions. The outliers are removed; then, expectation maximization technique is used to impute them [85].

4.2.5 Data imputation

Medical data usually include missing values. This returns to several reasons, including sensor failure, recording data at different time intervals, etc. The simple way to handle

missing values is to remove them. However, this way may lead to losing significant information. Therefore, several algorithms have been developed to impute missing values based on the other records such as hot-deck encoding [86], and expectation maximization [85]. In the preprocessing stage of MIMIC III, large proportion of data (40–55%) in important features are lost, but we could not eliminate them due to their importance in the prediction process. Considering this issue, we decide to choose cases that have at least 2 values in each measurement, then applying expectation maximization [85, 87] to impute other missing values.

4.3 Feature extraction

In this step, we mainly depend on medical expert opinions in specifying the most important features that could contribute to predict GCS and assure the ability of predicting GCS from vital signs. As mentioned before, GCS is a medical score that used to specify the consciousness of the patient through three main measures: verbal response, eye opening and motor response [88, 89]. These measurements are highly correlated with changing in patient's vital signs. For example, low blood pressure level may lead to

Table 7 ML algorithms hyperparameters

| Algorithm | Coefficient |
|-----------|--|
| LR | fit_intercept = True, normalize = True, copy_X = True, n_jobs = -1 |
| SVR | C = 1.0, epsilon = 0.1, kernel = 'rbf' |
| DT | criterion = 'mse', splitter = 'best', max_depth = 20, presort = False |
| k-NN | n_neighbors = 7, algorithm = 'auto', leaf_size = 30, p = 2, metric = 'minkowski' |
| Ridge | alpha = 1.0, fit_intercept = True, normalize = False, copy_X = True, tol = 0.001, solver = 'auto', random_state = 33 |
| RF | n_estimators = 100, max_depth = 16, random_state = 33 |
| GBR | n_estimators = 100, max_depth = 16, learning_rate = 1.5, random_state = 33 |

hypotensive; therefore, patient will not be able to respond correctly through verbal or motor response [90]. This is because decreasing blood pressure will result in generating metabolites that cause problem in circulation and tissue functions as well. The same for heart rate, the rapid or slow heart rate will mainly affect the cardiac functions that may cause cardiac arrest or blockage and subsequently affect the circulation. On the other hand, when the heart rate exceeds normal range, it increases the probability of tachyarrhythmia that may be reflected in fibrillation or atrial flutter [91, 92]. Temperature also affects the response of the patients; if the patient has hyperthermia (Temp > 40 C), the temperature autoregulation centers in the brain will be affected [93, 94]. If patient has hypothermia (Temp < = 35 C), this will affect patient response.

Deceasing in O₂ saturation may also lead to cardiac arrest or lactic acid accumulation [95]. Low hemoglobin level may also consider a risk indicator, especially with kidney diseases patients, while high level may lead to strokes, clots and heart attacks [91, 96, 97]. There are other features such as PCO₂, HCO₃, PO₂. These features used to specify the percentage of carbon and oxygen in the blood, indicate the level of blood PH and patient’s acid–base balance [98]. All of these are critical situations, affect

patient response and may lead to sudden death [99, 100]. Appendix A shows the feature names, Id, normal range and unit of measurement (UoM).

4.4 Results and discussion

The simulation results are carried out on the collected data from the patients using a server with NVIDIA GPU, Intel Core i7 CPU and 32 GB RAM. We used this facility to make sure that the proposed framework can be commonly used in the real application. This work proposes several ML techniques, including Linear Regression (LR), Support Vector Machine (SVM), Decision Tree (DT), k-Nearest Neighbor (*k*-NN) and Ridge Regression (RG). In addition, we deployed some ensemble ML methods, including Random Forest (RF) and Gradient Boosting Regression (GBR). Table 7 illustrates the hyperparameters of each one of the proposed models. These parameters are selected based on several iterations using grid search algorithm in terms of the optimal performance.

4.4.1 Evaluation metrics

The following metrics are used to evaluate the performance of the proposed model, which are Mean Absolute Error (MAE), Mean Square Error (MSE), Median Absolute Error (MedAE) and R² score. These metrics are computed as follows:

$$MAE = \frac{1}{n} \sum_{i=1}^n |y_i - \hat{y}_i| \tag{1}$$

$$MSE = \frac{1}{n} \sum_{i=1}^n (y_i - \hat{y}_i)^2 \tag{2}$$

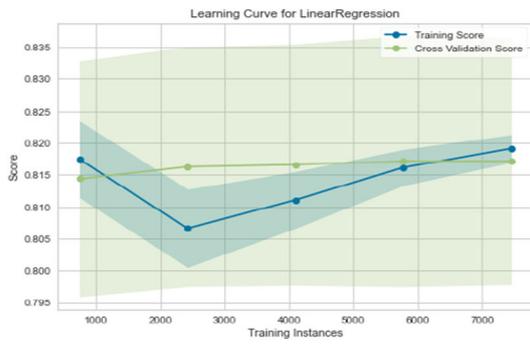
$$MedAE = \text{median}(|y_1 - \hat{y}_1|, |y_2 - \hat{y}_2|, \dots, |y_n - \hat{y}_n|) \tag{3}$$

$$R^2 = 1 - \frac{\sum_{i=1}^n (y_i - \hat{y}_i)^2}{\sum_{i=1}^n (y_i - \bar{y})^2} \tag{4}$$

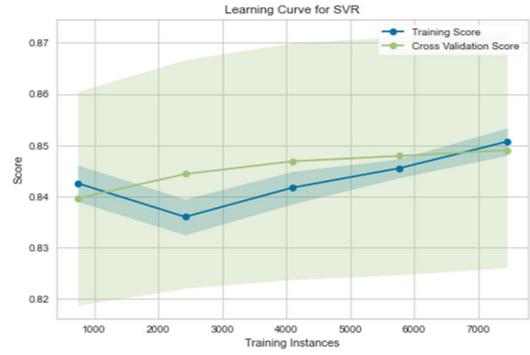
where y_i is the *i*th true value. \hat{y}_i is the corresponding predicted value. *n* is the number of observations. \bar{y} is the mean of the true values.

Table 8 Results of ML models without feature selection

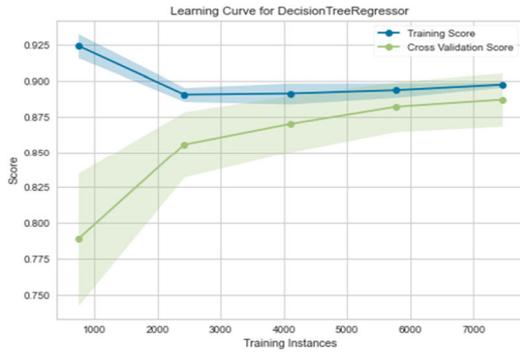
| Algorithm | Training score | Testing score | MAE | MSE | MedAE | R ² score |
|-----------|----------------|---------------|--------------|---------------|---------------|----------------------|
| LR | 0.819 | 0.816 | 0.701 | 1.995 | 0.381 | 0.816 |
| SVM | 0.850 | 0.849 | 0.3926 | 1.642 | 0.08591 | 0.8491 |
| DT | 0.897 | 0.896 | 0.503 | 1.168 | 0.348 | 0.8926 |
| KNN | 0.891 | 0.835 | 0.570 | 1.795 | 0.119 | 0.835 |
| Ridge | 0.819 | 0.816 | 0.711 | 1.993 | 0.392 | 0.816 |
| (GBR) | 0.885 | 0.880 | 0.572 | 1.3031 | 0.2578 | 0.880 |
| RF | 0.929 | 0.922 | 0.406 | 0.8385 | 0.2839 | 0.9229 |



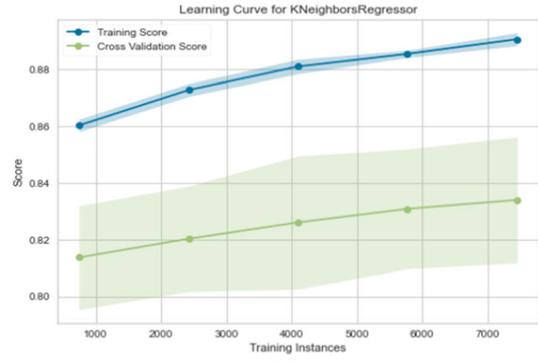
(a)



