HCIS | Human-centric Computing and Information Sciences

RESEARCH (Research Manuscript) Open Access

Privacy Preservation in Artificial Intelligence-Enabled Healthcare Analytics

Shancang Li[1*], Muddesar Iqbal[2], Ali Kashif Bashir[3,5,6] and Xinheng Wang[4]

Abstract

Emerging techniques such as the Internet of Things (IoT), machine learning, and artificial intelligence (AI) have revolutionized healthcare analytics by offering a multitude of significant benefits, including real-time process, enhanced data efficiency and optimization, enabling offline operation, fostering resilience, personalized and context-aware healthcare, etc. However, privacy concerns are indeed significant when it comes to edge computing and machine learning-enabled healthcare analytics. The training and validation of AI algorithms face considerable obstacles due to privacy concerns and stringent legal and ethical requirements associated with datasets. This work has proposed a healthcare data anonymization framework to address privacy concerns and ensure compliance with data regulations by enhancing privacy protection and anonymizing sensitive information in healthcare analytics, which can maintain a high level of privacy while minimizing any adverse effects on the analytics models. The experimental results have unequivocally showcased the effectiveness of the proposed solution.

Keywords

Data security, privacy preserving, artificial intelligence, healthcare analytics

1. Introduction

As the digitization of healthcare continues to accelerate, concerns regarding the privacy and security of sensitive patient data have come to the fore [1, 2]. The emerging techniques, such as machine learning (ML), artificial intelligence (AI), and Internet of Things (IoT), have shown great potential in healthcare analytics. Using enormous amounts of historic and real-time data (medical records, etc. [3, 4]), these techniques can help healthcare analytics algorithms create valuable clinical forecasts, predictions, and recommendations for both personalized care and public health, which can significantly benefit healthcare analytics in predictive analytics, clinical decision support systems, real-time monitoring and alerts, operational analytics, healthcare fraud detection, *etc.* [5]. Recently, AI technology has shown great potential to improve healthcare applications and patient care using computer vision techniques, deep neural networks, etc. in tumor detection, genomic characterization, tumor subtyping, grading prediction, outcome risk assessment, etc. [6, 7].

In particular, empowered health predictive analytics can do the following: (1) inform and alert clinicians/caregivers about the likelihood of events before they occur; (2) detect early signs of patient deterioration and help proactively intervene at the early stage; (3) identify and predict patients; (4) identify and predict the status of medical systems to minimize unscheduled workflow disruptions. However, the privacy concerns in healthcare analytics needs to be addressed, including patient confidentiality, data security, de-identification and anonymization, data sharing, data governance and compliance, ethical use of data, etc. [8].

***Corresponding Author:** Shancang Li (shancang.li@ieee.org)
[1] School of Computer Science & Informatics, Cardiff University, Cardiff, UK
[2] College of Engineering, Prince Sultan University, Riyadh, SA.
[3] Department of Computing & Mathematics, Manchester Metropolitan University, Manchester, UK.
[4] School of Advanced Technology, Xi'an Jiaotong-Liverpool University, Suzhou, CN.
[5] Woxsen School of Business, Woxsen University, Hyderabad 502 345, India
[6] Department of Computer Science & Mathematics, Lebanese American University, Beirut, Lebanon

However, there are several key challenges that need to be addressed in healthcare analytics, including data privacy and security, secure data access and sharing, data quality and integration, data governance and standards, regulatory compliance, ethical considerations, etc. [9, 10] Emerging privacy-enhancing techniques (PETs) in healthcare analytics have empowered healthcare analytics by enhancing privacy, improving data security, and enabling more accurate and insightful analysis [11]. Techniques like sensitive data anonymization, secure multi-party computation, federated learning, and homomorphic encryption protect patient privacy while allowing collaborative analysis [10]. Blockchain and distributed ledger technologies ensure data integrity and secure sharing, and differential privacy injects statistical noise to preserve individual privacy in aggregated results [12]. These techniques build trust, comply with regulations, and encourage data sharing, resulting in larger and more diverse datasets [13]. By striking a balance between privacy and data utility, these emerging techniques are poised to empower healthcare analytics to unlock valuable insights and improve patient outcomes [14, 15].

Striking a balance between extracting valuable insights from healthcare data and protecting patient privacy is crucial to foster trust, maintain compliance, and ensure the ethical use of data in healthcare analytics [16, 17]. This work focuses on enhancing privacy, improving data security, and enabling more accurate and insightful analysis based on the existing works. A secure and encrypted predictive analytics scheme is proposed to address the above challenges using the above privacy-enhancing techniques. In the context of healthcare, DP can maximize the accuracy of queries from statistical databases and minimize the chances of identifying its entries, especially in publishing and mining medical data. The SMPC allows multiple users to conduct healthcare analysis, *such as* joint modelling, feasibility analysis of clinical research, cohort study with large sample, disease prediction, etc. Also, the FL is a machine learning method that trains algorithms across multiple devices without exchanging samples. In relation, these techniques can be used separately or in combination, depending on the specific scenarios of the healthcare analysis task. The main contributions include the following.

1) This work proposes a privacy-enhancing framework for machine learning-based healthcare analytics, which supports differential privacy, secure multiparty computation, federate learning, and anonymize data.

