



Utilizing fog computing and explainable deep learning techniques for gestational diabetes prediction

Nora El-Rashidy¹ · Nesma E. ElSayed^{2,5} · Amir El-Ghamry^{2,3,4} · Fatma M. Talaat¹

Received: 11 November 2021 / Accepted: 26 October 2022 / Published online: 20 December 2022
© The Author(s) 2022

Abstract

Gestational diabetes mellitus (GDM) is one of the pregnancy complications that poses a significant risk on mothers and babies as well. GDM usually diagnosed at 22–26 of gestation. However, the early prediction is desirable as it may contribute to decrease the risk. The continuous monitoring for mother's vital signs helps in predicting any deterioration during pregnancy. The originality of this paper is to provide comprehensive framework for pregnancy women monitoring. The proposed Data Replacement and Prediction Framework consists of three layers which are: (i) IoT Layer, (ii) Fog Layer, and (iii) Cloud Layer. The first layer used IOT sensors to aggregate vital sings from pregnancies using invasive and noninvasive sensors. Then the vital signs transmitted to fog nodes to processed and finally stored in the cloud layer. The main contribution in this paper is located in the fog layer producing GDM module to implement two influential tasks which are: (i) Data Finding Methodology (DFM), and (ii) Explainable Prediction Algorithm (EPM) using DNN. First, the DFM is used to replace the unused data to free the cache space for the new incoming data items. The cache replacement is very important in the case of healthcare system as the incoming vital signs are frequent and must be replaced continuously. Second, the EPM is used to predict the incidence of GDM that may occur in the second trimester of the pregnancy. To evaluate our model, we extract data of 16,354 pregnancy women from medical information mart for intensive care (MIMIC III) benchmark dataset. For each woman, vital signs, demographic data and laboratory tests was aggregated. The results of the prediction model superior the state of the art (ACC = 0.957, AUC = 0.942). Regarding to explainability, we utilized Shapley additive explanation framework to provide local and global explanation for the developed models. Overall, the proposed framework is medically intuitive, allow the early prediction of GDM with cost effective solution.

Keywords Gestational diabetes mellitus (GDM) · Deep learning (DL) · Fog computing · Healthcare system · Effective data caching · Machine learning

1 Introduction

GDM is an irregular glucose level status that occurred during pregnancy. It is a common pregnancy complication that recognized in 3–10% of the pregnancies [1–3]. GDM is usually diagnosed between 22 and 26 of gestational, may result in high-risk complications for both women and infants. These risks include respiratory problems, metabolic complications, premature delivery, and the fetus may gain increased weight that may hamper the birthing process. Despite the fact that GDM usually goes away after birth, women are still at a significant risk of developing type 2 diabetes, with a cumulative incidence of 30–50%

within 5–10 years following the index pregnancy [3, 4]. Several studies reported that the high-risk complications could be avoided if the medical intervention started at the first or the begging of the second trimesters [5].

Hence early detection of GDM is critical for avoiding a variety of problems and it is important in issues such as: (i) Evidence suggests that pre-diabetes treatment response varies when GDM history is taken into consideration [5, 6]. (ii) Individualized risk prediction and treatment response estimation could also help guide pre-diabetes treatment selections [7, 8]. (iii) Women with a history of GDM can learn about their diabetes risk in future and how metformin and/or ILI might help [8, 9]. Individual risk estimation could help doctors make better clinical decisions and make

diabetes prevention programs more efficient, cost-effective, and patient-centered [7, 8].

For the general population, there are a variety of models available to estimate the risk of acquiring diabetes [10–13]. However, few people employ multivariable models to help customize preventive treatments to specific people [7, 8, 13, 14]. Predictors in models designed specifically for women with prior GDM frequently incorporate measurements taken during or shortly after pregnancy (e.g., insulin use during pregnancy or breastfeeding history) [13, 15, 16].

In the last decades, several studies have been utilized data from electronic health record (EHR) to diagnosis and predict patient future events such as mortality prediction [17, 18], sepsis prediction [19–21], predict heart problems [22], and GDM complications [23, 24]. However, little number of studies work on predicting GDM [25–27]. For example: (i) a recent study develop formula to predict GDM, this formula includes pregnancy body mass index (BMI), gestational age fasting glucose. (ii) Xiong et al. [28] decided to use support vector Machine (SVM) and light gradient boosting Machine to create a risk prediction mechanism for the first 19 weeks using high-potential GDM predictors (light GBM). (iii) Using biochemical markers and the ML method, Zheng et al. [29] presented a straightforward method for detecting GDM in early pregnancy. (iv) In a study conducted by Shen et al. [30], it was stated that the exploration of the best AI approach for GDM prediction required the least number of clinical devices and trainees in order to construct an AI-based application (AI). [31, 32].

Other studies [33, 34] have attempted to develop models based on risk factors discovered in the first trimester that can predict an abnormal OGTT at 24–28 weeks. They considered several factors predict GDM include scoring methods, glucose biochemistry assays, and glycosylated hemoglobin (HbA1c) levels have all been used in different populations with varying degrees of success [35, 36]. GDM can be averted, according to clinical study, if a comprehensive lifestyle modification is performed before the 20th week of pregnancy [37, 38].

Unless the good performance of preciously developed models for GDM, all of them neglecting the explainability issue, centering on optimizing the performance of the ML model. Therefore, most of them are not accepted in the medical domain. Explainability become important in ML applications, as it help to provide transparent model that could explain the output decision. Traditional models that deal with hundreds of variables find it difficult to understand the impact of each feature on the overall decision, and the features that make the developed model move toward one of the classes. Ongoing explainability of models is another important issue, as it used to detect the variable importance and the effect of changes in the

developed model. Our goal is to develop a clinical diabetes risk prediction model that is specific to women with GDM who have already been diagnosed.

The prediction model depends on the vital signs come from set of sensors connected to the woman. And when we talk about sensors sending data, it leads to talking about IoT [39]. IoT generates big data, which requires sending these data to cloud-based Data Centers. To minimize the latency which is very critical issue in such cases like healthcare [40], Fog Computing (FC) is a mandatory decision. FC is not a replacement for Cloud Computing, but rather an extension of it that makes use of resources from devices near the edge [41]. Hence, The FC increases QoS parameters such as bandwidth efficiency and energy usage while also lowering latency [42].

The originality of this paper is to provide comprehensive framework for pregnancy women monitoring. The proposed Data Replacement and Prediction Framework (DRPF) consists of three layers which are: (i) IoT Layer, (ii) Fog Layer, and (iii) Cloud Layer. The first layer used IOT sensors to aggregate vital sings from pregnancies using invasive and noninvasive sensors. Then the vital signs transmitted to fog nodes to processed and finally stored in the cloud layer. The main contribution in this paper is located in the fog layer producing GDM module to implement two influential tasks which are: (i) Data Finding Methodology (DFM), and (ii) Explainable Prediction Algorithm (EPM) using DNN. First, the DFM is used to replace the unused data to free the cache space for the new incoming data items. The cache replacement is very important in the case of healthcare system as the incoming vital signs are frequent and must be replaced continuously. Second, the EPM is used to predict the incidence of GDM that may occur in the second trimester of the pregnancy.

The rest of paper is organized as follows: Sect. 2 gives a background for some basic concepts. Section 3 introduces the recent previous efforts in the field of deep learning algorithms used to analyze and predict deterioration during pregnancy. Section 4 introduces a proposed Data Replacement and Prediction Framework (DRPF) with more details about each contribution. Section 5 introduces the implementation and evaluation. Our conclusion is discussed in Sect. 6.

2 Background and basic concepts

This section introduces some concepts in the field of Probabilistic Neural Networks (PNN), fog in healthcare applications, and data caching in fog.

2.1 Probabilistic neural networks (PNN)

A probabilistic neural network (PNN) is a type of feed forward neural network that is commonly used to solve classification and pattern recognition tasks. A Parzen window and a non-parametric function are used to approximate the parent probability distribution function (PDF) of each class in the PNN method. PNN are organized into a four-layer multilayered feed forward network [43, 44]: (i) Input layer: is made up of nodes that each have a collection of metrics. (ii) Pattern layer: Each example in the training data set has its own neuron. It calculates the test case's Euclidean distance from the neuron's center point, then uses the sigma values to apply the Radial Basis Function (RBF) kernel function. (iii) Summation layer: For each class, executes a sum operation on the outputs from such as the highest-scoring label node. PNN has a number of advantages, including: [43]: (1) PNN networks predict target probability scores with high accuracy, and (2) As the size of the representative training set grows, it is guaranteed to converge to an optimal classifier second layer. (iv) Output layer: takes all of the summation nodes' outputs and outputs the maximum.

In classification and pattern recognition applications, PNNs offer a scalable alternative to traditional back-propagation neural networks. They do not necessitate the huge forward and backward calculations that ordinary neural networks necessitate. They can also handle various sorts of training data. When applied to a classification task, these networks use the concept of probability theory to reduce misclassifications.

2.2 Explainability and interpretability of deep learning models

Researchers sometimes interchange the terms interpretability and explainability; nevertheless, while these terms are extremely similar, other works discriminate between them. There is no definite mathematical definition for interpretability or explainability, and they have never been measured; however, several attempts have been made to define not only these two terms, but other related notions like comprehensibility. All of these definitions, however, lack mathematical formality and rigor.

Explainable AI (XAI) is a framework that used to open the black box of the machine learning, help in understanding the output of the machine learning models [45]. Explainability also defined as the degree in which the human could understand the ML decision [46], provide insights on how the ML model, discuss the logic goes to take this decision. Applying XAI provide three main advantages include (1) provide clear explanation increase

the trust in the developed model. (2) Enable model troubleshooting, (3) specifying the source of model basis. Explainability and accuracy considered two separate issues that should maintained when building ML models. Generally, algorithms with high accuracy performance are not able to give a clear explanation to their decision and vice versa. Two main types of AI explainability include global method that apply to understand the overall behavior of the model and effect of each feature in the output decision, local method that used to clarify the decision of the model for each instance [47]. The interpretable model is very critical, especially in medical domains to translate the output decision to human understandable language.