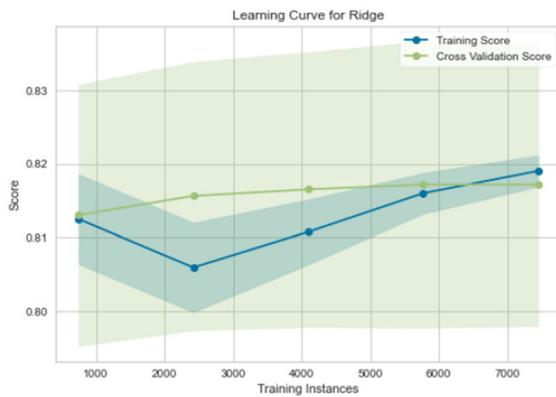
(b)



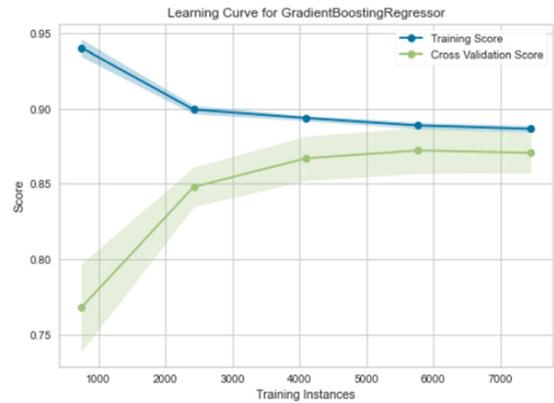
(c)



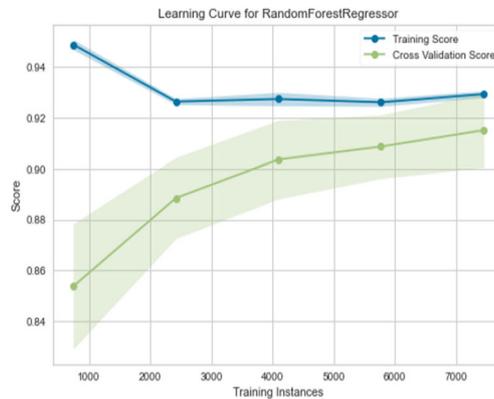
(d)



(e)



(f)



(g)

◀**Fig. 5** Results of machine learning models without feature selection for both training and testing (a) linear regression model, (b) support vector regression model, (c) decision tree model, (d) k-nearest neighbor, (e) ridge regression, (f) XGBoost, (g) random forest regressor

4.4.2 Results without feature selection

The proposed models are performed on the dataset without feature selection. This scenario is performed to highlight the performance of the proposed models without any preprocessing. Table 8 illustrates the training and testing scores. In addition, it contains the evaluation metrics including MAE, MSE, MedAE and R^2 score. It can be observed that the tree algorithm which is deployed using DT achieved a quite high performance with 0.503, 1.186, 0.348 and 0.8926 for MAE, MSE, MedAE and R^2 score, respectively. Moreover, the ensemble method (RF) appears an optimal performance with 0.406, 0.8385, 0.2839 and 0.9229 for MAE, MSE, MedAE and R^2 score, respectively. Furthermore, Fig. 5a–g shows the learning curves of the proposed methods containing the train and validation curves. It can be observed that the proposed DT and RF methods achieved their optimal performance at number of iterations of 7000, while the other deployed method achieved theirs at 8000. This means that they have a high performance with a low complexity. So, these techniques (DT and RF) can be considered as acceptable automatic GCS prediction solutions without preprocessing. However, there is a need to enhance their performance with some preprocessing like feature selection which is discussed in the next subsection.

4.4.3 Results with feature selection

In this section, we perform feature selection algorithm which utilized to magnify the impact of the input features by selecting the most important features to estimate the output value. For this purpose, we used the recursive feature elimination (RFE) algorithm with feature scaling to figure out the most impacted feature by a recursive elimination process which leads to a high performance. Table 9

illustrates the simulation results of the proposed models. It can be observed that DT model achieved 0.39, 0.961, 0.25 and 0.917 for MAE, MSE, MedAE and R^2 score, respectively. In addition, RF model achieved 0.269, 0.625, 0.0784 and 0.946 for MAE, MSE, MedAE and R^2 score, respectively. Thus, it can be noticed the performance of both DT and RF model is considerably increased by deploying feature selection. Furthermore, the proposed SVM regressor achieved 0.283, 0.813, 0.0844 and 0.929 for MAE, MSE MedAE and R^2 score, respectively. So, the proposed models can be considered as efficient solutions for detection of level of consciousness. Figure 6a–g shows the training and testing performance of the developed model performance.

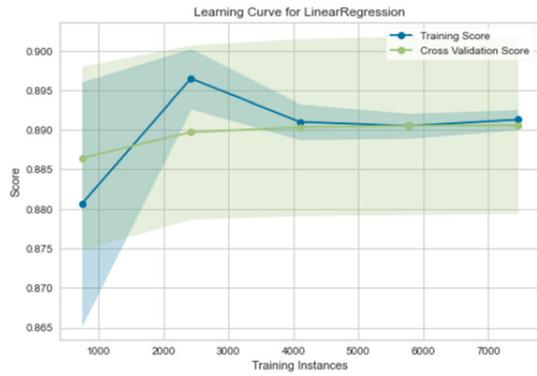
4.5 Model explainability

ML and deep learning (DL) have been widely used in predicting and diagnosing various diseases such as predicting hypertension [101], diabetes [102], sepsis [103]. Unfortunately, most of these studies concentrated on achieving advances in the overall performance of the developed models, while disregarding the interpretability issues. ML and DL are considered as a black box that is unable to provide answers for the medical straightforward inquiries (i.e., why it developed this decision, what the correlation between patient's medical features and the developed output, etc.). Therefore, the physicians find it complex to depend on such models without a clear and understandable explanation. As a result, there exists a serious gap between the developed models and their utilization in medical practice.

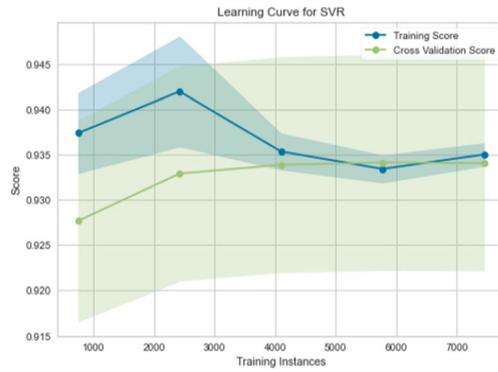
In the latest years, a quite number of studies tried to justifying this issue by explaining the developed models using what is known as explainable artificial intelligence [73, 74, 104]. Explainable artificial intelligence or explainability (XAI) is the ability of ML and DL models to open the black box and provide natural language explanation for the developed decisions [105–107], explain what occurred in the developed model from input features to the final output. It utilized to help non-ML experts to understand the solutions developed by ML models. Therefore,