2) A data anonymization approach is proposed that can anonymize sensitive information before being fed into machine learning-based healthcare analytics without significantly compromising the data utility.

3) A use case of diabetes prediction using machine learning is addressed to validate the proposed framework and approach.

2. Related Work

Privacy-enhancing techniques play a crucial role in safeguarding the confidentiality of healthcare analytics due to key privacy concerns in terms of patient confidentiality, legal and ethical obligations, data minimization, risk of re-identification, trust and patient engagement, discrimination, and stakeholder collaboration, etc. [18, 19]. By prioritizing privacy, healthcare organizations can ensure that analytics initiatives contribute to improved healthcare outcomes and safeguard patient rights and interests.

Recently, researchers have developed various techniques and frameworks for privacy-preserving secure data sharing in different domains, such as blockchain, cloud computing, biomedical research, online social networks, and IoT-enabled healthcare systems [20]. They address the challenges of protecting sensitive data in the course of enabling efficient and controlled sharing among authorized entities. Liang and Lu proposed a privacy-preserving data sharing framework using blockchain technology and smart contracts to ensure data confidentiality and integrity concurrently with allowing controlled access to shared data through smart contract enforcement [21].

Secure multi-party computation (SMPC) has also been used in healthcare analytics to protect user privacy in collaborative analytics scenarios. Ma et al. developed an efficient protocol for privacy-preserving logistic regression using MPC to achieve high accuracy in classification in tandem with ensuring the privacy of individual data during the model training process [22, 23]. Anonymization techniques have been utilized in healthcare analytics to protect patient privacy in the course of maintaining data utility [24, 25].

Sowmiya *et al*. proposed a privacy-preserving data sharing framework for healthcare using k-anonymity and blockchain to ensure that sensitive patient data is shared with a reduced risk of re-identification in tandem with leveraging the immutability and transparency [26, 27, 28].

There are still a few challenges that need to be addressed in healthcare analytics, including balancing privacy and data utility, complying with privacy regulations and legal considerations, imbalanced and incomplete data, etc. Addressing these challenges requires multidisciplinary collaboration, robust data governance frameworks, advanced data preprocessing techniques, and development of interpretable and validated machine learning models.

3. Re-Identification Risks Evaluation

Privacy protection in healthcare analytics is paramount to ensure the confidentiality and security of sensitive patient information, in which data anonymization is one of the critical techniques used to protect patient privacy along with enabling valuable analysis. In machine learning-based analytics, a lack of data used for training and validation due to privacy protection considerations makes the analytics process a highly challenging one.

There are a few secure and private AI technologies to solve problems in the medical imaging field, which can be classified into private AI (*re-identification & differential privacy*) and secure AI (*federated learning, homomorphic encryption and SMPC*).

Conversely, improperly protected healthcare data can lead to re-identification risks, including *privacy breach, identity inference, data linkage, unintended discrimination, regulatory compliance, etc.* Mitigating these risks requires the implementation of robust privacy protection measures, including appropriate anonymization techniques, strict access controls, encryption, and comprehensive data governance frameworks.

It is complicated to assess privacy in real-world systems [29]. In this work, we have assumed that re-identification risk complements the data anonymization procedures (e.g. k-anonymity, differential privacy, *etc.*). For a representation matrix $\mathbf{P} \in [0, 1]^{n \times m}$, it is assumed an attacker will attempt to re-identify users based on representations sampled from $\mathbf{P}$ using a prediction rule φ, which is the probability that an attacker correctly re-identifies the random user conditioned on the representation matrix

$$Acc_{\mathrm{R}}(\varphi_R) = \mathbb{P}(\varphi_R(O) = I|\mathbf{P}) \tag{1}$$

where I denotes a uniformly random sample, φ_R: O → I is a possibly randomized prediction rule, let O be sampled from P[I,:].

As an example, in k-anonymity, given that each user cannot be distinguished from at least $k - 1$ other users, an attack prediction rule φ_R should satisfy the equation below as follows:

$$Acc_{\mathrm{R}}\ (\varphi_R) \leq \frac{1}{k} \tag{2}$$

In data anonymization, asymmetric similarity metric can be used to measure individual attribute similarity for two anonymized records $\mathbf{x}$ and $\mathbf{y}$, expressed mathematically as follows:

$$T(\mathbf{x}(i), \mathbf{y}(i)) = 1 - \frac{|\mathbf{x}(i) - \mathbf{y}(i)|}{p(i)} \tag{3}$$

where $p(i)$ denotes the range of attribute i. The similarity metric can be defined mathematically as follows:

$$S(\mathbf{x}, \mathbf{y}) = \sum_{i \in supp(\mathbf{x})} \frac{T(\mathbf{x}(i), \mathbf{y}(i))}{|supp(\mathbf{x})|} \tag{4}$$

where $|supp(\mathbf{x})|$ denotes non null attributes in $\mathbf{x}$.

3.1 Data anonymization in healthcare analytics

Data anonymization is a critical technique used in healthcare analytics to protect patient privacy in tandem with enabling valuable insights to be derived from healthcare data. Data anonymization techniques includes the following key steps as specified:

Step 1: **Data identification**: Identify the specific attributes within the healthcare dataset that need to be anonymized. These attributes typically include sensitive information such as patient identifiers (e.g. name, address, social security number), medical record numbers, or other personally identifiable information (PII).