2.3 Fog in healthcare applications

The cloud, on the other hand, is unsuited for mission-critical applications. High bandwidth requirements, periodic delays, and safety and security difficulties are all issues that cloud-based applications face. Real-time monitoring is required for healthcare applications. Real-time requirements cannot be met by the cloud. Delays occur when data are sent to the cloud and then returned to the application.

Healthcare services and applications are delay sensitive. They deal with private data of the patients [48]. The patients' data contain very sensitive and personal data so the data location should be secured. High latency may cause many problems in tele-health and telemedicine applications, which makes FC a suitable paradigm in healthcare applications. As for many applications in health informatics, a simple sensor-to-cloud architecture is impractical. Regulations prohibit the storage of patient data outside of a hospital in specific instances. Because of patient safety concerns in the event of network and data center failures, reliance exclusively on remote data centers is also unsuitable for some applications [49]. Fog computing is one possible option for bridging the gap between sensors and analytics in health informatics.

2.4 Data caching in fog

Reduced latency is a critical issue in the fog computing paradigm as the number of time-sensitive applications grows. As a result, one of the goals of an effective IoT application is to reduce fog computing latency. This approach uses popularity-based caching to achieve this goal, with a strong emphasis on the users' interests.

For boosting data availability and lowering access latency, data caching is a critical issue in FC. Each Fog Node (FN) is so small, therefore cache replenishment is a critical concern. In the FC context, cache replacement achieves load balancing by ensuring data availability.

Cooperative caching is the most frequent data caching strategy in FC. In this case, each FN's local cache is shared with its neighbors, resulting in a big unified cache. Each node in a cooperative caching system can get data not only from its own local cache but also from the caches of its neighbors.

As a result, the data availability is maximized, the access delay is minimized, and the response time for the end-user layer is reduced. FNs share data in a variety of fog applications, including healthcare [50], smart homes [51], industrial systems [52], and intelligent traffic signs [53]. As a result, sharing cache contents among FNs has a lot of advantages. By picking a suitable set of data for caching, the cache replacement method plays a key role in lowering response time. When the cache fills up, a data item must be removed to make room for the data that needs to be fetched. The performance will be improved if the least used data object is chosen.

3 Related work

This section will divide in to main section as follows (i) provide literature review of using fog computing in reducing latency, (ii) introduces some of the recent previous efforts in the field of deep learning algorithms used to analyze and predict deterioration during pregnancy.

3.1 Utilizing fog computing in healthcare systems

Real-time monitoring is required for healthcare applications. Real-time requirements cannot be met by the cloud [54, 55]. For latency-sensitive applications, the cloud is ineffective. Fog computing has been presented as a solution to these issues. Ahmad et al. [56] suggested a health fog system in which fog computing serves as an intermediary layer between the cloud and the end-user. Communication expenses are reduced thanks to this three-layer architecture.

Shukla et al. [57] presented a smart fog computing architecture to reduce latency and network traffic. Requests can be processed locally before being sent to the cloud in this three-layer design. Fog computing serves as a middle layer that improves network services while reducing the downsides of IoT health. In the healthcare IoT, fog nodes are employed to reduce latency. Greco et al. [58] presented a layered architecture aimed at solving health monitoring issues. There are two types of health monitoring problems: static and dynamic monitoring.

An IoT-Fog-Cloud ecosystem was proposed by Alli et al. [59]. It is an intriguing architecture in which IoT devices respond to user requests. End devices are on the bottom, the fog layer is in the center, and the cloud layer is

at the top. Localized computation, fog-edge computing, and remote computing are all supported by this architecture. Abdelmoneem et al. [60] present a system that dynamically distributes healthcare tasks across cloud and fog computing. This architecture will handle a wide range of health issues and a large number of individuals.

3.2 Utilizing deep learning in predicting GDM

Predicting pregnancies deterioration considered critical issue in medical domain. Recently, various studies utilizing ML and DL to predict GDM and its consequence. For example, j. Wang [61] utilizing various ML algorithms include random forest (RF), support vector machine (SVM), and artificial neural network (ANN) to predict GDM. The model evaluated on a data collected from different hospitals in eastern China, result in accuracies ranged from 81 to 86%. Another study [32] utilized patient electronic health record (EHR) to predict GDM during early pregnancy based. The authors first employed six ML include (SVM, NN, logistic regression (LR), bayesian network, and CHAID tree), then they developed cost effective hybrid model to improve accuracy. Result in accuracy of 86.5% and 84.7% for training and testing, respectively. The same in [62], A.Sumathi provide voting ensemble classifier based on various ML techniques include (LR, SVM, RF, k-nearest neighbor (KNN))result in accuracy of 94.24%.

In [53], Y. Liu et al first studied the impact of different types of features and multiclass feature combination on predicting GDM. They devised a feature screening method for determining the importance of traits in order to automatically filter the appropriate number of features. Then, they vectorized features using depth representation methods like network embedding, analyzed the relationship between features using a similarity measurement method, and finally applied it to the classification model for prediction. This approach could learn some aspects automatically based on both domain knowledge rather than artificial rules, resulting in superior results. Unless the enhanced performance of the developed model, It necessitates extra time and money to manage data by humans.

When the features filtered by Wideband Bandpass Filters (WBFs) as in [54], the accuracy, F1 value, and AUC value of logistic regression are 0.809, 0.881, and 0.825, respectively, which was a 12 percent gain when compared to when the feature is not used. The findings showed that a data drive based on electronic medical records can significantly enhance the accuracy of forecasting gestational diabetes. Y. Zhong et al. [55] created a method to assess the risk of GDM in second-trimester pregnancy. This model based on a variety of risk factors, that has a high predictive value for the development of GDM in pregnant

women in China and may be useful in directing future clinical practice. However, In terms of liver function, there was no significant difference between the two groups, which is an essential indicator of visceral fat metabolism (especially hepatic fat metabolism).

For women with past GDM, M.Schwartz et al. [63] built and internally verified a therapeutically useful prediction model that includes fasting glucose, HbA1c, BMI, treatment arm, and BMI by treatment arm interaction. Integrating personalized diabetes risk prediction into prediabetes therapy decision-making should help researchers better grasp the benefits of ILI and/or metformin in diabetes prevention. For personalized decision-making in the management of prediabetes in women with past GDM, a clinical prediction model was devised. For women with a prior GDM, the estimated incidence of diabetes without therapy was 37.4%, compared to 20.0 percent with comprehensive lifestyle modification or metformin treatment. It is officially predicated on the presumption of a lady who has had a previous GDM. In most circumstances, it is not very accurate. In [64], F. Guo et al. created a simple nomogram for pregnant Chinese women that may be used to predict the likelihood of getting GDM during the first antenatal visit. This method could detect GDM early allowing more effective management to improve maternal outcomes. On the other hand, the AUC statistic is just concerned with prediction accuracy. A model with a higher AUC but a little lower sensitivity might be a better choice for clinical application. As a result, we employed decision-analytic approaches to assess the worthiness of a model or alternatives, using our findings and theory.

4 The proposed data replacement and prediction framework (DRPF)

One of the most significant applications related to the aims of IoT is an efficient healthcare system. In this regard, many factors should be taken into consideration such as time, privacy of data, and accuracy. The healthcare system should be reliable and available at any time. Accordingly, this paper is concerned with designing an IoT-Fog based healthcare system as shown in Fig. 1. The proposed Data Replacement and Prediction Framework (DRPF) consists of three layers which are: (i) IoT Layer, (ii) Fog Layer, and (iii) Cloud Layer. The IoT layer combines the IoT devices (pulse oximeter, ECG monitor, etc.) to observe the user status. The fog layer is considered with handling the incoming requests and forwards them to the suitable Fog Node (FN). The fog layer is divided into set of fog regions, and layer 3 is the cloud datacenters. The following subsections details the roles of the proposed layers.

4.1 IoT layer

IoT devices utilized as it provides wide range of flexibility, for example, if a patient requires constant care, he or she can remain at home rather than in a hospital and be monitored on a frequent basis utilizing IoT technology. The data transferred from the sensor to the control device and then to the monitoring center will be affected by noise, lowering the data quality. On the IoT, monitoring a large number of users necessitates more storage and infrastructure, which may be avoided by keeping data in the cloud.

4.2 Cloud layer

Cloud datacenters locate at remote distance from IoT devices which leads to high latency. This issue adversely affects the response time for real time applications such as critical health monitoring systems, traffic monitoring, and emergency fire. In addition, IoT sources are geographically extended and can generate a large amount of data sent to cloud for processing which lead to overloading. The edge computational resources can handle the previous mentioned challenges in IoT systems.

The patient data which come from the IoT sensors is sent to the used application which uses the proposed GDM module. The used application sends its data to be processed in the fog layer. The main module called GDM Module is implemented in the fog layer as shown in Fig. 2 is running in the fog layer. GDM module is used to predict GDM with low latency.