Table 9 Results of ML models with feature selection

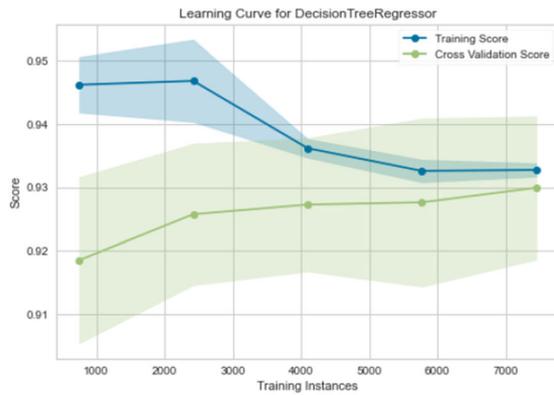
| Algorithm | Training score | Testing score | MAE | MSE | MedAE | R^2 score |
|------------|----------------|---------------|--------------|--------------|---------------|--------------|
| LR | 0.8912 | 0.888 | 0.652 | 1.304 | 0.408 | 0.888 |
| SVM | 0.934 | 0.929 | 0.283 | 0.813 | 0.0844 | 0.929 |
| DT | 0.932 | 0.917 | 0.390 | 0.961 | 0.25 | 0.917 |
| KNN | 0.949 | 0.934 | 0.350 | 0.762 | 0.0 | 0.934 |
| Ridge | 0.8911 | 0.885 | 0.658 | 1.305 | 0.414 | 0.888 |
| (GBR) | 0.927 | 0.916 | 0.475 | 0.974 | 0.240 | 0.916 |
| RF | 0.960 | 0.940 | 0.269 | 0.625 | 0.0784 | 0.946 |



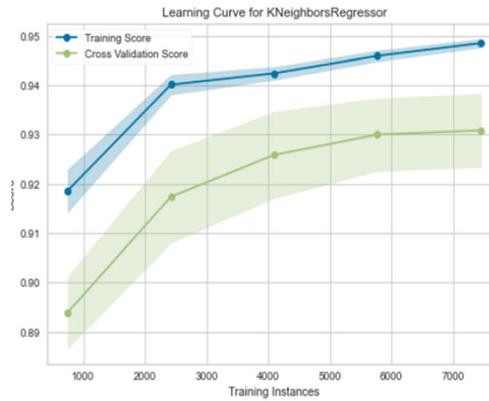
(a)



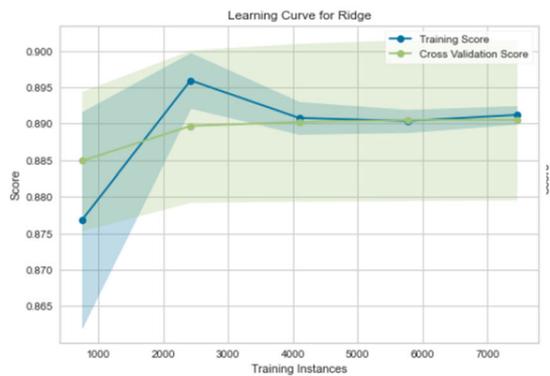
(b)



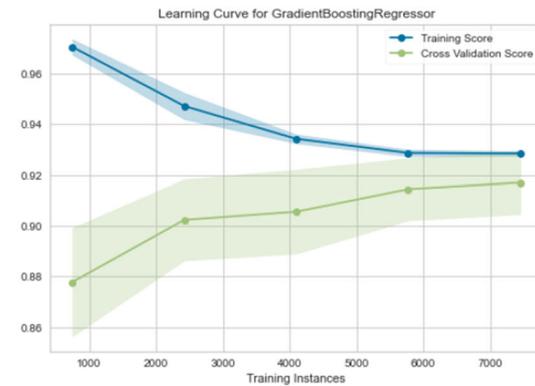
(c)



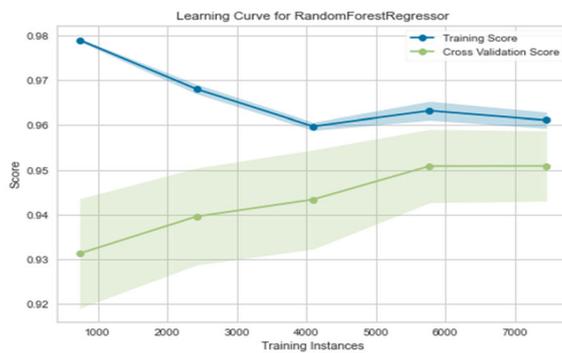
(d)



(e)



(f)



(g)

◀**Fig. 6** Results of machine learning models with feature selection for both training and testing (a) linear regression model, (b) support vector regression model, (c) decision tree model, (d) k-nearest neighbor, (e) ridge regression, (f) XGBoost, (g) random forest regressor

Table 10 Importance of features according to RF model

| Feature name | Correlation using RF |
|-----------------------------------|----------------------|
| Heart rate | 0.09850 |
| SpO ₂ | 0.09146 |
| Temperature C (calc) | 0.04965 |
| Arterial blood pressure systolic | 0.03043 |
| CVP alarm [low] | 0.04972 |
| O ₂ flow | 0.02934 |
| CVP alarm [high] | 0.02933 |
| Respiratory rate | 0.08911 |
| HCO ₃ (serum) | 0.02906 |
| Hemoglobin | 0.02817 |
| Blood urea nitrogen (BUN) | 0.07517 |
| Creatine kinesis | 0.06293 |
| WBC | 0.02269 |
| Insulin | 0.02019 |
| Cardiac index | 0.01621 |
| Cholesterol, HDL | 0.0349 |
| Respiratory effort | 0.06302 |
| Arterial blood pressure diastolic | 0.06255 |
| Glucose finger stick | 0.05606 |
| Arterial pH | 0.01021 |
| weight | 0.01822 |
| Albumin | 0.02703 |
| Urine out void | 0.00257 |

we not only concentrate in developing ML model that could predict GCS based on vital signs and laboratory tests, but also provide an explanation for the developed decision. In this work, we depend on the SHAP library in the explanations issue [108]. SHAP explainers usually depend on a tree-based classifier (i.e., DT, RF, XGBoost, LightGBM, etc.) to calculate the contribution of each feature in the decision [109, 110]. Furthermore, we utilize the internal logic of the RF regressor to discuss the explainability of features as well as cases.

4.5.1 Explainability of features (globally)

Feature importance gives a general view of the rank of all features and the impact of each feature on the final decision. In this work, we depend on RF to calculate the importance of all features. Table 10 shows the feature importance for all features. Unfortunately, we cannot depend on it to specify the direction of each feature. For

instance, we cannot specify if increasing the cardiac index will contribute to increasing the overall score of consciousness or not.

In this section, we utilized SHAP summary plots to show the rank of each feature. As shown in Fig. 7, each line represents one feature, and each dot represents the effect of this feature in a specific instance. Feature correlation is represented by colors (blue for low correlation, and red for high correlation). From the summary plot, we can observe the following: (1) heart rate has a significant impact on the overall decision. (2) Increasing heart rate and O₂ flow value have a positive impact in increasing the overall score. (3) On the contrary, decreasing the value of cholesterol, respiratory effort and CVP alarm [high] have a positive impact on the overall performance of the calculated score. (4) Summary plot also allows us to specify the impact of the outliers. For example, the respiratory effort is not the global critical feature, but it has a high negative effect on some cases. This is indicated in the long-tail that is distributed along the left direction. Features that have a long tail in the right direction are likely to have a positive effect on the total decision. Our medical expert reported that this is a medically intuitive issue that increases the confidence in our model.

To ensure the impact of the chosen feature in the developed model, we extract the feature importance using RF model. Table 10 illustrates the correlation among the whole features and the estimated output. It can observe that there are some features have high correlation such as heart rate, temperature and blood pressure. Others have low correlation values such as cardiac index and arterial PH.

4.5.2 Explainability of cases (locally)

In this section, we will discuss the explainability for each case. For example, as we show in Fig. 8, each example represents a horizontal line. It firstly shows the final prediction for this case (GCS = 11.15). It also shows the features that have a positive impact on the final decision (heart rate = 68, albumin = 6), and features that push the total prediction away from the optimal values (SpO₂ = 90.6%, O₂ flow = 90.83% and BUN = 1.6). This represents the effect of each feature in the final output by colors (red = supported, blue = not supported).