Step 2: **Select anonymization algorithm**: Choose an appropriate anonymization algorithm based on the data and privacy requirements. Common techniques include generalization, suppression, pseudonymization, and permutation. More advanced techniques like differential privacy or secure multi-party computation can also be considered depending on the privacy level and data utility required.

Step3: **Re-identification risk assessment**: In this stage, to evaluate the potential re-identification risk associated with the anonymized dataset, we can use matching attack or record linkage analysis to assess the vulnerability of the dataset to re-identification.

Step 4: **Data anonymization**: Implement the selected anonymization techniques on the identified variables, which may involve modifying or removing certain identifiers, replacing identifiers with pseudonyms, or transforming sensitive values to preserve privacy along with retaining data integrity and utility.

Step 5: **Evaluate data utility**: Assess the impact of anonymization on the data utility for analytics purposes. Evaluate whether the anonymized data still retains sufficient quality and accuracy to derive meaningful insights and perform the desired analyses. Balancing privacy protection with data utility is crucial to ensure that anonymized data remains valuable for analytics purposes.

Step 6: **Compliance with regulations**: Ensure compliance with relevant privacy regulations, such as HIPAA or GDPR, applying data anonymization techniques. Understand the legal requirements and guidelines pertaining to data de-identification, and implement necessary measures to adhere to regulatory frameworks.

To evaluate the quality of anonymization, in this work, we use mutual information (MI), which considers the dependencies between data vectors. For data $\mathbf{x}$ and anonymized data $\mathbf{x}'$, we mathematically derive the pertinent equation as follows:

$$MI(\mathbf{x}';\mathbf{x}) = \sum_{x'\in\mathbf{x}'}\sum_{x\in\mathbf{x}} p(x',x)\log\frac{p(x',x)}{p(x')p(x)} \tag{5}$$

where $p(x)$ is the probability of observing x', $p(x)$ is the probability of observing x, and $p(x',x)$ is the probability of observing x' given x mathematically derived as follows:

$$MI(\mathbf{x}';\mathbf{x}) = H(\mathbf{x}) - H(\mathbf{x}|\mathbf{x}') \tag{6}$$

where $H = -\sum_i p_i \, log_2 p_i$ is the entropy contained in each variable. The *MI* score can be calculated by deriving the pertinent equation as follows:

$$MI_{score} = \sum_{i=1}^{N}\sum_{j=1}^{N}\left[\frac{MI\,(x_i,x_j)}{MI(\hat{x_i},\hat{x_j})}\right] \tag{7}$$

where $x_i \in \mathbf{x}$ is the original data and $x\hat{}_i$ is the anonymized data. The success metric can be defined mathematically by deriving the pertinent equation as follows:

$$Acc_{\mathrm{R}}\left(\varphi_R\right) \leq \frac{1+MI(I;O|P)}{log(n)} \tag{8}$$

3.2 Image Anonymization against Re-Identification Risks

Image re-identification can be used in healthcare analytics to enhance the privacy of applications of image re-identification, including *biometric identification, patient monitoring and tracking, clinical trials, medical image analysis, etc.* In this work, we employ a re-identification method focusing on the DICOM dataset IXI of MRI scans by combining FSL library and PyDace to extract and process MRI images of the human brain in the DICOM dataset IXI. This allows a user to extract meaningful information and uncover insights related to brain structure, function, and abnormalities.

1) **DICOM format**. DICOM (**digital imaging & communications in medicine**) is a widely used standard format for storing, transmitting, and sharing medical images and related information in healthcare analytics. DICOM plays a crucial role in facilitating interoperability and seamless integration of medical imaging data across various healthcare systems and applications. The DICOM includes key components like *metadata and structured reporting, integration with healthcare systems, etc.*, and medical images have the four key components of *image depth, photometric representation, metadata, and pixel data*. These components are related to image size and image resolution. The DICOM 3D position can be represented as Eqs. (6) and (7) above [30]

$$\mathbf{V} = \mathbf{IPP} + \mathbf{D} + \mathbf{E} \tag{9}$$

where **IPP** is the **image position patient**, $\mathbf{D} = h_{mm}(v_1 - v_2)$. **IOP2** and $\mathbf{E} = w_{mm}(v_2 - p_2)$.**IOP1**.

$$\begin{bmatrix} v_{\mathrm{L}} \\ v_{\mathrm{P}} \\ v_{\mathrm{H}} \end{bmatrix} = \begin{bmatrix} p_{\mathrm{L}} \\ p_{\mathrm{P}} \\ p_{\mathrm{H}} \end{bmatrix} + \mathrm{h_{mm}}\left(v_1 - p_1\right) \begin{bmatrix} b_{\mathrm{L}} \\ b_{\mathrm{P}} \\ b_{\mathrm{H}} \end{bmatrix} + \mathrm{w_{mm}}\left(v_2 - p_2\right) \begin{bmatrix} a_{\mathrm{L}} \\ a_{\mathrm{P}} \\ a_{\mathrm{H}} \end{bmatrix} \tag{10}$$