4.3 Fog layer

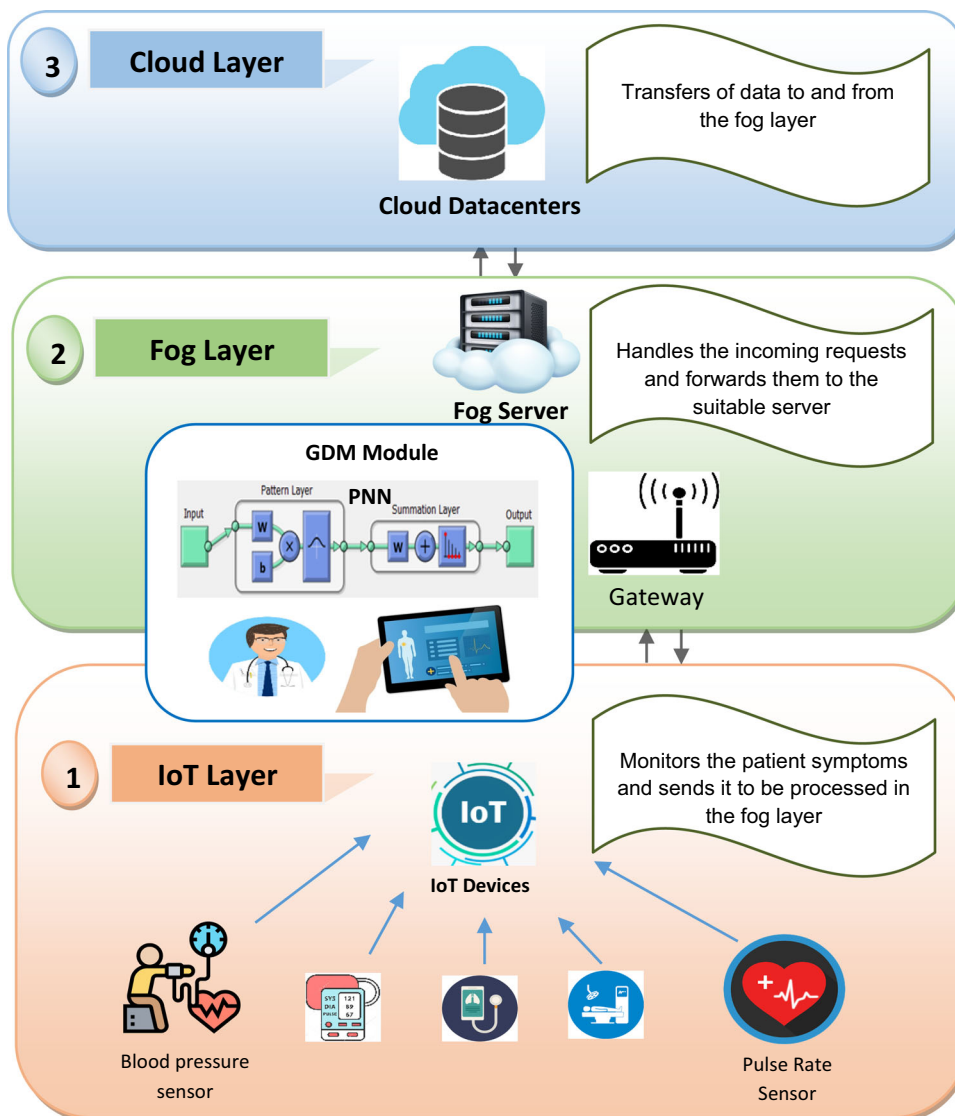
Fog can be considered as a computing paradigm which performs IoT applications at the edge of the network. The Fog improves the QoS metrics such as (bandwidth efficiency and energy consumption) and reduces latency. The main mission of fog is to deliver data and place it closer to the user.

4.3.1 The proposed GDM module

The proposed GDM module is composed of two main sub-modules: (i) Data Finding Methodology (DFM), and (ii) Explainable Prediction Algorithm (EPM) using DNN.

4.3.1.1 (i) Data finding methodology (DFM) The DFM is used to replace the unused data to free the cache space for the new incoming data items. The cache replacement is very important in the case of healthcare system as the incoming vital signs are frequent and must be replaced continuously. Caching in a fog environment is constrained

Fig. 1 The Proposed Effective Prediction Methodology (EPM)



by bandwidth limitations, power limitations, and cache space limitations. To discriminate between data items that should be preserved in the cache and those that should be discarded when the cache is full, a decent replacement mechanism is necessary.

The network is divided into fog regions and each region has a Master Node (MN) that manages the communication in each fog region. The MN collects the required features about each fog node such as: (i) Existing Data (ED), (ii) Time-To-Live (TTL), and (ii) Cache Size (CS). The MN periodically checks each data features to delete the data items with zero TTL. If the cache of the Fog sever is full and there is an incoming data, the MN can decide to remove a data item according to some criteria. Each FN has a table called Data Cache Table (DCT) which contains information about each di in its cache memory such as:

(data item (di), Access Time (TA), Size of data (S), Access Frequency (FA), Access Count (AC), Time-To-Live (TTL), and Cache Free Size (CFS) as shown in Table 1.

The DFM leads to periodically update the cache and decrease the latency. The suggestive measures have been taken into account to measure the performance of the caching schemes are: (i) Hit Ratio (HR), (ii) access latency, and (iii) power consumption. The access latency is defined as the average packet delay over a multi-hop route. It is used as a measure of the accessibility of the nodes.

Using PNN, algorithm can decide to remove a data item and replace it with new incoming data according to its features. The input to the PNN is: TA, AC, and FA. The output of PNN is Data Replace (DR). DR can be Yes or No. The steps of PNN-based cache replacement strategy are shown in Algorithm 1.

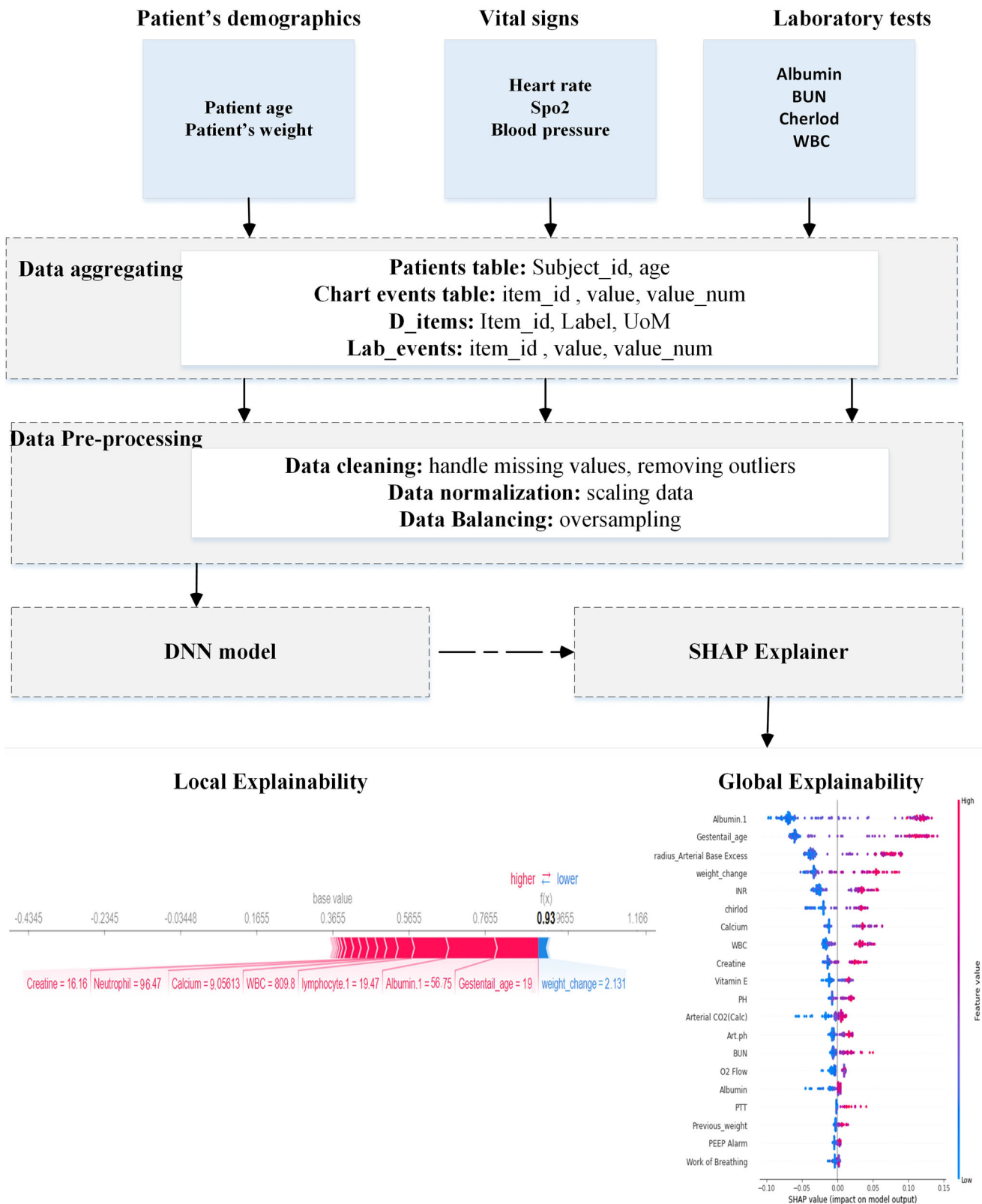


Fig. 2 Explainable Prediction Algorithm (EPM) using DNN

Table 1 Data cache table (DCT)

Data item (di)	Access time (TA)	Size of data (S)	Access frequency (FA)	Access count (AC)	Time-to-live (TTL)	Cache free size (CFS)
Refers to data item's number	The time at which data item enters the cache of the FN	Size of data item	The value of how many times di was accessed. It indicates the importance of the di	Each data item maintains a count which gives the number of FNs having the same data	TTL is a time period given to each di when it is located at this cache	The free space in the cache

PNN based cache replacement strategy

- **Input:**

- Accessing Time (TA), Accessing Count (AC), Accessing Frequency (FA)

- **Output:**

- Data Replacement (DR)