5 Discussion

5.1 Comparison between model performance before and after feature selection (FS)

In this section, we compare the performance of our proposed model in terms of the two-feature list (before and

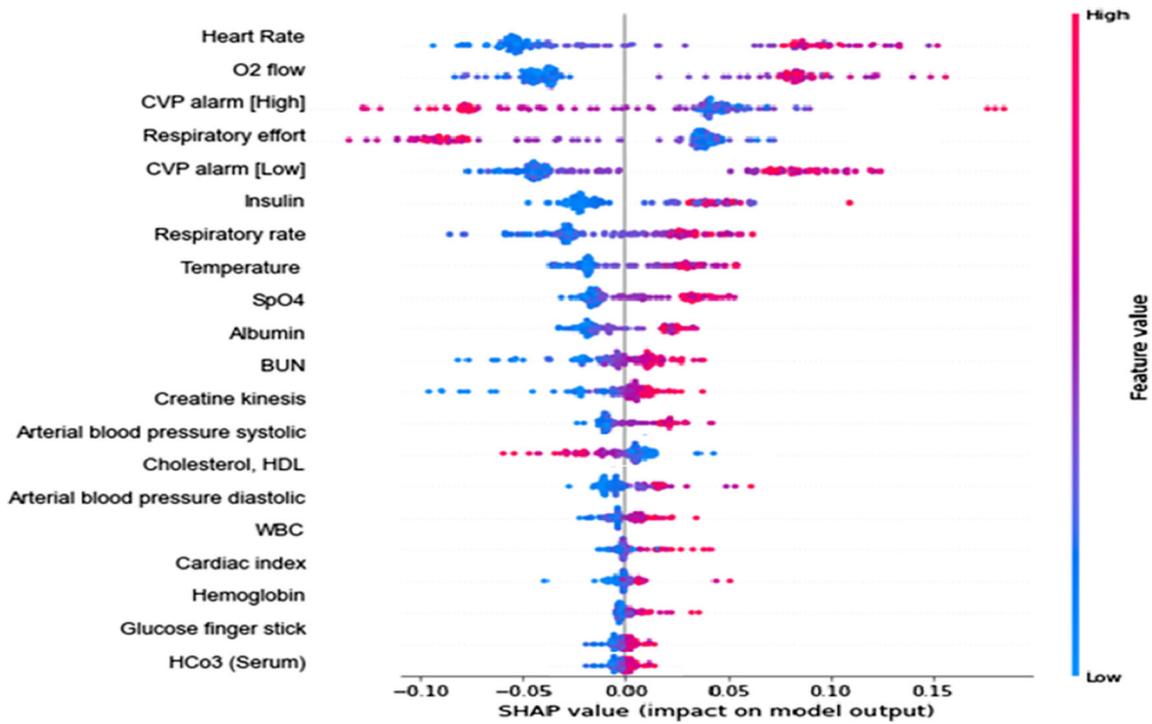


Fig. 7 SHAP summary plots for the proposed model

Fig. 8 SHAP model behavior for cases

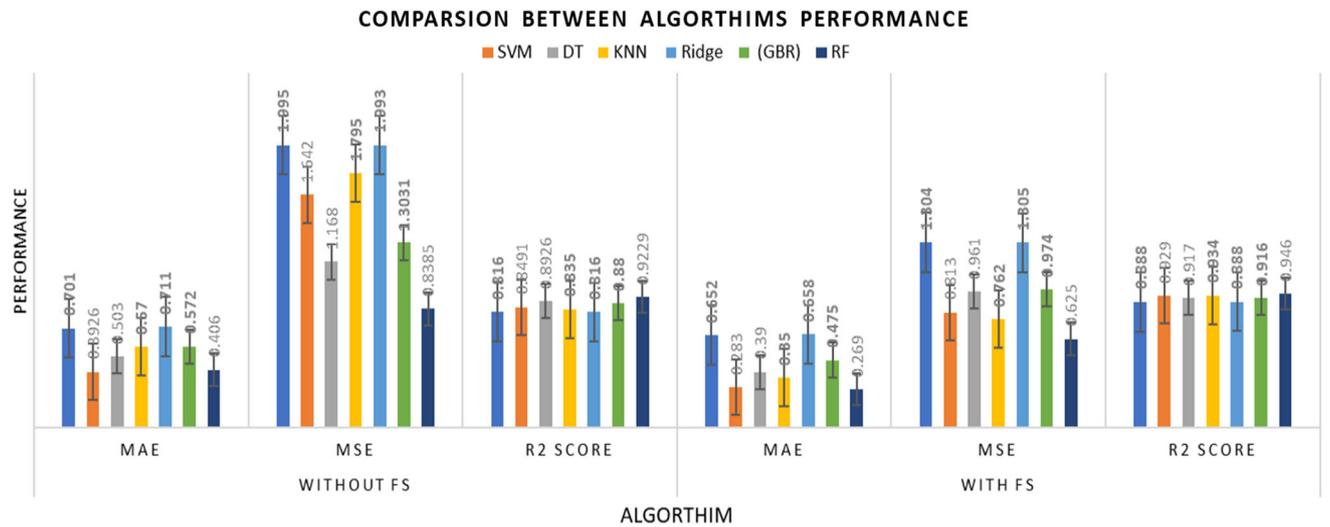
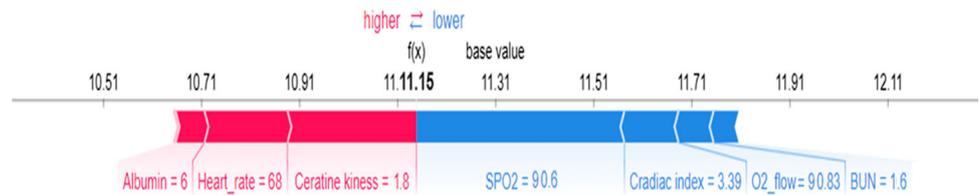


Fig. 9 Comparison between models performance (before and after FS step)

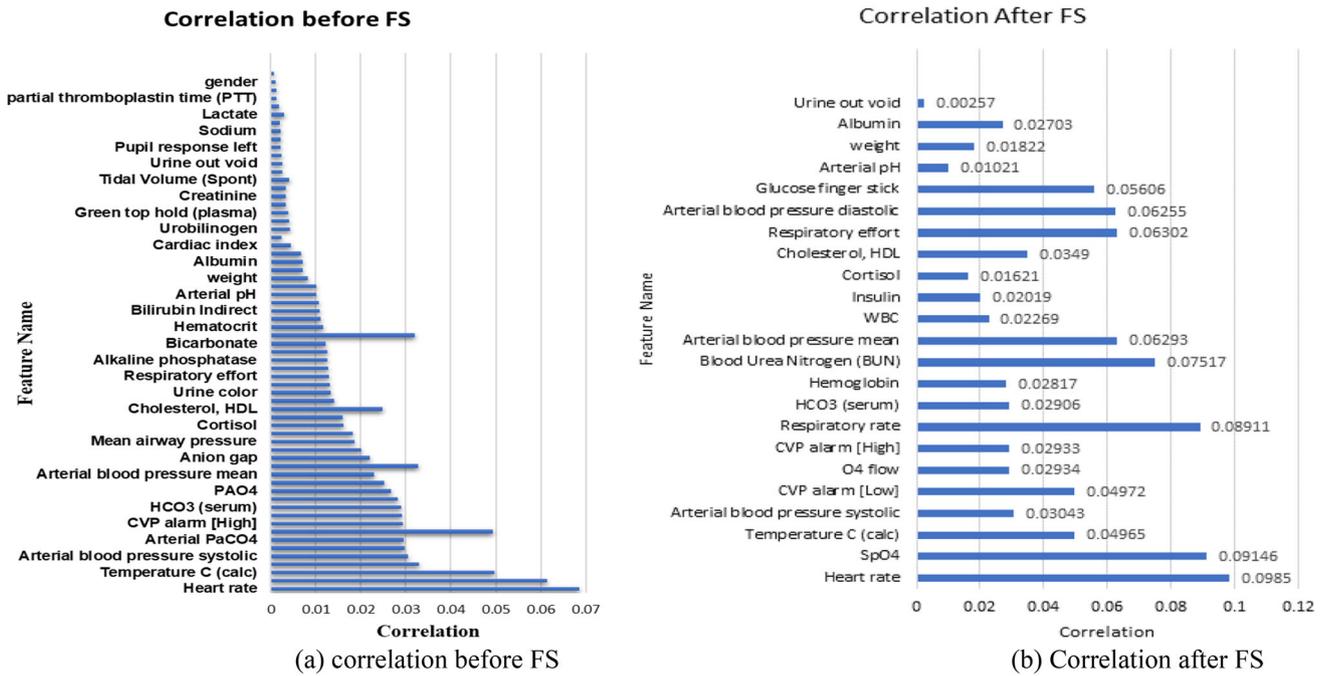


Fig. 10 a Correlation before FS, b Correlation after FS

Table 11 Illustration of the proposed work and the works in the relation

| Work | Method | Model | Metric | Performance |
|----------|------------|-------------------------|-------------|----------------|
| Proposed | Regression | KNN | R^2 score | 0.835 |
| | | SVM | | 0.894 |
| | | RF | | 0.923 |
| | | KNN + feature selection | | 0.934 |
| | | SVM + feature selection | | 0.929 |
| | | RF + feature selection | | 0.946 |
| | | [2] | | Classification |
| SVM | 0.831 | | | |
| RF | 0.925 | | | |