Most of today's MRI instruments acquire reconstructed data in DICOM format. DICOM is not only a storage format for images, but also a mode of conversion between different forms of data from different imaging systems, among which MRI images are only one form. Typically, DICOM treats each image layer as a separate file, which includes multiple files, and each of them is named numerically to reflect the corresponding number of image layers (with some variations from system to system). The files contain header information and must be opened by specific software. In all formats, DICOM contains a large amount of metadata information in the file header, including instrument information, image acquisition parameters, and patient information. In the DICOM context, a *transverse plane, coronal plane, sagittal plane* are used to help users communicate the location and orientation of structures within in the body accurately: (a) transverse plane (axial plane) is perpendicular to the longitudinal axis of the body or the image field; (b) coronal plane is perpendicular to both the longitudinal axis and the transverse plane; and (c) sagittal plane is perpendicular to both the longitudinal axis and the transverse plane.

The final image size equation is shown below, where *DH* denotes the size of the data header (including metadata), *R* and *C* denote the number of rows and columns of data, and *PD* denotes the pixel depth (number of image frames), respectively, denoted mathematically as follows.

$$S_{img} = DH + R * C * PD \tag{11}$$

2) **FSL and PyDeface library**. The FSL Library was developed for the analysis of structural MRI, functional MRI (task, resting), diffusion MRI, pre-processing and analysis of MRI and CT data. FSL

also includes an intuitive but powerful 3D/4D image display tool, FSLView, which allows multiple color overlays, multiple orthogonal or lightbox views, time series display (graph-based and cine-based), image editing, vector field diffusion direction display (color coded or arrow) and histogram viewing. PyDeface supports pre-defined face voxel masks that are localized on the input image and removed (set to zero) using linear registration. Fig. 1(a) below indicates original DICOM image slices and Fig. 1(b) below shows the anonymized DICOM image slices, Fig. 1(b) removes key meta data and features that can be used to identify the subject. Eq. (12) below can be used to check the similarity of an original image and anonymized image, mathematically denoted by the pertinent equation as follows:

$$S_{sq} = \sum_{(n,m)\in N^{M\times N}} (J[n,m] - I[n,m])^2 \tag{12}$$

and the cross-correlation can be obtained by using the pertinent equation denoted as follows:

$$C_{crr} = \sum_{(n,m)\in N^{M\times N}} (J[n,m] \times I[n,m])^2 \tag{13}$$

3) **FSL and PyDeface-based DICOM Anonymization**. In this work, we use FSL and PyDeface-anonymized IXI-T1 MRI dataset. Fig. 1(a) and Fig. 1(b) below show the portrait images before and after image anonymization, in which the key human features have been anonymized. Unlike existing methods, which only remove the meta data included in the DICOM files, the proposed methods are able to anonymize the features in the content of the image. Similarly, Fig. 2(a) and Fig. 2(b) below present the original images and anonymized images in the context of DICOM.

Image similarity metric. To compute the similarity of two images I_i, I_j, we can use the sum of Euclidean distance of each histogram, mathematically denoted as follows:

$$D(I_i, I_j) \sum_{i=1}^{N} \frac{\sqrt{\sum_{j=1}^{M} (q_{ij} - s_{ij})^2}}{N} \tag{14}$$

where N is the number of histograms, M is the number of bins per histogram.

3.3 Data Anonymization in ML Model-Based Healthcare Analytics

Machine learning-based data analysis and decision-making techniques have the potential to revolutionize healthcare analytics by providing more efficient, accurate, and cost-effective analytics, especially in predictive healthcare analytics, medical image analysis, etc.

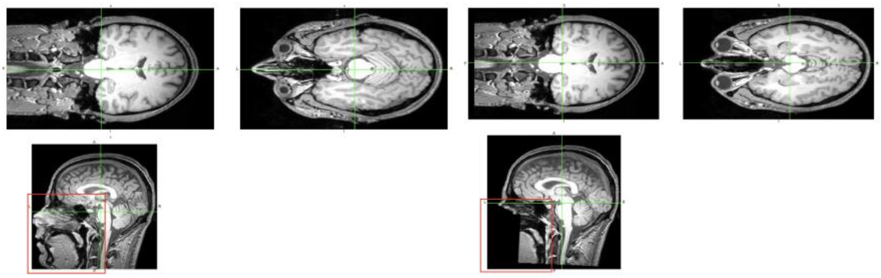
The widespread adoption of machine learning techniques has raised significant concerns regarding data security and privacy. To address these concerns, it is crucial to implement effective strategies that eliminate sensitive data before training machine learning models. This section introduces a comprehensive approach for identifying and anonymizing sensitive data before being fed into machine learning models, which include the following four key steps:

Step 1: Densifying sensitive data: The process of identifying sensitive data is crucial for effective privacy protection. This section discusses various types of sensitive data commonly encountered in ML applications, such as personally identifiable information (PII), medical records, and financial data, and also explores techniques and methodologies for identifying and categorizing sensitive data within a given dataset.