- **Steps:**

// for training:

```
int PNN(int X, int Y, int Z, float S, bool test_example[T], float Examples[Y][T]){
    int MS = -1; float largest = 0;
    float sum[ X ], Product[ d ], sump[ d ], SumS[ Y_L ];
    float sump, ExSum;
    // The OUTPUT layer which computes the Gaussian functions (G) for each class X
    for ( int L=0; L<X; L++ ){
        sum[ L ] = 0;
        // The SUMMATION layer that accumulates the Gaussian functions (G)
        // for each example from the particular class L
        for ( int M=0; M<Y_L; M++ ){
            float product = 0;
            // The PATTERN layer that computes eculidean distance
            // for each parameter from the particular class
            for ( int N=0; N<d; N++ ){
                Product[N] = test_example[N] - Examples[M][N];
                Product[N] = (product[N]* product[N]) / (S * S);
                Product[N] = exp( product[N] );
                sump += product[N];
            }
            SumS[M] = sump;
            ExSum+=SumS[M];
        }
        sum[ L ] = ExSum;
        sum[ L ] /= Y_L;
    }
    Target=sum[0];
    for ( int L=1; L<=X; L++ ){
        if ( sum[ L ] > largest ){
            largest = sum[ L ];
            MS = L;
        }
    }
    MS }}

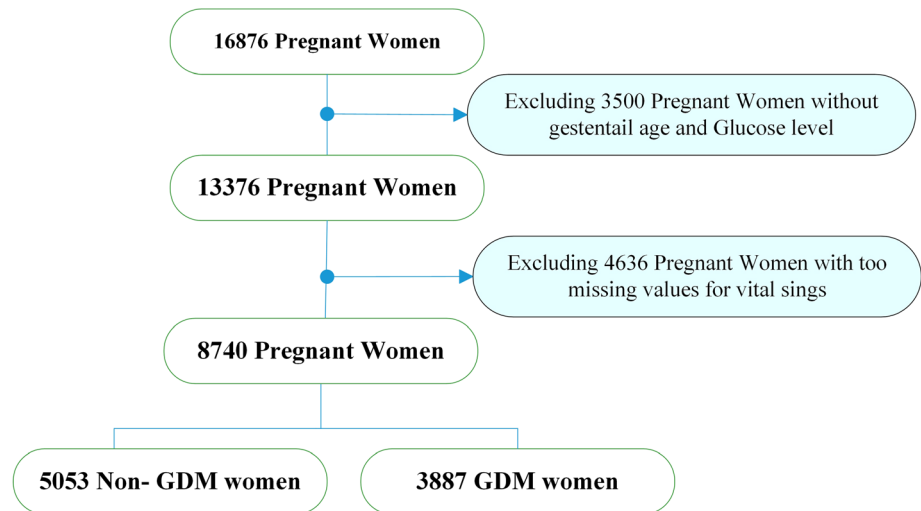
```

4.3.1.2 (ii) Explainable prediction algorithm (EPM) using DNN

This section proposed EPA model that aimed to detect the incidence of GDM among pregnancies. In addition to provide understandable explanation to the predicted output. We evaluated our model based on MIMIC III dataset. As shown in Fig. 3, The proposed EPM consists of four main steps: (a) Data Collection: collecting the required dataset using PostgreSQL, extracting data from various tables include (patients, chartevents, D_itmes, lab-events, and input_events), (b) Data Preprocessing: The output from

the first step cleaned and preprocessed using different steps include (removing outliers, standardization and balancing), (c) Feature Extraction: utilizing DNN to build classification model that could detect the incidence of GDM. (d) Developing DL model: The output decision then utilized SHAP explainer to provide understandable explanation to the developed decision. The performance of our model evaluated using unseen data to ensure the efficiency the proposed model is promising, accurate, and explainable.

Fig. 3 Selected Data



(a) *Data collection*: Medical Information Mart for Intensive Care III (MIMIC III) is a benchmark dataset that developed by MIT Lab. It includes HER data for patients inside ICU. MIMIC III accessible by getting confirmation from Physionet Organization. MIMIC III includes the data for 53,422 distinct patients. 4750 measurement and 390 laboratory tests included in MIMIC III dataset. As shown in Fig. 3, in this study we extract the data from MIMIC III dataset include patient's demographics (i.e., age, gender, BMI), vital signs (i.e., heart rate, respiratory rate, glucose level, etc.) and laboratory test (i.e., Albumin, Creatine, Cholesterol, sodium, etc.) The present study was conducted on 8740 pregnant women.

according to inclusion criteria includes: (i) female gender that was adult (age > 20). (ii) Recorded as pregnant in mimic iii database (item_id (pregnant = 225,082, pregnant due date = 225,083). Gestational age between 6 to 26 weeks. Existing of required vital signs and laboratory tests. Features used in EPM is detailed in Table 2.

(b) *Data preprocessing*: The output from the first step cleaned and preprocessed using different steps include removing outliers, standardization and balancing [65]. The steps of data preprocessing are as follow: (i) Data balancing: Class imbalance is a common problem, especially with medical dataset. In MIMIC III a minor number of pregnant women have GDM which may lead to the problem of imbalanced dataset. Two main techniques commonly used to handle this issue include oversampling [66] and under-sampling [67]. Oversampling techniques used to increase number of samples in the minority class such as synthetic minority oversampling technique, where under sampling used to remove samples from the majority class such as Tomek link and random under sampling. In this study we used the random under sampling technique to keep the data balanced. The main advantage of using under sampling technique is that it does not add any noise to the dataset.

(ii) Handle missing values: MIMIC dataset includes about 15–20% of missing data. Several statistical techniques used to impute the missing values such as expectation maximization [68], hot decking encoding [69], etc. in this study we removed data with more 50% missing data. We only selected patients that have at least one record for each vital signs per day. Then, forward and backward filling used to fill patient's data. (iii) Scaling data: The extracted features have different values which may vary in their value. These variations usually affect classifier performance. Therefore, in this study we scaled all features to be ranged from 0 to 1 using Minmax scaling [70]

(c) *Feature extraction*: In this section we extracted two feature subset A, and B as shown in Table 3. Feature set A: include the main vital signs include (heart rate, glucose level, SpO2, blood pressure, etc.), and some laboratory tests include (PCT, total bilirubin, etc.). feature set B: include all features in feature A, in addition to some features related to pregnancy such as Gestational age, weight change, and other laboratory features such as Lymphocyte, Sodium, Vitamin E, and Neutrophil these features have a critical effect of GDM detection. For example, Vitamin E is a critical measure to maintain the metabolic of the body and scavenging radical activities. The deficiency of Vitamin E among pregnancies may lead to vascular endothelial, incidence of GDM, and hypertension, in addition to placental and premature birth [51]. Therefore. Considering vitamin E is important in GDM prediction. The same for Lymphocyte, the count decreases during the first and the second trimesters and increased during the third one. Increasing lymphocyte may also contribute to irregular glucose level.

(d) *Developing DL model*: DL model includes 20 input dimensions using dense and dropout layers. Dense layers considered a neural network that connected deeply. Each neuron in each layer receives the output from the previous

Table 2 Features used in EPM

Feature_ID	Feature_name	UOM	Average for GDM	Average for non-GDM	P_value
	BMI	–	28 ± 6.2	21.66 ± 3.2	< 0.05
3692	Weight Change	kg	12 ± 12.8	10 ± 7.9	< 0.05
3583	Previous Weight	Kg	75 ± 15.3	66 ± 7.2	< 0.05
3446	gestenail age		24.4 ± 1.2	18 ± 2.3	< 0.05
1127	WBC (4–11, 000)	(*103 / μm)	9.48 ± 2.6	8.87 ± 1.3	< 0.05
626	Neutrophil	%	69.21 ± 8.9	71 ± 8.7	< 0.05
220,635	PCT	%	0.20 ± 0.05	0.17 ± 0.61	< 0.01
220,645	Sodium	mEq/L	142 ± 3.2	135 ± 4.2	< 0.05
223,830	PH (Arterial)	–	7.45 ± 0.2	7.35 ± 2.2	< 0.01
223,751	Noninvasive Blood Pressure	mmHg	125 ± 5.8/90 ± 5.6	115 ± 3.8/75 ± 3.4	< 0.05
2381, 220,045	Heart rate	Bit per M	70 ± 23	60 ± 22	< 0.05
646, 5820	Spo2	%	95 ± 4.2	95 ± 5	
1126	Platelet	(× 103 / μm)	231.0 ± 62.6	198.0 ± 62.6	< 0.05
783	lymphocyte	%	25.9 ± 7.4	24.8 ± 6.9	< 0.05
772,227,456	Albumin (> 3.2)	(g/L)	44 ± 9.8	3412.2	< 0.05
1529	Glucose	mg/dL	100 ± 25.3	90 ± 22.7	< 0.05
1525	Creatinine	mg/dL	0.7 ± 0.5	. 0.6 ± 0.2	< 0.02
1523	Chloride	(mEq/L)	100 ± 3.2	96 ± 4.4	< 0.05
3684	Vitamin E	mg/l	9.20 ± 2.37	10.80 ± 5.01	< 0.05
1522	Calcium	mg/dL	9.3 ± 1.8	8.6 ± 2.2	< 0.05

Table 3 Features used in model A and model B

Model	Features
Model A	Age, BMI, Respiratory rate, Heart Rate, Glucose level, SPO2, blood pressure, Calcium, Sodium, PH (Arterial), Total Bilirubin, PCT
Model B	Age, BMI, Weight change, Gestenail age, Respiratory rate, Heart Rate, Glucose level, blood pressure, BUN, PH (Arterial), SPO2, PTT, Vitamin E, Neutrophil, Lymphocyte, Glucose, Creatinine, Creatinine, Calcium, Sodium, PCT

layers. Dense layers also utilized to change the vector dimension. Dropout layer is regularization approach that used to randomly ignoring some neurons during training

process to avoid overfitting [71]. As shown in Fig. 4, in the hidden layers, we used activation function rectified linear activation function or “ReLU,” it is a liner activation function that out the input directly if it is positive, otherwise, it will output zero. Learning process done using backpropagation algorithm [72]. This method helps calculate the gradient of a loss function with respect to all the weights in the network. In the last layer, we utilized the sigmoid activation function for binary classification [73]. This result in a robust network that have a good generalization ability and less likely to overfit.

Deep learning model

• **Defintion:**

W_{ij}^k : weight for neuron j in layer l_k for incoming node i (a neuron if $K \geq 1$) in layer l_{k-1}

B_i^k : bias for neuron i in layer l_k

h_i^k : product sum plus bias i in layer l_k

o_i^k : Output for node i in layer l_k

r_k : number of nodes in layer l_k

w_i^{-k} weight vector for neuron i in layer l_k

o^{-k} : output vector for layer l_k

• **Steps:**

- 1- Initialize the input layer l_0
Set the values of the outputs o_i^0 for nodes in input layer l_0 to their associated inputs in the vectors $\vec{x} = \{x_1, \dots, x_n\}$
- 2- Calculate the product of sums and outputs of each hidden layer in order from l_1, l_{m-1}
For k from 1 to $m - 1$
 - a. Compute $h_i^k = w_i^{\vec{k}} \cdot O^{k-1} + b_i^k = b_i^k + \sum_{j=1}^{r_{k-1}} w_{ji}^k o_j^{k-1}$ for $i = 1, \dots, r_k$;
 - b. Compute $O_i^k = g(h_i^k)$ for $i = 1, \dots, r_k$;
- 3- Compute the output y for output l_m
 - a. Compute $h_1^m = w_1^{\vec{m}} \cdot O^{m-1} + b_1^m = b_1^m + \sum_{j=1}^{r_{m-1}} w_{j1}^m o_j^{m-1}$
 - b. Compute $o = O_1^m = g_0(h_1^m)$

Learning process

Given a set of N input-output pairs $X = \{(x_1, y), \dots (x_n, y_n)\}$, learning consists of iteratively updating the values of w_i^k and b_j^k to minimize the cost function. in our model, we used the cross-entropy cost function that derived from the maximum likelihood principle through the following equation

$$C_{EC} = \sum_{i=1}^N \sum_{k=1}^C t_{k,i} \log(y_{k,i})$$

5 Results

5.1 Evaluation metrics

The performance of the proposed intrusion detect method has been evaluated using the confusion matrix consisting of: True Positive (TP), True Negative (TN), False Positive (FP), and False.