after feature selection). Figure 9 shows comparison between all models in terms of MAE, MSE and R^2 score. From this figure we can observe that the performance improved about 2–6% for all evaluation metrics. This enhancement ensures the importance of this step in the final result. Figure 10a, b shows correlation between the feature list and the output before and after FS stage.

5.2 Study limitations

Although our proposed model provides a promising solution for GCS automatic calculation, it still has some limitations that need to be handled. *First*, there are some situations in which changes in vital signs will not affect GCS, such as small changes in blood pressure that did not reach hypotension, hypothermia that did not affect thermoregulation center. *Second*, MIMIC dataset was extracted

from one institute; therefore, using MIMIC dataset to evaluate the developed model does not guarantee the generalization ability of the model. *Third*, the imputing process for several important features could negatively affect model performance. Therefore, we intend to investigate several imputation techniques. All of these limitations will be addressed in the future studies.

5.3 Comparison with the works in the literature

This paper proposes a GCS prediction system to estimate the level of consciousness of the patients based on their vital signs. For this purpose, several machine learning techniques are deployed to achieve the optimal method. The simulation results are carried out on the MIMIC III dataset with the interest of the vital signs and the level of consciousness. This paper proposes several machine

learning models to handle this issue. The simulation results are carried out on the data with and without feature selection. The feature selection is performed based on the importance of the features and their impact on the output according to their correlation with the output. The simulation results reveal that the proposed SVM, KNN, DT and RF models achieved the optimal performance prior to GCS value prediction. To highlight the performance of the proposed system, we illustrate the impact of the proposed system with the works in the literature. The objective of the proposed system is to predict the value of the GCS using regression. The works in the literature focused on the classification as a solution for this issue. As proposed in [4], they categorized the GCS into three ranges and performed machine learning techniques to classify among them. We show the performance of the proposed system with the proposed in [4] in terms of the mutual machine learning techniques, including KNN, SVM and RF. We compare their performance as classifiers and regressors to solve the problem of diagnosis of level of consciousness. The regression method is evaluated by R^2 score, while the classification method is evaluated by accuracy. Both of the evaluation metrics are within range of 0–1. As shown in Table 11, the proposed models without feature selection achieve 0.835, 0.849 and 0.923 for KNN, SVM and RF, respectively. On the other hand, these models with feature selection achieved 0.934, 0.929 and 0.946 for KNN, SVM and RF, respectively. Therefore, it can be observed that the regression trend achieved a quite high performance rather than classification trend prior to diagnosis of level of consciousness.

6 Conclusion

The problem of detection of level of consciousness has been discussed in this work. This issue has been handled in the presence of IoT system and cloud/edge environment. The proposed framework is based on deploying machine learning for automatic prediction of the level of consciousness based on some vital signs and laboratory tests. Therefore, several machine learning models including both ensemble and kernel models have been implemented to provide a judgeable comparison and extensive study. The simulation results reveal that the proposed ensemble models present a superior performance prior to prediction

of the level of consciousness. Therefore, it can be considered as an efficient solution for consciousness level prediction in IoT and cloud/edge environments.

Appendix A: Features description

| Item_ID | Label | Normal range | Unit of measurement |
|------------------|---------------------------|-------------------------|----------------------|
| 619, 224,690 | Respiratory rate | 12–16 | Breaths per minute |
| 2381, 220,045 | Heart rate | 60–100 | Beats per minute |
| 646, 5820 | SpO ₂ | > 90 | % |
| 470 | O ₂ flow | 80–100 | % |
| 1525 | Creatine | 0.74–1.35 | mg/dL |
| 3066,772 | Albumin | 0.4–5.4 | g/dL |
| 116, 7610 | Cardiac index | 2.5–4.0 | l/min/m ² |
| 1162,781 | Blood urea nitrogen (BUN) | 6–24 | mg/dL |
| 220,050, 220,179 | Blood pressure | 120/80 | mmHg |
| 676, 223,762 | Temperature | 36.1–37.2 | Celsius |
| 440,546, 227,062 | WBC | (4–11)000 | N/μL |
| 225,664 | Glucose finger stick | 80–130 | mg/dL |
| | Insulin | 80–130 | mg/dL |
| 50,909 | Cortisol | 6–8 am, 10–40 pm | mcg/dL |
| 50,904 | Cholesterol, HDL | 40–59 | mg/dL |
| 616 | Respiratory effort | – | – |
| 780 | Arterial pH | 7.25–7.55 | mmHg |
| 40,069 | Urine out void | 0.3–0.5 ml for kg per H | mL |
| | Weight | | Kg |
| | Age | | Y |
| 198, 227,015 | GCS | 11–15 | |

(mg/dl) milligrams/deciliter

(g/dl) grams/deciliter

Appendix B: List of abbreviations

| Term | Abbreviations |
|------|--|
| DL | Deep learning |
| LR | Linear regression |
| SVR | Support vector regression |
| KNN | K nearest neighbor |
| MLP | Multilayer perceptron |
| RF | Random forest |
| BUN | Blood urea nitrogen |
| WBC | White blood cells |
| GCS | Glasgow coma scale |
| UoM | Unit of measurement |
| XAI | Explainability artificial intelligence |
| SHAP | Shapley additive explanations |
| SDN | Software defined network |
| IoT | Internet of things |
| PCA | Principle component analysis |
| FACM | Fog-Assisted Consciousness Management |
| MAE | Mean absolute error |
| MSE | Mean square error |

Funding Open access funding provided by The Science, Technology & Innovation Funding Authority (STDF) in cooperation with The Egyptian Knowledge Bank (EKB).

Data availability Data are available on request due to ethical restrictions.

Declarations

Conflict of interest The authors declare that they have no known competing financial interests or personal relationships that could have appeared to influence the work reported in this paper.

Open Access This article is licensed under a Creative Commons Attribution 4.0 International License, which permits use, sharing, adaptation, distribution and reproduction in any medium or format, as long as you give appropriate credit to the original author(s) and the source, provide a link to the Creative Commons licence, and indicate if changes were made. The images or other third party material in this article are included in the article's Creative Commons licence, unless indicated otherwise in a credit line to the material. If material is not included in the article's Creative Commons licence and your intended use is not permitted by statutory regulation or exceeds the permitted use, you will need to obtain permission directly from the copyright holder. To view a copy of this licence, visit <http://creativecommons.org/licenses/by/4.0/>.