Step 2: **Data anonymization**: Best practices and techniques. To ensure the removal of sensitive data, this section outlines the best practices and techniques for data anonymization and sanitization, while discussing popular Python libraries, such as pandas and scikit-learn, and presenting methods such as data aggregation, de-identification, generalization, and suppression. Additionally, it explores the concept of differential privacy and its application to preserve data privacy along with maintaining statistical accuracy.

Step 3: **Implementation framework and case study**: In this section, we provide a step-by-step implementation framework for removing sensitive data from a dataset using Python libraries. Also, a case study is presented to illustrate the practical application of the proposed techniques, demonstrating their effectiveness in safeguarding privacy and enhancing data security.

Step 4: **Evaluation and performance analysis**: To evaluate the performance of the proposed approach, this section presents a comprehensive analysis of the implemented methods. Key metrics such as data loss, utility preservation, and privacy guarantees are assessed, providing insights into the effectiveness and trade-offs associated with the elimination of sensitive data.



(a) Pre-anonymization

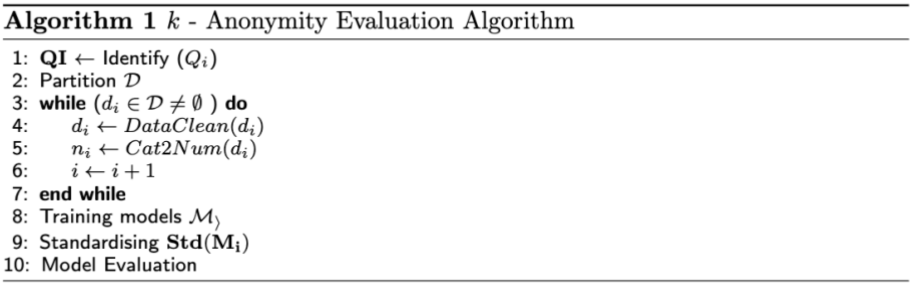
(b) Post-anonymization

Fig. 1. Image of Human Brain before/after Anonymization Processing

4. Diabetes Prediction Using Anonymized Machine Learning

Given that diabetes is a chronic metabolic disorder that affects a significant portion of the global population, early detection and prediction of diabetes can lead to timely interventions and improved patient outcomes. In recent years, ML algorithms have shown promising results in diabetes prediction. In relation, it is important to remove sensitive information in tandem with preserving the utility and integrity of diabetes data for analysis and diabetes prediction model training.

Using different machine learning models on a diabetes prediction dataset [11], such as MLP, LR, RFC, KNN, *etc.* is possible to predict diabetes. In the context of diabetes prediction, sensitive data refers to any information that, if disclosed or accessed by unauthorized individuals, could potentially harm an individual's privacy, lead to discrimination, or compromise their personal well-being.

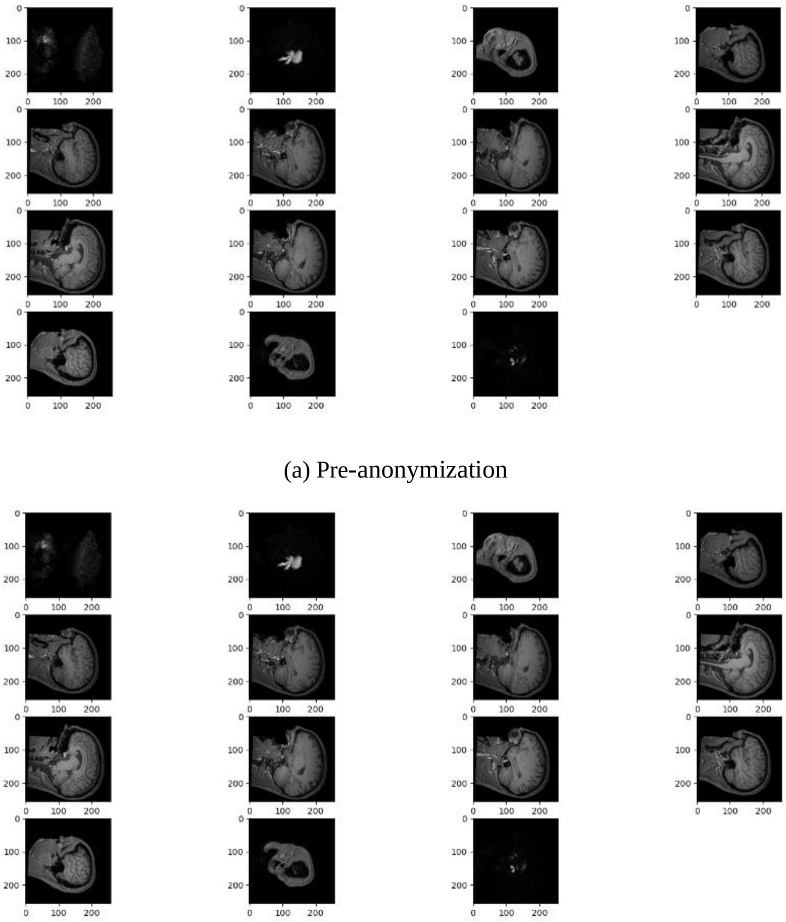


Algorithm 1 k - Anonymity Evaluation Algorithm

```
QI ← Identify ($Q_i$)
Partition $\mathcal{D}$
while ($d_i \in \mathcal{D} \neq \emptyset$ ) do
    $d_i \leftarrow DataClean(d_i)$
    $n_i \leftarrow Cat2Num(d_i)$
    $i \leftarrow i + 1$
end while
Training models $\mathcal{M}_{\rangle}$
Standardising $\mathbf{Std}(\mathbf{M_i})$
Model Evaluation
```

where the *DataClean(dᵢ)* function conducts data cleaning to ensure the accuracy and reliability of the dataset, which focuses on the following processes of identifying/handling missing data, ensuring consistency, dealing with outliers, validating data types, removing duplicated entries, and addressing data entry errors.