True Positive (TP): # of records successfully recognized as injection attack.

- True Negative (TN): # of records classified as a normal class.
- False Positive (FP): # of records wrongly classified as injection attack
- False Negative (FN): # of injection attacks undetected by IDS

For our proposed model, we used various metrics include accuracy, Accuracy, Precision, Recall, and F1

metrics, and AUC. The cross-validation (CV) results are calculated based on the training data, and the generalization performance is measured based on the testing data. Table 4 details the used evaluation metrics.

5.2 Results of DL model

In model A, we used the basic feature set such as patient’s age, heart rate, blood pressure and other vital signs to predict GDM. EPM achieves adequate performance (Accuracy = 0.902%, AUC = 0.912%). Model B used the same features in model A, in addition to other features include gestenail age, weight change and other laboratory tests such as Albumin, vitamin E which have impact on GDM incidence. The results demonstrate that the performance increased when adding weight-change and gestenail change (Accuracy = 0.957%, AUC = 0.942%). From the previous experiments, we observed the following: (i) patients with GDM ranged from 25 to 45 years (average

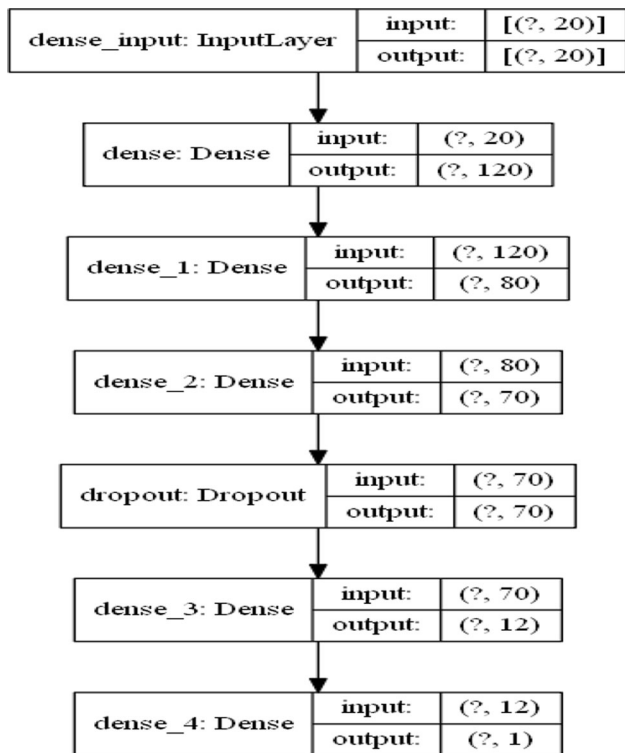


Fig. 4 Deep learning model

32.12 ± 5.6). (ii) GDM usually appeared in gestenail age between (19–26 weeks). (iii) Both BMI and weight change during pregnancy is highly associated with GDM (GDM average for BMI was 28 ± 6.2 and for non GDM was 21.66 ± 3.2). (iv) GDM pregnant women associated with significance difference ($P < 0.05$) in liver and kidney functions that reflected in high values for Albumin, BUN, SBP, TC, etc. The overall results are illustrated in Table 5 and Fig. 5.

5.3 Statistical analysis

To ensure the superiority of the developed DNN model, both model a and b are compared using Freidman test [74]. Freidman test is non-parametric test that is used to

Table 5 Features used in model A and model B

DL Model	Accuracy	Precision	Recall	F-score	AUC
Model A	0.926%	0.9321%	0.8431%	0.902%	0.912%
Model B	0.957%	0.949%	0.892%	0.937%	0.942%

determine if there is significant difference between models. In order to choose the best performance model according to statistical test, the average rank for each model is calculated based on the Nemenyi test [75]. Results of the Nemenyi test could be visualized using the critical difference diagram. Figure 6 shows a comparison between classification models based on the critical difference calculated based on the results of the Nemenyi test for all models. The test shows a significance difference between the developed models (Statistics = 9.855, $P < 0.005$). Figure 6 shows that model B give the improved performance over model A (i.e., AUC = 0.942, $P < 0.005$) followed by the same feature set after.

5.4 Evaluation of explainability of DL model

5.4.1 Global explainability

In this section, we used SHAP summary plots to show the behavior of the developed model in terms of different values of several features. As shown in Fig. 7, each horizontal line represents one feature and the number of dots represent the correlation between the feature and the overall decision. Color of the dots represent the nature of correlation (Red for high, and blue for low). From Fig. 7, we make the following observations. (i) Albumin and weight change have a significant correlation with GDM prediction, higher values have a positive impact of predicting GDM. (ii) High values of PTT have negative effect of predicting GDM. (iii) Summary plot allows specifying the effect of the outliers. For example, weight change is not the most critical

Table 4 Evaluation metrics

Metric	Abbreviation	Equation	Definition
Accuracy	ACC	$\frac{tp+tn}{tp+fp+tn+fn}$	The percentage between number of cases that are correctly classified and the total number of cases
Precision	P	$\frac{m}{m+fp}$	The percentage of the negative class cases that classified correctly
Recall	R	$\frac{m}{m+fn}$	The ratio of the number of positive records that are classified correctly to the total number of classifications in the actual class. This means that the higher the FN rate
F1-score	F1	$\frac{2(P * R)}{P + R}$	Measure is known as the harmonic mean of the precision and the recall measures. It is considered a good evaluation measure for imbalanced data

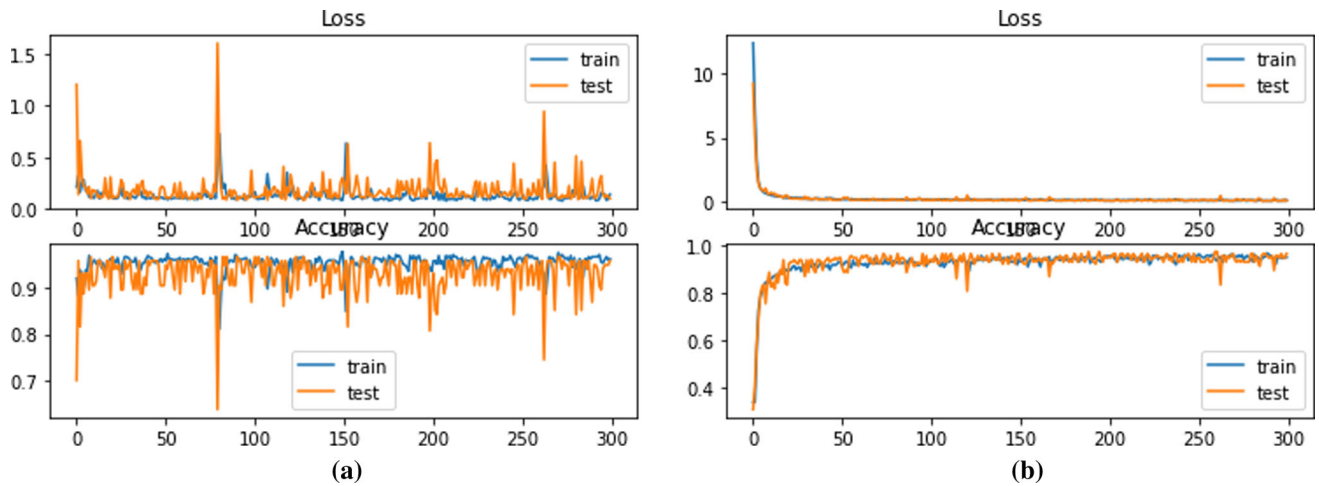


Fig. 5 Results of DL model **a** accuracy and loss results for model A, **b** accuracy and loss results for model B

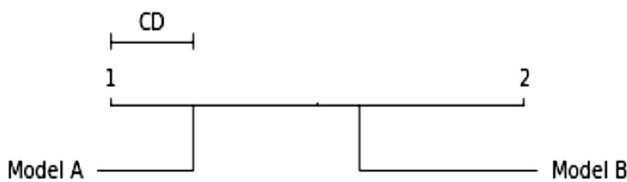


Fig. 6 Critical difference between the Model A and model B

feature, it impacted for some cases. This appeared in the long tail that distrusted in both directions.

5.5 Local explainability

In this section, we utilized SHAP plots to show explanation for the output decision for each case (local explainability). Figure 8 shows a GDM case with probability of 73% to have GDM. It also shown the most impact feature values that move the result toward positive class such as gestenail age = 19, albumin = 56.75, Neutrophil = 96.47 and other factors that move the decision to not have GDM include weight change = 2.31.

The all-previous mentioned abbreviations are listed as shown in Table 6.