References

1. Tang KJW, Ang CKE, Theodoros C et al (2020) Artificial intelligence and machine learning in emergency medicine. *Biocybern Biomed Eng* 41:156
2. Azimi M, Eslamlou AD, Pekcan G (2020) Data-driven structural health monitoring and damage detection through deep learning: state-of-the-art review. *Sensor* 20:2778
3. Tindall SC (1990) Level of consciousness. In: Walker HK, Hall WD, Hurst JW (eds). Boston
4. Latifoğlu F, Altıntop ÇG, Akin AK, et al (2020) Evaluation of glasgow coma score using electroencephalogram signals. In: 2020 IEEE International Symposium on Medical Measurements and Applications (MeMeA). IEEE, pp 1–6
5. Minami Y, Mishima S, Oda J (2020) Prediction of the level of consciousness using pupillometer measurements in patients with impaired consciousness brought to the emergency and critical care center. *Acute Med Surg* 7:e537. <https://doi.org/10.1002/ams2.537>
6. Yang X, Wang T, Ren X, Yu W (2017) Survey on improving data utility in differentially private sequential data publishing. *IEEE Trans Big Data* 7:729
7. Liang F, Yu W, An D et al (2018) A survey on big data market: pricing, trading and protection. *IEEE Access* 6:15132–15154
8. Stankovic JA (2014) Research directions for the internet of things. *IEEE Internet Things J* 1:3–9
9. Siam AI, Almaiah MA, Al-Zahrani A et al (2021) Secure health monitoring communication systems based on IoT and cloud computing for medical emergency applications. *Comput Intell Neurosci* 2021:1–23. <https://doi.org/10.1155/2021/8016525>
10. Sujith AVLN, Sajja GS, Mahalakshmi V et al (2022) Systematic review of smart health monitoring using deep learning and artificial intelligence. *Neurosci Inform* 2:100028. <https://doi.org/10.1016/j.neuri.2021.100028>
11. Magi N, Prasad BG (2020) Activity monitoring for ICU patients using deep learning and image processing. *SN Comput Sci* 1:123. <https://doi.org/10.1007/s42979-020-00147-6>
12. Chen X-W, Lin X (2014) Big data deep learning: challenges and perspectives. *IEEE Access* 2:514–525
13. Nguyen ND, Nguyen T, Nahavandi S (2017) System design perspective for human-level agents using deep reinforcement learning: a survey. *IEEE Access* 5:27091–27102
14. Siam AI, Abou Elazm A, El-Bahnasawy NA et al (2019) Smart health monitoring system based on IoT and cloud computing. *Menoufia J Electron Eng Res* 28:37–42. <https://doi.org/10.21608/mjeer.2019.76711>
15. Wu X, Liu C, Wang L, Bilal M (2021) Internet of things-enabled real-time health monitoring system using deep learning. *Neural Comput Appl*. <https://doi.org/10.1007/s00521-021-06440-6>
16. Papernot N, McDaniel P, Sinha A, Wellman M (2016) Towards the science of security and privacy in machine learning. *arXiv Prepr arXiv:161103814*
17. Papernot N, McDaniel P, Sinha A, Wellman MP (2018) Sok: security and privacy in machine learning. In: 2018 IEEE European Symposium on Security and Privacy (EuroS&P). IEEE, pp 399–414
18. Li X, He J, Vijayakumar P et al (2021) A verifiable privacy-preserving machine learning prediction scheme for edge-enhanced HCPSS. *IEEE Trans Ind Inform* 18:5494–5503
19. Deo RC (2015) Machine learning in medicine. *Circulation* 132:1920–1930. <https://doi.org/10.1161/CIRCULATIONAHA.115.001593>
20. Myszczyńska MA, Ojames PN, Lacoste AMB et al (2020) Applications of machine learning to diagnosis and treatment of

- neurodegenerative diseases. *Nat Rev Neurol* 16:440–456. <https://doi.org/10.1038/s41582-020-0377-8>
21. Vellido A (2019) Societal issues concerning the application of artificial intelligence in medicine. *Kidney Dis* 5:11–17. <https://doi.org/10.1159/000492428>
 22. Buch VH, Ahmed I, Maruthappu M (2018) Artificial intelligence in medicine: current trends and future possibilities. *Br J Gen Pract* 68:143–144. <https://doi.org/10.3399/bjgp18X695213>
 23. Rajula HSR, Verlato G, Manchia M et al (2020) Comparison of conventional statistical methods with machine learning in medicine: diagnosis, drug development, and treatment. *Medicina (B Aires)* 56:455. <https://doi.org/10.3390/medicina56090455>
 24. Siam AI, Sedik A, El-Shafai W et al (2021) Biosignal classification for human identification based on convolutional neural networks. *Int J Commun Syst*. <https://doi.org/10.1002/dac.4685>
 25. Huang L, Shea AL, Qian H et al (2019) Patient clustering improves efficiency of federated machine learning to predict mortality and hospital stay time using distributed electronic medical records. *J Biomed Inform* 99:103291. <https://doi.org/10.1016/j.jbi.2019.103291>
 26. Detection D, Infections C-, Sedik A, et al (2020) Deploying machine and deep learning models for efficient data-augmented detection of COVID-19 infections
 27. El-Rashidy N, Abdelrazik S, Abuhmed T et al (2021) Comprehensive survey of using machine learning in the covid-19 pandemic. *Diagnostics* 11:1155
 28. Alballa N, Al-Turaiki I (2021) Machine learning approaches in COVID-19 diagnosis, mortality, and severity risk prediction: a review. *Inform Med Unlocked* 24:100564. <https://doi.org/10.1016/j.imu.2021.100564>
 29. Islam MN, Inan TT, Rafi S et al (2020) A systematic review on the use of AI and ML for fighting the COVID-19 pandemic. *IEEE Trans Artif Intell* 1:258–270. <https://doi.org/10.1109/TAI.2021.3062771>
 30. Alyasseri ZAA, Al-Betar MA, Doush IA et al (2022) Review on COVID-19 diagnosis models based on machine learning and deep learning approaches. *Expert Syst*. <https://doi.org/10.1111/essy.12759>
 31. Qin J, Chen L, Liu Y et al (2020) A machine learning methodology for diagnosing chronic kidney disease. *IEEE Access* 8:20991–21002. <https://doi.org/10.1109/ACCESS.2019.2963053>
 32. Kate RJ, Perez RM, Mazumdar D et al (2016) Prediction and detection models for acute kidney injury in hospitalized older adults. *BMC Med Inform Decis Mak* 16:39. <https://doi.org/10.1186/s12911-016-0277-4>
 33. Safdar S, Zafar S, Zafar N, Khan NF (2018) Machine learning based decision support systems (DSS) for heart disease diagnosis: a review. *Artif Intell Rev* 50:597–623. <https://doi.org/10.1007/s10462-017-9552-8>
 34. Ahsan MM, Siddique Z (2022) Machine learning-based heart disease diagnosis: a systematic literature review. *Artif Intell Med* 128:102289. <https://doi.org/10.1016/j.artmed.2022.102289>
 35. Mahmood H, Shaban M, Rajpoot N, Khurram SA (2021) Artificial Intelligence-based methods in head and neck cancer diagnosis: an overview. *Br J Cancer* 124:1934–1940. <https://doi.org/10.1038/s41416-021-01386-x>
 36. Saxena S, Gyanchandani M (2020) Machine learning methods for computer-aided breast cancer diagnosis using histopathology: a narrative review. *J Med Imaging Radiat Sci* 51:182–193. <https://doi.org/10.1016/j.jmir.2019.11.001>
 37. Zou Q, Qu K, Luo Y et al (2018) Predicting diabetes mellitus with machine learning techniques. *Front Genet*. <https://doi.org/10.3389/fgene.2018.00515>
 38. Khalil H, El-Hag N, Sedik A et al (2019) Classification of diabetic retinopathy types based on convolution neural network (CNN). *Menoufia J Electron Eng Res* 28:126–153
 39. Kassem MA, Hosny KM, Damaševičius R, Eltoukhy MM (2021) Machine learning and deep learning methods for skin lesion classification and diagnosis: a systematic review. *Diagnostics* 11:1390. <https://doi.org/10.3390/diagnostics11081390>
 40. Orooji A, Kermani F (2021) Machine learning based methods for handling imbalanced data in hepatitis diagnosis. *Front Heal Inform* 10:57. <https://doi.org/10.30699/fhi.v10i1.259>
 41. Spann A, Yasodhara A, Kang J et al (2020) Applying machine learning in liver disease and transplantation: a comprehensive review. *Hepatology* 71:1093–1105. <https://doi.org/10.1002/hep.31103>
 42. Alharbey R, Dessouky MM, Sedik A et al (2022) Fatigue state detection for tired persons in presence of driving periods. *IEEE Access*. <https://doi.org/10.1109/ACCESS.2022.3185251>
 43. Teasdale G, Jennett B (1974) Assessment of coma and impaired consciousness. A practical scale. *Lancet* 304:81–84. [https://doi.org/10.1016/S0140-6736\(74\)91639-0](https://doi.org/10.1016/S0140-6736(74)91639-0)
 44. Sternbach GL (2000) The glasgow coma scale. *J Emerg Med* 19:67–71. [https://doi.org/10.1016/S0736-4679\(00\)00182-7](https://doi.org/10.1016/S0736-4679(00)00182-7)
 45. Teasdale G, Jennett B (1976) Assessment and prognosis of coma after head injury. *Acta Neurochir (Wien)* 34:45–55. <https://doi.org/10.1007/BF01405862>
 46. Estévez-Báez M, Machado C, García-Sánchez B et al (2019) Autonomic impairment of patients in coma with different glasgow coma score assessed with heart rate variability. *Brain Inj* 33:496–516. <https://doi.org/10.1080/02699052.2018.1553312>
 47. Young AJ, Hare A, Subramanian M et al (2021) Using machine learning to make predictions in patients who fall. *J Surg Res* 257:118–127. <https://doi.org/10.1016/j.jss.2020.07.047>
 48. Durga S, Nag R, Daniel E (2019) Survey on machine learning and deep learning algorithms used in internet of things (IoT) healthcare. In: 2019 3rd International Conference on Computing Methodologies and Communication (ICCMC). IEEE, pp 1018–1022
 49. Tang KJW, Ang CKE, Constantinides T et al (2021) Artificial intelligence and machine learning in emergency medicine. *Biocybern Biomed Eng* 41:156–172. <https://doi.org/10.1016/j.bbe.2020.12.002>
 50. Kim J, Chang H, Kim D et al (2020) Machine learning for prediction of septic shock at initial triage in emergency department. *J Crit Care* 55:163–170. <https://doi.org/10.1016/j.jcrc.2019.09.024>
 51. Liu N, Zhang Z, Wah Ho AF, Ong MEH (2018) Artificial intelligence in emergency medicine. *J Emerg Crit Care Med* 2:82–82. <https://doi.org/10.21037/jccm.2018.10.08>
 52. Raita Y, Goto T, Faridi MK et al (2019) Emergency department triage prediction of clinical outcomes using machine learning models. *Crit Care* 23:64. <https://doi.org/10.1186/s13054-019-2351-7>
 53. Gravesteijn BY, Nieboer D, Ercole A et al (2020) Machine learning algorithms performed no better than regression models for prognostication in traumatic brain injury. *J Clin Epidemiol* 122:95–107. <https://doi.org/10.1016/j.jclinepi.2020.03.005>
 54. Matsuo K, Aihara H, Nakai T et al (2020) Machine learning to predict in-hospital morbidity and mortality after traumatic brain injury. *J Neurotrauma* 37:202–210. <https://doi.org/10.1089/neu.2018.6276>
 55. Tsiklidis EJ, Sims C, Sinno T, Diamond SL (2020) Using the National Trauma Data Bank (NTDB) and machine learning to predict trauma patient mortality at admission. *PLoS One* 15:166. <https://doi.org/10.1371/journal.pone.0242166>
 56. Hall AN, Weaver B, Liotta E et al (2021) Identifying modifiable predictors of patient outcomes after intracerebral hemorrhage