In this work, we have developed a machine learning model using TensorFlow to predict whether a person has diabetes. The dataset includes the fields {*name, address, gender, age, hypertension, heart disease, smoking history, bmi, HbA1c level, blood glucose level*}, in which {*name, gender, age*} are key PII that can be targeted by hackers or malicious actors using a 're-identification' attack. In this work, we have used the *k–Anonymity* approach to protect {*name, gender, age*}. *k*-anonymization can help prevent the re-identification of {*name, gender, age*}, which include the following key procedures of *Generalization, Suppression*. The *k* refers to the minimum number of individuals that must be in each group to provide anonymity [31].



(a) Pre-anonymization

(b) Post-anonymization

Fig. 2. DICOM Layers of Human Brain Anonymization Processing

The algorithm includes the following key eight steps:

Step 1: Data anonymization: In this work, we use k-anonymity to anonymize key QIs in the dataset.

Step 2: Partitioning: The dataset will be partitioned into groups of at least *k* individuals such that each group has indistinguishable quasi-identifiers (QIs) [32].

Step 3: Data cleaning: It aims to identify *missing values, remove duplicates, handle outliers, encode labels, etc.*

Step 4: Convert categorical to numeric: It is usually needed to transform the categorical columns into numeric to match machine learning algorithms.

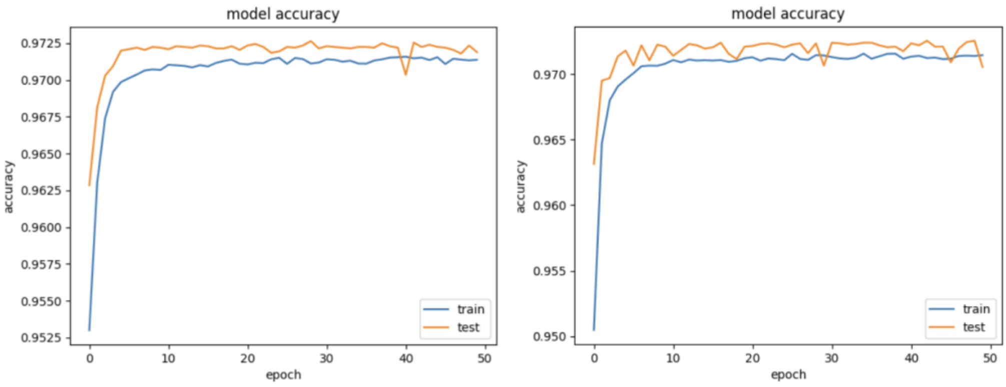
Step 5: Splitting: The dataset needs to be split into training and testing sets. In this work, we use 80% data for training and 20% for testing.

Step 6: Standardizing: It transforms the data so that the mean is 0 and the standard deviation is 1 to ensure that all features have the same scale, which is important for machine learning algorithms.

Step 7: In this work, we created the random forests and artificial neural network with TensorFlow and use the partitioned subsets of the dataset to train the model. The model learns patterns and relationships between the diabetes features and the target labels through an integrative process.

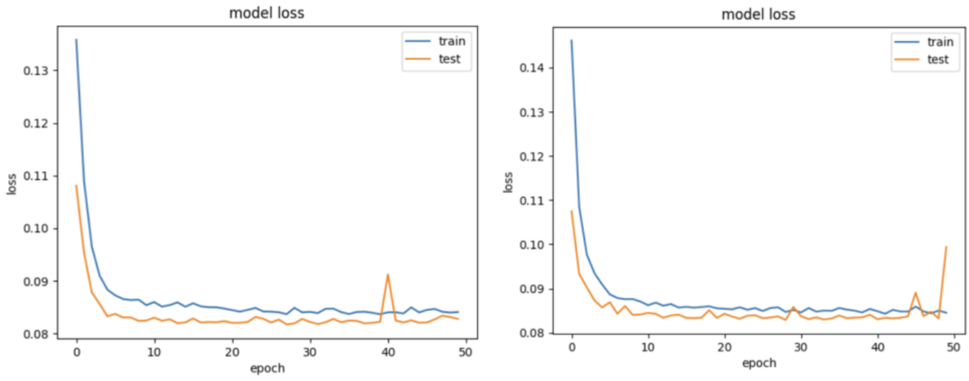
Step 8: Model evaluation: Using the testing subset for the dataset, we can evaluate the trained model with appropriate evaluation metrics, such as accuracy, precision, recall, F_1 score, etc. Figs. 3 and 4 below show the accuracy and loss of trained model when using the anonymized data versus the original dataset.