Fig. 7 Global explainability of proposed DL model

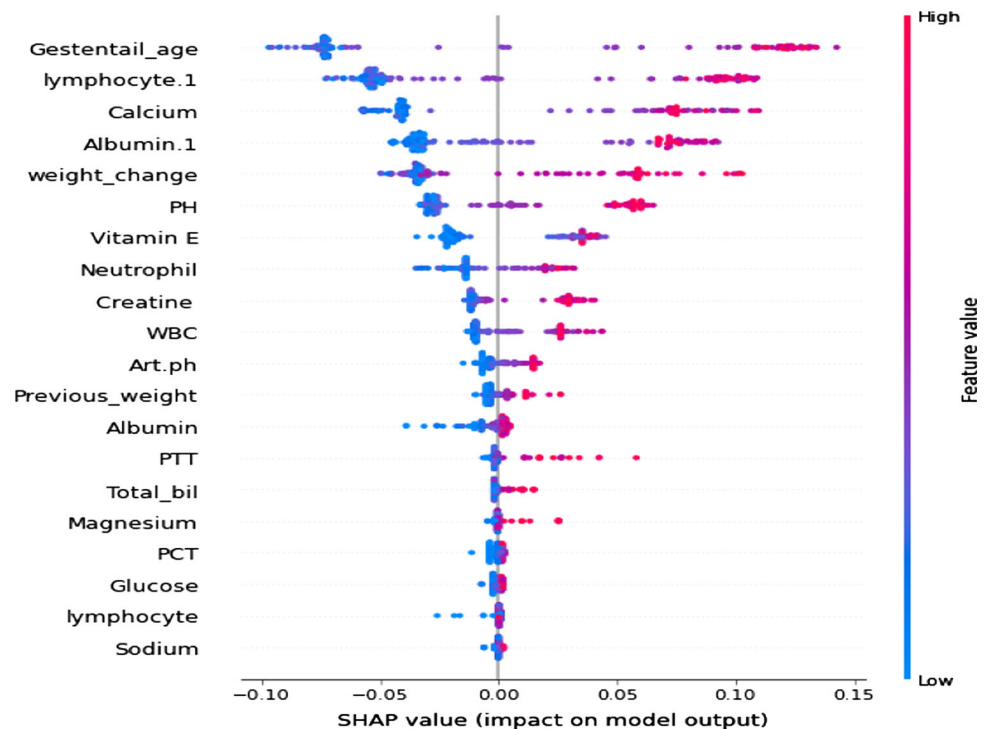
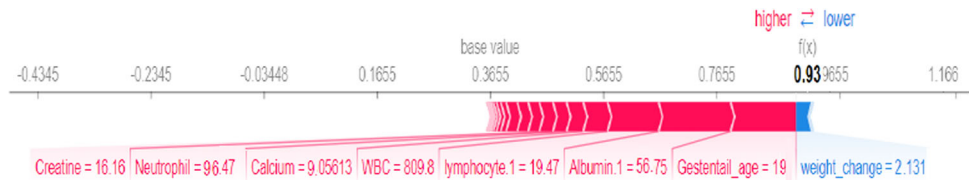


Fig. 8 Local explainability of proposed DL model**Table 6** List of abbreviations

Term	Description
GDM	Gestational diabetes mellitus
DL	Deep learning
AUC	Area under the roc curve
PHR	Patient health record
HER	Electronic health record
PA	Predictive analytics
PNN	Probabilistic neural networks
FC	Fog computing
CC	Cloud computing
IOT	Internet of things

Table 7 The Performance metrics to evaluate the proposed DSP scheme

Metric	Definition
Hit Ratio (HR)	The number of successful requests per time interval T to the total number of requests during the same time interval
Access Latency	Defined as the average packet delay over a multi-hop route. It is used as a measure of the accessibility of the nodes
Power Consumption	The average power consumed over a multi-hop route

5.6 The performance metrics for DFM

The common performance metrics which are used to measure the performance of the caching schemes are: (i) Hit Ratio (HR), (ii) access latency, and (iii) power consumption. Table 7 summarizes the definitions of the performance metrics.

Assume the four data items located at the DCT have the parameters values shown in Table 8. And a new incoming data item ($d_{i_{new}}$) needs to be located at the DCT. The size of $d_{i_{new}}$ is 0.204 MB.

The performance of DFM comparing with the top state-of-the-art caching strategy is shown in Table 9.

From Table 8, it is shown that DFM has achieved the highest HR, the lowest access latency, and the lowest power consumption due to the high accuracy of using the PNN.

6 Study strengths and limitations

6.1 Study strengths

The key strengths of the present study could be summarized in the following points. *First*, the methodology used to predict the GDM and the clinical criteria for choosing the appropriate features for prediction, including vital signs, laboratory tests and biomedical abnormalities that contribute to increase the predication accuracy. *Second*, Utilizing IoT sensors, fog, and cloud computing provide real time system for GDM monitoring with low latency. Utilizing SHAP library to explain the decision of the model and determine the effect of each feature. These contributions provide significant step over state of the art.

6.2 Study limitations

Although our proposed model adds promising achievements to sepsis prediction, we still have many limitations that need further handling. *First*, because MIMIC III dataset is extracted from one institution, we cannot claim the generalization of our results. *Second*, predicting GDM may fail due to the unavailability of the required features. *Third*, the imputation process in which we average all measurement per hour may lead to loss of some temporal values which may negatively affect model performance. Therefore, we intend to work on better imputation techniques that could capture this missing data is a main point of future exploration. Fourth, summarizing time series data and working with feed forward neural networks may discard many temporal features in these multivariate series. Utilizing other deep learning models such as LSTM and CNN are expected to improve the performance. These limitations will be handled in our future studies.

Table 8 Four data items parameters located at DCT

Data item (di)	Access time (TA)	Size of data (S)	Access frequency (FA)	Access count (AC)	Time-to-live (TTL)	Energy (E)	Distance (D)	Cache free size (CFS)
d1	0.936	0.213	0.769	0.0455	0.245	0.0818	3	0.120
d2	0.5	0.0833	0.324	0.264	0.748	0.136	2	0.0005
d3	0.155	0.824	0.287	0.973	0.985	0.791	5	0.610
d4	0.5	0.62	0.528	0.973	0.145	0.482	2	0.487

Table 9 Comparing of DFM with the top state-of-the-art caching strategy

Algorithm	Year	Hit ratio (HR) %	Access latency	Power consumption
DFM	2022	92.11%	1.007	3747
DSP [73]	2019	91.54%	1.541	3815

7 Comparison with literature

As shown in Table 10, we make a comparison of our model with the recent literature according to different criteria such as number of weeks, number of features, and model performance. Note that we choose to compare with studies that not only depend only on the MIMIC dataset due to the small number of studies that used mimic III dataset. As show in Table 9, most of the-state-the-art [63, 76–78] depend on large number of features in predicting GDM. Even though they achieved adequate results, these studies considered not medically accepted. They did not take in consideration the unavailability of such features in most cases. Furthermore, most studies depend on the data aggregated from the first (18–22) weeks to predict gestation typically diagnosed at 24–28 weeks of gestation, but earlier detection is desirable as this may prevent or considerably reduce the risk of adverse pregnancy outcomes. Therefore, in our study we depend on the data aggregated from the first 12 weeks to make the prediction. Compared

with [79], this study depend on small number of features (12 features) and large sample of 6092 women. However, they achieved result 0.750 in terms of AUC. This returns to depending on logistic regression model that that could not consider the changes in several important features. The same in [80], this study also depended on logistic regression algorithm with sample size of 6444 patients, result in 0.721 in terms of AUC. In [62] authors use DL model techniques for predicting, result in AUC = 0.889, 0.849, respectively, but these studies neglected the role of time in predicting gestation and its affect in patient progression. As in our study, authors in [81] DL to predict GDM using only data from the first 14 week. unless their acceptable result AUC = 0.880, the chosen features considered insufficient in the medical domain. One of the strengths of our proposed model is the ability to predict GDM using only the data from the first 12 week of gestation, Result in the most superior model over the state-of-the-art (AUC = 0.906). Overall, our models not only allow early-stage intervention in high-risk women, but also provide a cost-effective

Table 10 Comparison with other work

#	Sample size	Data source	Features	Nu of weeks	AUC	Accuracy
[76]	588,622	Pregnancies in Israel	134 features	20	0.68	0.67
[77]	487	Women's Hospital, School of Medicine,	52 features	22	0.727	69.9
[78]	1385		71 features	21	0.728	0.716
[63]	317	UIC Department of Medicine	111 features	16	0.82	0.68
[82]	1000	Provincial Administration of Traditional Chinese Medicine	83 features	22	0.825	0.809
[79]	6092	university-affiliated tertiary medical center between 2007 to 2014	12 features	20	0.74	0.85
[80]	6444	collected longitudinally using self-administered questionnaires	23 feature	21	0.710	0.712
[62]	25	medical centers and maternity hospitals		22	0.889	0.849
Our work	8740 pregnant	MIMIC III dataset	22		0.942	0.957

screening approach that could avoid the need for glucose tolerance tests. Future prospective studies and studies on additional populations are needed to assess the real-world clinical utility of the model.

8 Conclusion

In this study, we provide comprehensive framework for pregnancy women monitoring. The proposed Data Replacement and Prediction Framework (DRPF) consists of three layers which are: (i) IoT Layer, (ii) Fog Layer, and (iii) Cloud Layer. The first layer used IOT sensors to aggregate vital sings from pregnancies using invasive and noninvasive sensors. Then the vital signs transmitted to fog nodes to processed and finally stored in the cloud layer. The main contribution in this paper is located in the fog layer producing GDM module to implement two influential tasks which are: (i) Data Finding Methodology (DFM), and (ii) Explainable Prediction Algorithm (EPM) using DNN. First, the DFM is used to replace the unused data to free the cache space for the new incoming data items. The cache replacement is very important in the case of healthcare

system as the incoming vital signs are frequent and must be replaced continuously. Second, the EPM is used to predict the incidence of GDM that may occur in the second trimester of the pregnancy. The first DL model (model A) based on vital signs, laboratory tests, and patient's demographics. The second DL model (model B) used the same features, in addition to other pregnant features include weight change, gestenail age, Lymphocyte, Sodium, Vitamin E, Neutrophil, etc. Our study reported that patient's age, BMI, blood pressure, Lymphocyte vitamin E are mainly associated with GDM diagnosing. The proposed model achieves accurate and promising result for academic perspective. However, we are still need to close from real-world secaris. Therefore, in future, we intend to apply our model on a large scale of pregnant patients to ensure the generalization ability of our study.