- with machine learning. *Neurocrit Care* 34:73–84. <https://doi.org/10.1007/s12028-020-00982-8>
57. Amorim RL, Oliveira LM, Malbouisson LM et al (2020) Prediction of early TBI mortality using a machine learning approach in a LMIC population. *Front Neurol*. <https://doi.org/10.3389/fneur.2019.01366>
 58. de Toledo P, Rios PM, Ledezma A et al (2009) Predicting the outcome of patients with subarachnoid hemorrhage using machine learning techniques. *IEEE Trans Inf Technol Biomed* 13:794–801. <https://doi.org/10.1109/TITB.2009.2020434>
 59. Kamruzzaman MM, Alanazi S, Alruwaili M et al (2022) Fuzzy-assisted machine learning framework for the fog-computing system in remote healthcare monitoring. *Measurement* 195:111085. <https://doi.org/10.1016/j.measurement.2022.111085>
 60. Tuli S, Basumatary N, Gill SS et al (2020) HealthFog: an ensemble deep learning based smart healthcare system for automatic diagnosis of heart diseases in integrated IoT and fog computing environments. *Futur Gener Comput Syst* 104:187–200. <https://doi.org/10.1016/j.future.2019.10.043>
 61. Abdulkareem KH, Mohammed MA, Gunasekaran SS et al (2019) A Review of fog computing and machine learning: concepts, applications, challenges, and open issues. *IEEE Access* 7:153123–153140. <https://doi.org/10.1109/ACCESS.2019.2947542>
 62. Ali ZH, Badawy MM, Ali HA (2020) A novel geographically distributed architecture based on fog technology for improving Vehicular Ad hoc Network (VANET) performance. *Peer-to-Peer Netw Appl* 13:1539–1566
 63. Ali ZH, Hagrass S, Ali HA (2021) Distributed computing architecture using fog technology for improving intelligent transportation systems in smart city. *Int J Comput Appl* 183:42–45. <https://doi.org/10.5120/ijca2021921351>
 64. Verma P, Tiwari R, Hong W-C et al (2022) FETCH: a deep learning-based fog computing and IoT integrated environment for healthcare monitoring and diagnosis. *IEEE Access* 10:12548–12563. <https://doi.org/10.1109/ACCESS.2022.3143793>
 65. Kishor A, Chakraborty C, Jeberson W (2021) A novel fog computing approach for minimization of latency in healthcare using machine learning. *Int J Interact Multimed Artif Intell* 6:7. <https://doi.org/10.9781/ijimai.2020.12.004>
 66. Sudqi Khater B, Abdul Wahab AW, Bin IMYI et al (2019) A lightweight perception-based intrusion detection system for fog computing. *Appl Sci* 9:178. <https://doi.org/10.3390/app9010178>
 67. Hu P, Dhelim S, Ning H, Qiu T (2017) Survey on fog computing: architecture, key technologies, applications and open issues. *J Netw Comput Appl* 98:27–42. <https://doi.org/10.1016/j.jnca.2017.09.002>
 68. Karakus M, Duresi A (2017) Quality of service (QoS) in software defined networking (SDN): a survey. *J Netw Comput Appl* 80:200–218
 69. Shaukat U, Ahmed E, Anwar Z, Xia F (2016) Cloudlet deployment in local wireless networks: motivation, architectures, applications, and open challenges. *J Netw Comput Appl* 62:18–40
 70. Zhang P, Zhou M, Fortino G (2018) Security and trust issues in fog computing: a survey. *Futur Gener Comput Syst* 88:16–27
 71. Hou X, Li Y, Chen M et al (2016) Vehicular fog computing: a viewpoint of vehicles as the infrastructures. *IEEE Trans Veh Technol* 65:3860–3873
 72. Liu H, Cocea M (2019) Nature-inspired framework of ensemble learning for collaborative classification in granular computing context. *Granul Comput* 4:715–724. <https://doi.org/10.1007/s41066-018-0122-5>
 73. El-Sappagh S, Alonso JM, Islam SMR et al (2021) A multilayer multimodal detection and prediction model based on explainable artificial intelligence for Alzheimer's disease. *Sci Rep* 11:1–26. <https://doi.org/10.1038/s41598-021-82098-3>
 74. Vellido A (2020) The importance of interpretability and visualization in machine learning for applications in medicine and health care. *Neural Comput Appl* 32:18069–18083
 75. Fawagreh K, Gaber MM, Elyan E (2016) An outlier ranking tree selection approach to extreme pruning of random forests. *Commun Comput Inf Sci* 629:267–282. https://doi.org/10.1007/978-3-319-44188-7_20
 76. Ameixieira C, Cardote A, Neves F et al (2014) Harbornet: a real-world testbed for vehicular networks. *IEEE Commun Mag* 52:108–114
 77. Openfog reference architecture for fog computing. <https://www.openfogconsortium.org/>. Accessed 17 Apr 2020
 78. Johnson EWA, Pollard TJ, Shen L (2016) Data descriptor: MIMIC-III, a freely accessible critical care database. *Thromb Haemost* 76:258–262. <https://doi.org/10.1038/sdata.2016.35>
 79. Johnson A, Pollard T, Mark R MIMIC-III Clinical Database v1.4
 80. Caicedo-Torres W, Gutierrez J (2019) ISeeU: visually interpretable deep learning for mortality prediction inside the ICU. *J Biomed Inform* 98:1–24. <https://doi.org/10.1016/j.jbi.2019.103269>
 81. Adams RP, Mayaud L, Poincare HR et al (2015) A physiological time series dynamics-based approach to patient monitoring and outcome prediction. *IEEE J Biomed Heal Inform* 19:1068–1076. <https://doi.org/10.1109/JBHI.2014.2330827.A>
 82. El-rashidy N, El-sappagh S, Abuhmed T, Abdelrazek S, El-Bakry HM (2020) Intensive care unit mortality prediction: an improved patient-specific stacking ensemble model. *IEEE Access*. <https://doi.org/10.1109/ACCESS.2020.3010556>
 83. El-rashidy N, El-sappagh S, Abdelrazik S, El-bakry H (2022) Ensemble machine learning model for mortality prediction inside intensive care unit. Springer International Publishing
 84. Greco L, Luta G, Krzywinski M, Altman N (2019) Analyzing outliers: robust methods to the rescue. *Nat Methods* 16:275–276. <https://doi.org/10.1038/s41592-019-0369-z>
 85. Moon TK (1996) The expectation-maximization algorithm. *IEEE Signal Process Mag* 13:47–60. <https://doi.org/10.1109/79.543975>
 86. Joensuu DW, Bankhofer U (2015) Hot deck methods for imputing missing data hot deck methods for imputing missing data the effects of limiting donor usage <https://doi.org/10.1007/0097836.4231.53746>
 87. Caballero-Ruiz E, García-Sáez G, Rigla M et al (2017) A web-based clinical decision support system for gestational diabetes: automatic diet prescription and detection of insulin needs. *Int J Med Inform* 102:35–49. <https://doi.org/10.1016/j.ijmedinf.2017.02.014>
 88. Wright J (2018) Glasgow coma scale. pp 1–2
 89. Cook N (2020) The glasgow coma scale. *Crit Care Nurs Clin North Am*. <https://doi.org/10.1016/j.cnc.2020.10.005>
 90. Ko J, Deprez D, Shaw K et al (2020) Stretching is superior to brisk walking for reducing blood pressure in people with high-normal blood pressure or stage I hypertension. *J Phys Act Health* 18:21–28. <https://doi.org/10.1123/jpah.2020-0365>
 91. Berntson GG, Bigger JTJ, Eckberg DL et al (1997) Heart rate variability: origins, methods, and interpretive caveats. *Psychophysiology* 34:623–648. <https://doi.org/10.1111/j.1469-8986.1997.tb02140.x>
 92. Andršová I, Hnatkova K, Šišáková M et al (2021) Influence of heart rate correction formulas on QTc interval stability. *Sci Rep* 11:1–21. <https://doi.org/10.1038/s41598-021-93774-9>