In practice, it is important to update the model periodically with new data to ensure its accuracy and adaptability to changes in the population or healthcare environment. It can be seen from Figs. 3 (a) and 3(b) below that the model accuracy stays between 97%-97.25% for a dataset without anonymization, while the model accuracy stays above 97.0% for anonymized dataset. This infers that data anonymization does not significantly affect the performance of model accuracy. A similar result can be found for the model loss for both the original dataset and anonymized dataset in Figs. 4(a) and 4(b) below.



(a) Without anonymization (b) With anonymization

Fig. 3. Diabetes Prediction Model Accuracy for Original Dataset vs. Anonymized Dataset



(a) Without anonymization (b) With anonymization

Fig. 4. Diabetes Prediction Model Accuracy for Original Dataset vs. Anonymized.

Numerous machine learning models can be used in healthcare analysis: *Multilayer Perceptron*

(MLP), which is a type of artificial neural network widely used in machine learning. It is a feedforward neural network that consists of multiple layers of interconnected nodes, also known as artificial neurons or perceptron. MLPs are known for their ability to learn complex patterns and make predictions or classifications based on input data.

Logistic regression (LR) is a popular machine learning model used for classification tasks. Despite its name, it is primarily used for binary classification, where the target variable has two classes.

Random forest classifier (RFC) is a popular machine learning algorithm used for classification tasks, belonging to the ensemble learning family of algorithms, which combine the predictions of multiple individual models to make more accurate predictions. RFC is based on decision tree models and provides several benefits over single decision trees.

K-nearest neighbors (KNN) is a popular machine learning algorithm used for both classification and regression tasks, and also is a non-parametric and instance-based algorithm, meaning that it does not make any assumptions about the underlying data distribution and learns directly from the training examples. Fig. 5 below shows the number of components needed to explain variance versus cumulative variance, while Table. 1 below shows the machine learning models in diabetes prediction.

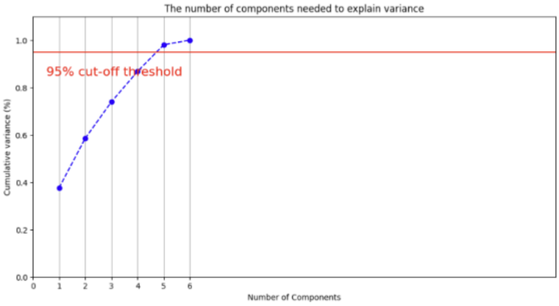
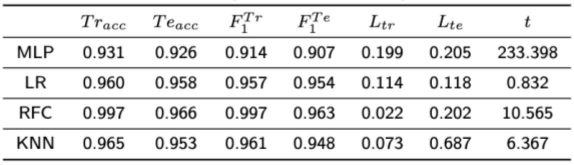


Fig. 5. Number of Components

Table 1: Using machine learning to test

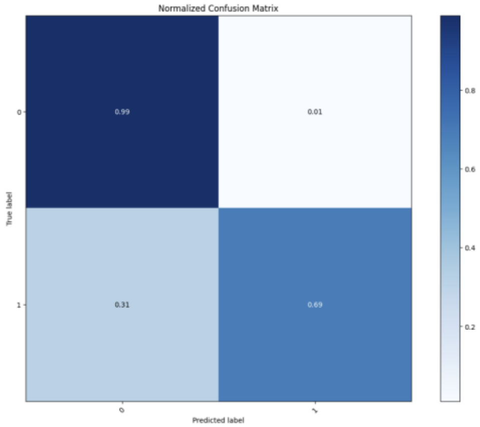


	Tr_{acc}	Te_{acc}	F_1^{Tr}	F_1^{Te}	L_{tr}	L_{te}	t
MLP	0.931	0.926	0.914	0.907	0.199	0.205	233.398
LR	0.960	0.958	0.957	0.954	0.114	0.118	0.832
RFC	0.997	0.966	0.997	0.963	0.022	0.202	10.565
KNN	0.965	0.953	0.961	0.948	0.073	0.687	6.367

The anonymized data needs to be re-standardized with a re-calculated covariance matrix of the features. The number of components determines the amount of information retained from the original dataset. In this work, we use scikit-learn to analyze the anonymized dataset and selected six components in the diabetes analysis, as shown in Figure. 5 above.

The heatmap is a powerful visual tool that presents data patterns, relationships, and distributions in a clear and intuitive manner. By leveraging color gradients, it offers a concise and visually appealing summary of the data to enable effective data exploration, decision-making, and an understanding of the underlying data structure in machine learning-based data analytics. Fig. 6 below shows the heatmap of diabetes analysis on the following features of {*'blood_glucose level', 'HbA1c level', 'age', 'bmi', 'hypertension', 'heart disease'*}.

The confusion matrix is a valuable tool in predictive analytics since it provides a comprehensive evaluation of classification model performance and is especially useful when dealing with binary or multi-class classification problems. The confusion matrix is constructed based on the actual and predicted class labels of a dataset and helps in assessing the model's accuracy, precision, recall, and other important metrics. Fig. 7 below presents the confusion matrix for the relevant features of *'blood glucose _ level', 'HbA1c level', 'age', 'bmi', 'hypertension', 'heart disease'*.