Appendix

Appendix a

Step 1 – Data collection

1.1 collected data = Querying (MIMIC III)

Step 2- Data Preprocessing

2.1: RawDataset = ReadDataset(Path(collected data))

2.2: procedure Cleaning Preprocessing(RawDataset)

2.3: Features, Labels = split dataset to features and labels

2.4: Encode(Labels)

: dataset = Replace(Features,"?",None)

: for each column col in dataset do

: if nullvalues \geq 50% then

: delete(col)

: end if

: end for

: for each row r in dataset do

: for each value v in r do

: if v = null then

: delete(r)

: end if

: end for

: end for

: for each data instance in dataset do

: cast into integer value

: end for

: scaling(dataset)

: split dataset to training and testing subsets using 70:30 percentage

Step 3- feature extraction (Extract the average the feature value per day)

3.1 extract the required features from the collected dataset

3.2 for each data instance in dataset N do

 For each column S do

 Average(S for each N)

Step 4 deep neural network building

4.1 providing extracted features

4.2 add dense layer_1 (conv layer)

4.3 add dropout layer_1

Step 5 model learning

5.1 learn on training data (70% of the total data)

5.2 test the model on test data (30% of the total data)

Step 6 model evaluation

6.1 calculate accuracy metric

6.2 calculate F1 score metric

6.3 calculate precision metric

6.4 calculate recall metric

Step 7 interpreting deep neural network model and results

7.1 provide local explainability

7.2 provide global explainability

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

Data availability The data used in this paper are publicly accessible after agreement, and details are available in the references section.

Declarations

Conflict of interest The authors declared that they have no conflict of interest.

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

- Risk P, Monitoring A, Prams S, Desisto CL, Kim SY, Sharma AJ (2014) "Prevalence Estimates of Gestational Diabetes Mellitus in the United States, Prevalence Estimates of Gestational Diabetes Mellitus in the United States , Pregnancy Risk Assessment Monitoring System (PRAMS), 2007–2010," no. September, pp 2007–2010, doi: <https://doi.org/10.5888/pcd11.130415>
- Xia Y and Gerber BS (2016) "Diabetes Screening in US Women With a History of Gestational Diabetes, National Health and Nutrition Examination Survey," pp 1–11
- Zhu Y, Zhang C (2016) Prevalence of gestational diabetes and risk of progression to type 2 diabetes: a global perspective. *Curr Diabetes Rep* 16(1):1–11. <https://doi.org/10.1007/s11892-015-0699-x>
- Egan AM, Enninga EAL, Alrahmani L, Weaver AL, Sarras MP, Ruano R (2021) Recurrent gestational diabetes mellitus: a narrative review and single-center experience
- Christophi C, Dabelea D, Pi-sunyer X (2008) Prevention of diabetes in women with a history of gestational diabetes: effects of metformin and lifestyle interventions. *J Clin Endocrinol Metabol* 93(12):4774–4779. <https://doi.org/10.1210/jc.2008-0772>
- Aroda VR et al., "The effect of lifestyle intervention and metformin on preventing or delaying diabetes among women with and without gestational diabetes: the diabetes prevention program outcomes study 10-Year Follow-Up," 100, 1646–1653, 2015, <https://doi.org/10.1210/jc.2014-3761>
- Kent D (2015) "Improving diabetes prevention with benefit based tailored treatment: risk based reanalysis of Diabetes," no. February, 2015, doi: <https://doi.org/10.1136/bmj.h454>
- Herman WH et al (2017) Impact of lifestyle and metformin interventions on the risk of progression to diabetes and regression to normal glucose regulation in overweight or obese people with impaired glucose regulation. *Diabetes Care* 40(12):1668–1677. <https://doi.org/10.2337/dc17-1116>
- Intervention L and Metformin OR (2006) "NIH Public Access," 346(6), 393–403
- Mathur R, Mary Q, Meads C, Greenhalgh T (2011) Risk models and scores for type 2 diabetes: systematic. *BMJ*. <https://doi.org/10.1136/bmj.d7163>
- Bang H, Edwards AM, Bomback AS, Ballantyne CM, Brillon D (2009) Annals of internal medicine article development sand validation of a patient self-assessment score for diabetes risk. *Ann Intern Med*. <https://doi.org/10.1059/0003-4819-151-11-200912010-00005>
- Köhler ABM, Ziegler AG, and Institute, Development of a simple tool to predict the risk of postpartum diabetes in women with gestational diabetes mellitus, no. 0, pp 1–15
- Lindström J et al (2008) Determinants for the effectiveness of lifestyle intervention in the finnish Diabetes prevention study. *Diabetes Care* 31(5):857–862. <https://doi.org/10.2337/dc07-2162>
- Costa B, Barrio F, Cabré JJ, Piñol JL, Cos X (2012) Delaying progression to type 2 diabetes among high-risk Spanish individuals is feasible in real-life primary healthcare settings using intensive lifestyle intervention international federation of clinical chemistry. *Diabetologia* 55(5):1319–1328. <https://doi.org/10.1007/s00125-012-2492-6>
- Ekelund M, Shaat N, Berntorp K (2009) Prediction of postpartum diabetes in women with gestational diabetes mellitus. *Diabetologia*. <https://doi.org/10.1007/s00125-009-1621-3>
- Ignell C, Ekelund M, Anderberg E, Berntorp K (2016) Model for individual prediction of diabetes up to 5 years after gestational diabetes mellitus. Springerplus. <https://doi.org/10.1186/s40064-016-1953-7>
- Awad A, Bader-El-Den M, McNicholas J, Briggs J (2017) Early hospital mortality prediction of intensive care unit patients using an ensemble learning approach. *Int J Med Inform* 108(October):185–195. <https://doi.org/10.1016/j.ijmedinf.2017.10.002>
- 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* 8:133541–133564. <https://doi.org/10.1109/ACCESS.2020.3010556>
- Adams RP et al (2015) A physiological time series dynamics-based approach to patient monitoring and outcome prediction. *IEEE J Biomed Heal Informatics* 19(3):1068–1076. <https://doi.org/10.1109/JBHI.2014.2330827.A>
- Calvert J et al (2016) High-performance detection and early prediction of septic shock for alcohol-use disorder patients. *Ann Med Surg* 8:50–55. <https://doi.org/10.1016/j.amsu.2016.04.023>
- El-Rashidy N et al (2021) Sepsis prediction in intensive care unit based on genetic feature optimization and stacked deep ensemble learning, vol 1. Springer, London
- Forkan ARM and Khalil I (2016) A probabilistic model for early prediction of abnormal clinical events using vital sign correlations in home-based monitoring, doi: <https://doi.org/10.1109/PERCOM.2016.7456519>
- Zhang Y et al (2020) A mid-pregnancy risk prediction model for gestational diabetes mellitus based on the maternal status in combination with ultrasound and serological findings. *Exp Ther Med* 20(1):293–300. <https://doi.org/10.3892/etm.2020.8690>
- Ahmadi M, Mirbagheri E (2019) Designing data elements and minimum data set (MDS) for creating the registry of patients with gestational diabetes mellitus. *J Med Life* 12(2):160–167. <https://doi.org/10.25122/jml-2019-0011>
- Burlina S, Dalfrà MG, Chilelli NC, Lapolla A (2016) Gestational diabetes mellitus and future cardiovascular risk: an update, *Int J Endocrinol* vol. 2016
- Mohammadbeigi A, Farhadifar F, Soufi N, Mohammadsalehi N, Rezaiee M, Aghaei M (2013) Fetal macrosomia: risk factors,