93. van der Ven WH, Schuurmans J, Schenk J et al (2022) Monitoring, management, and outcome of hypotension in intensive care unit patients, an international survey of the European Society of intensive care medicine. *J Crit Care* 67:118–125. <https://doi.org/10.1016/j.jcrc.2021.10.008>
94. Al-Rashed F, Sindhu S, Al Madhoun A et al (2021) Elevated resting heart rate as a predictor of inflammation and cardiovascular risk in healthy obese individuals. *Sci Rep* 11:13883. <https://doi.org/10.1038/s41598-021-93449-5>
95. Chan NC, Li K, Hirsh J (2020) Peripheral oxygen saturation in older persons wearing nonmedical face masks in community settings. *JAMA* 324:2323–2324. <https://doi.org/10.1001/jama.2020.21905>
96. Tapio J, Vähänikkilä H, Kesäniemi YA et al (2021) Higher hemoglobin levels are an independent risk factor for adverse metabolism and higher mortality in a 20-year follow-up. *Sci Rep* 11:1–13. <https://doi.org/10.1038/s41598-021-99217-9>
97. Lee SH, Kim M, Do HK, Lee JH (2021) Low hemoglobin levels and an increased risk of psoriasis in patients with chronic kidney disease. *Sci Rep* 11:1–7. <https://doi.org/10.1038/s41598-021-94165-w>
98. Forman JP, Rifas-Shiman SL, Taylor EN et al (2008) Association between the serum anion gap and blood pressure among patients at Harvard vanguard medical associates. *J Hum Hypertens* 22:122–125. <https://doi.org/10.1038/sj.jhh.1002286>
99. Anestis DM, Tsitsopoulos PP, Foroglou NG et al (2021) Cross-cultural adaptation and validation of the greek version of the “full outline of unresponsiveness score”: a prospective observational clinimetric study in neurosurgical patients. *Neurocrit Care*. <https://doi.org/10.1007/s12028-021-01342-w>
100. Schmidt WU, Lutz M, Ploner CJ, Braun M (2021) The diagnostic value of the neurological examination in coma of unknown etiology. *J Neurol* 268:3826–3834. <https://doi.org/10.1007/s00415-021-10527-4>
101. Andalib S, Lattanzi S, Di Napoli M et al (2020) Blood pressure variability: a new predicting factor for clinical outcomes of intracerebral hemorrhage. *J Stroke Cerebrovasc Dis Off J Natl Stroke Assoc* 29:105340. <https://doi.org/10.1016/j.jstrokecerebrovasdis.2020.105340>
102. Wagner R, Heni M, Tabák AG et al (2021) Pathophysiology-based subphenotyping of individuals at elevated risk for type 2 diabetes. *Nat Med* 27:49–57. <https://doi.org/10.1038/s41591-020-1116-9>
103. El-Rashidy N, Abuhmed T, Alarabi L et al (2021) Sepsis prediction in intensive care unit based on genetic feature optimization and stacked deep ensemble learning. Springer, London
104. Zheng Q, Delingette H, Ayache N (2019) Explainable cardiac pathology classification on cine MRI with motion characterization by semi-supervised learning of apparent flow. *Med Image Anal* 56:80–95. <https://doi.org/10.1016/j.media.2019.06.001>
105. Lee H, Yune S, Mansouri M et al (2019) An explainable deep-learning algorithm for the detection of acute intracranial haemorrhage from small datasets. *Nat Biomed Eng* 3:173–182. <https://doi.org/10.1038/s41551-018-0324-9>
106. Jiménez-Luna J, Grisoni F, Schneider G (2020) Drug discovery with explainable artificial intelligence. *Nat Mach Intell* 2:573–584. <https://doi.org/10.1038/s42256-020-00236-4>
107. Rudin C (2019) Stop explaining black box machine learning models for high stakes decisions and use interpretable models instead. *Nat Mach Intell* 1:206–215. <https://doi.org/10.1038/s42256-019-0048-x>
108. Gulum MA, Trombley CM, Kantardzic M (2021) A review of explainable deep learning cancer detection models in medical imaging. *Appl Sci*. <https://doi.org/10.3390/app11104573>
109. Kakogeorgiou I, Karantzas K (2021) Evaluating explainable artificial intelligence methods for multi-label deep learning classification tasks in remote sensing. *Int J Appl Earth Obs Geoinf* 103:102520. <https://doi.org/10.1016/j.jag.2021.102520>
110. Petch J, Di S, Nelson W (2021) Opening the black box: the promise and limitations of explainable machine learning in cardiology. *Can J Cardiol*. <https://doi.org/10.1016/j.cjca.2021.09.004>

Publisher's Note Springer Nature remains neutral with regard to jurisdictional claims in published maps and institutional affiliations.