- maternal, and perinatal outcome. *Ann Med Health Sci Res.* <https://doi.org/10.4103/2141-9248.122098>
27. Savvidou M, Nelson SM, Makgoba M, Messow C (2010) First-trimester prediction of gestational diabetes mellitus: examining the potential of combining maternal characteristics and laboratory measures. *Diabetes* 59(12):3017–3022. <https://doi.org/10.2337/db10-0688.N.S0>
 28. Zheng T et al. (2019) A simple model to predict risk of gestational diabetes mellitus from 8 to 20 weeks of gestation in Chinese women, 8, 1–11
 29. Zheng T et al. (2019) A simple model to predict risk of gestational diabetes mellitus from 8 to 20 weeks of gestation in Chinese women, 8, 1–10
 30. Shen J, Chen J, Zheng Z, Zheng J, Liu Z (2020) An innovative artificial intelligence—Based App for the diagnosis of gestational diabetes mellitus (GDM-AI): development study corresponding author. *J Med Internet Res* 22:1–11. <https://doi.org/10.2196/21573>
 31. Of S and Care diabetes M (2018) Standards of medical care in diabetes—2018,” vol. 41
 32. Qiu H et al (2017) Electronic health record driven prediction for gestational diabetes mellitus in early pregnancy. *Sci Rep* 7(1):16417. <https://doi.org/10.1038/s41598-017-16665-y>
 33. Wu Y et al (2021) Early prediction of gestational diabetes mellitus in the Chinese population via advanced machine learning. *J Clin Endocrinol Metabol* 106(3):1191–1205. <https://doi.org/10.1210/clinem/dgaa899>
 34. Nuzzo AM et al (2021) Placental and maternal sFlt1/PlGF expression in gestational diabetes mellitus. *Sci Rep.* <https://doi.org/10.1038/s41598-021-81785-5>
 35. Meertens LJE et al (2020) External validation and clinical utility of prognostic prediction models for gestational diabetes mellitus: a prospective cohort study. *Acta Obstetrica et Gynecologica Scandinavica* 99:891–900. <https://doi.org/10.1111/aogs.13811>
 36. Gracia VD, Olmedo J (2017) Diabetes gestacional: conceptos actuales. *Ginecología y Obstetricia de México* 85(6):380–390
 37. Shen H, Liu X, Chen Y, He B, Cheng W (2016) Associations of lipid levels during gestation with hypertensive disorders of pregnancy and gestational diabetes mellitus: a prospective longitudinal cohort study. *BMJ Open.* <https://doi.org/10.1136/bmjopen-2016-013509>
 38. Ramos G, Borges C, Figueiroa N, Alves LV, Alves JG (2019) Physical activity pattern in early pregnancy and gestational diabetes mellitus risk among low-income women: a prospective cross-sectional study. *SAGE Open Med.* <https://doi.org/10.1177/2050312119875922>
 39. Talaat FM, Ali SH, Saleh AI, Ali HA (2019) Effective load balancing strategy (ELBS) for real-time fog computing environment using fuzzy and probabilistic neural networks, vol. 27, no. 4. Springer US
 40. Atlam HF (2018) Fog computing and the internet of things: a review, pp 1–18, doi: <https://doi.org/10.3390/bdcc2020010>
 41. Bonomi F, Milito R, Zhu J, Addepalli S (2012) Fog computing and its role in the internet of things characterization of fog computing, pp 13–15
 42. Talaat FM, Saraya MS, Saleh AI, Ali HA, Ali SH (2020) A load balancing and optimization strategy (LBOS) using reinforcement learning in fog computing environment. *J Ambient Intell Humaniz Comput* 11(11):4951–4966. <https://doi.org/10.1007/s12652-020-01768-8>
 43. Karthikeyan B (2008) Partial discharge pattern classification using composite versions of probabilistic neural network inference engine. *Expert Syst Appl* 34:1938–1947. <https://doi.org/10.1016/j.eswa.2007.02.005>
 44. Venkatesh S, Gopal S (2011) Expert systems with applications robust heteroscedastic probabilistic neural network for multiple source partial discharge pattern recognition—Significance of outliers on classification capability. *Expert Syst Appl* 38(9):11501–11514. <https://doi.org/10.1016/j.eswa.2011.03.026>
 45. Vellido A (2019) The importance of interpretability and visualization in machine learning for applications in medicine and health care. *Neural Comput Appl.* <https://doi.org/10.1007/s00521-019-04051-w>
 46. 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>
 47. El-Sappagh S, Alonso JM, Islam SMR, Sultan AM, Kwak KS (2021) A multilayer multimodal detection and prediction model based on explainable artificial intelligence for Alzheimer’s disease. *Sci Rep* 11(1):1–27. <https://doi.org/10.1038/s41598-021-82098-3>
 48. Aazam M, Hung PP, Huh E (2014) Smart gateway based communication for cloud of things,” no. October 2015, doi: <https://doi.org/10.1109/ISSNIP.2014.6827673>
 49. Habibi P, Farhoudi M, Kazemian S, Khorsandi S, Leon-garcia A (2020) Fog computing: a comprehensive architectural survey. *IEEE Access.* <https://doi.org/10.1109/ACCESS.2020.2983253>
 50. Ali SH, Saleh AI, Ali HA (2020) Effective cache replacement strategy (ECRS) for real-time fog computing environment. *Cluster Comput* 23(4):3309–3333. <https://doi.org/10.1007/s10586-020-03089-z>
 51. Elhayatmy G, Dey N, Ashour AS (2018) Internet of things and big data analytics toward next-generation intelligence, 30, 3–20, doi: <https://doi.org/10.1007/978-3-319-60435-0>
 52. Pop P, Zarrin B, Barzegaran M, Schulte S (2020) The FORA fog computing platform for industrial IoT ★
 53. Rahul S and Aron R (2021) Fog computing architecture, application and resource allocation: a review, 4638, 0–2
 54. Verma P, Sood SK (2018) Fog assisted- IoT enabled patient health monitoring in smart homes. *IEEE Internet Things J.* <https://doi.org/10.1109/JIOT.2018.2803201>
 55. Khaloufi H, Jadida E, Abouelmehdi K (2020) Fog computing for smart healthcare data analytics: an urgent necessity fog computing for smart healthcare data analytics: an urgent necessity, no. March, doi: <https://doi.org/10.1145/3386723.3387861>
 56. Ahmad M, Bilal M, Hussain S, Ho B, Cheong T, Lee S (2016) Health Fog: a novel framework for health and wellness applications. *J Supercomput* 72(10):3677–3695. <https://doi.org/10.1007/s11227-016-1634-x>
 57. Science B (2020) Development of edge-IoMT computing architecture for smart healthcare monitoring platform, doi: <https://doi.org/10.1109/ISMSIT50672.2020.9254501>
 58. Yi S, Li C, Li Q (2015) A Survey of Fog Computing, In: *Proceedings of the 2015 workshop on mobile big data—Mobidata ’15*, pp 37–42, doi: <https://doi.org/10.1145/2757384.2757397>
 59. Abdelmoneem RM, Benslimane A, Shaaban E (2020) Mobility-aware task scheduling in cloud-fog IoT-based healthcare architectures. *Comput Netw* 179:107348
 60. Nasralla MM (2021) Sustainable virtual reality patient rehabilitation systems with IoT sensors using virtual smart cities
 61. J. Wang et al., Machine learning approaches for early prediction of gestational diabetes mellitus based on prospective cohort study, pp 1–14, [Online]. Available: <https://doi.org/10.21203/rs.3.rs-508626/v1>
 62. Sumathi A, Meganathan S (2022) Ensemble classifier technique to predict gestational diabetes mellitus (GDM). *Comput Syst Sci Eng* 40(1):313–325. <https://doi.org/10.32604/CSSE.2022.017484>
 63. Man B, Schwartz A, Pugach O, Xia Y, Gerber B (2021) A clinical diabetes risk prediction model for prediabetic women with prior gestational diabetes. *PloS One* 16(6):1–14. <https://doi.org/10.1371/journal.pone.0252501>

64. Guo F, Yang S, Zhang Y, Yang X, Zhang C, Fan J (2020) Nomogram for prediction of gestational diabetes mellitus in urban, Chinese, pregnant women, pp 1–9
65. Li D-C, Liu C-W, Hu SC (2010) A learning method for the class imbalance problem with medical data sets. *Comput Biol Med* 40(5):509–518. <https://doi.org/10.1016/j.combiomed.2010.03.005>
66. Mao W, Liu Y, Ding L, Li Y (2019) Imbalanced fault diagnosis of rolling bearing based on generative adversarial network: a comparative study. *IEEE Access* 7:9515–9530. <https://doi.org/10.1109/ACCESS.2018.2890693>
67. Kaur P and Gosain A (2018) Comparing the behavior of over-sampling and undersampling approach of class imbalance learning by combining class imbalance problem with noise comparing the behavior of oversampling and undersampling approach of class imbalance learning by combining class , no. January, doi: <https://doi.org/10.1007/978-981-10-6602-3>
68. Moon TK (1996) The expectation-maximization algorithm. *IEEE Signal Process Mag* 13(6):47–60. <https://doi.org/10.1109/79.543975>
69. Joenssen DW and Bankhofer U (2015) Hot deck methods for imputing missing data hot deck methods for imputing missing data the effects of limiting donor usage, no. July 2012, doi: <https://doi.org/10.1007/0097836.4231.53746>
70. Rachkidi E et al. (2015) Towards efficient automatic scaling and adaptive cost-optimized ehealth services in cloud, In: 2015 IEEE global communications conference GLOBECOM 2015, <https://doi.org/10.1109/GLOCOM.2014.7417751>
71. Smieja M, Struski Ł, Tabor J, Zielinski B, Spurek P (2018) Processing of missing data by neural networks, In: *Advances in neural information processing systems*, 2018, vol. 2018-Decem, no. Section 4, pp 2719–2729
72. Amari S (1993) Backpropagation and stochastic gradient descent method. *Neurocomputing* 5(4):185–196. [https://doi.org/10.1016/0925-2312\(93\)90006-0](https://doi.org/10.1016/0925-2312(93)90006-0)
73. Khan IU et al (2021) Computational Intelligence-based model for mortality rate prediction in COVID-19 patients. *Int J Environ Res Public Health*. <https://doi.org/10.3390/ijerph18126429>
74. Deñ J (2006) Statistical comparisons of classifiers over multiple data sets, 7, 1–30
75. Friedman M (1990) The use of ranks to avoid the assumption of normality implicit in the analysis of variance. *J Am Stat Assoc* 32(200):675–701. <https://doi.org/10.1080/01621459.1937.10503522>
76. Artzi NS et al (2020) Prediction of gestational diabetes based on nationwide electronic health records. *Nat Med* 26(1):71–76. <https://doi.org/10.1038/s41591-019-0724-8>
77. Wu Q et al (2022) An early prediction model for gestational diabetes mellitus based on genetic variants and clinical characteristics in China. *Diabetol Metab Syndr* 14(1):1–10. <https://doi.org/10.1186/s13098-022-00788-y>
78. Zhang X et al (2020) Risk prediction model of gestational diabetes mellitus based on nomogram in a Chinese population cohort study. *Sci Rep* 10(1):1–7. <https://doi.org/10.1038/s41598-020-78164-x>
79. Hourri O et al (2022) Prediction of type 2 diabetes mellitus according to glucose metabolism patterns in pregnancy using a novel machine learning algorithm. *J Med Biol Eng* 42(1):138–144. <https://doi.org/10.1007/s40846-022-00685-9>
80. Gao S et al (2020) Development and validation of an early pregnancy risk score for the prediction of gestational diabetes mellitus in Chinese pregnant women. *BMJ Open Diabetes Res Care*. <https://doi.org/10.1136/bmjdr-2019-000909>
81. Mitra A and Ashraf K (2018) Sepsis prediction and vital signs ranking in intensive care unit patients, [Online]. Available: <http://arxiv.org/abs/1812.06686>
82. Liu Y, Yu Z, Sun H (2021) Prediction method of gestational diabetes based on electronic medical record data. *J Healthc Eng*. <https://doi.org/10.1155/2021/6672072>

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

Authors and Affiliations

Nora El-Rashidy¹ · Nesma E. ElSayed^{2,5} · Amir El-Ghamry^{2,3,4} · Fatma M. Talaat¹

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

Nesma E. ElSayed
Nesma.e@dhic.edu.eg

Amir El-Ghamry
Amir_nabil@mans.edu.eg

Fatma M. Talaat
fatma.nada@ai.kfs.edu.eg

² Computer Science Department, Faculty of Computers and Information, Mansoura University, Mansoura, Egypt

³ School of Engineering and Computer Science, University of Hertfordshire, Egypt, hosted by Global Academic Foundation, Cairo, Egypt

⁴ Faculty of Computer Science and Engineering, New Mansoura University, New Mansoura, Egypt

⁵ Department of computer science, Faculty of Computers and Artificial Intelligence, Damietta University, Damietta, Egypt

¹ Machine Learning and Information Retrieval Department, Faculty of Artificial Intelligence, Kafrelsheikh University, Kafrelsheikh, Egypt