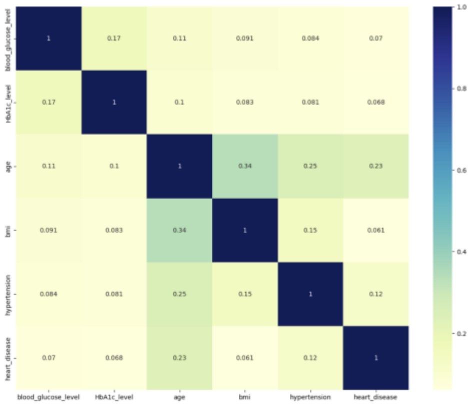


Fig. 6. Heatmap of Key Features

Fig. 7. Confusion Matrix

The precision-recall curve is used in predictive analysis to evaluate the performance of a binary classification model, particularly when dealing with imbalanced datasets. Also, it illustrates the trade-off between precision and recall at various thresholds or decision boundaries set by the model. Fig. 8 below shows the precision-recall curve regarding diabetes prediction.

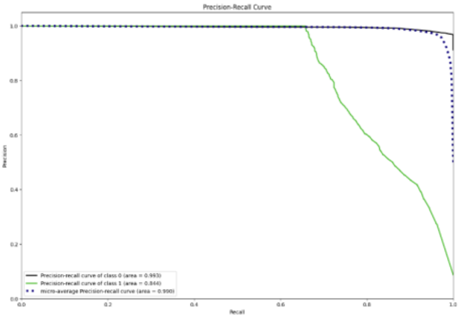


Fig. 8. Precision-Recall Curve of Diabetes Prediction

The feature distributions provide insights into the characteristics and properties of the variables used for prediction. In predictive healthcare analysis, the feature distributions can help in gaining a better understanding of the data, besides enhancing the data pre-processing. In diabetes detection, the feature distribution indicates the dependency on the nature of data and the problem at hand. The choice of feature distribution can significantly impact the performance and interoperability of machine learning models. Figure. 9 below shows the feature distributions in terms of *'blood _glucose _level', 'HbA1c _level', 'age', 'bmi', 'hypertension', 'heart disease'*. In relation, data anonymization should not be able to change the feature distributions. In Figure. 9 below, the blue lines denote the feature distribution in the original dataset, while the red lines denote the feature distribution in the anonymized dataset.

5. Conclusion

In healthcare analytics, PETs play a vital role in healthcare analytics, given that the need to balance data utility and patient privacy is paramount [36-40]. Similarly, advancements in technology and the increasing availability of healthcare data have raised concerns about the potential misuse or unauthorized access to sensitive information. This work focuses on the adoption of data anonymization techniques, such as image anonymization and *k*-anonymity, provision of methods for de-identifying sensitive information, and reduction of the risk of re-identification. In particular, we have proposed a privacy-enhancing framework for a machine learning model-based healthcare analysis and detailed image/data anonymization techniques to anonymize sensitive information before being fed into the model. In this work, we have presented a use case of diabetes prediction using machine learning to validate the proposed framework and approach while addressing re-identification risks. However, the proposed methods can enhance privacy but may lead to the loss of data utility, and in next stage, we plan to focus on addressing the data utility trade-off.

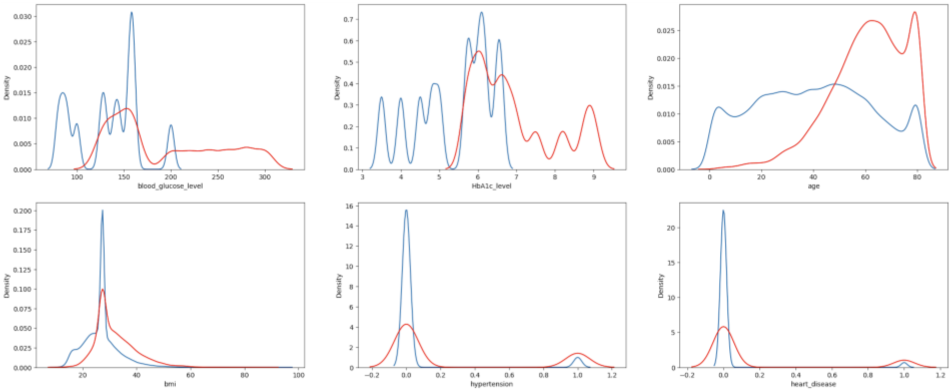


Fig. 9. Feature Distributions

Acknowledgement

This work was partially supported by the research grants from "Unmanned Ground Vehicles for Disaster Management and Recovery (PID-000085_01_03)" at Prince Sultan University.

Author's Contributions

Shancang Li developed the idea and drafted the manuscript. Muddesar Iqbal contributed to conducting experiments and validation. Ali Kashif Bashir contributed to forming the methodology and write-up. Xinheng Wang contributed to forming the conceptual design and validation.

References

[1] Krall, A., Finke, D., Yang, H.: Mosaic privacy-preserving mechanisms for healthcare analytics. IEEE Journal of Biomedical & Health Informatics 25(6), 2184–2192 (2021).

[2] Bugshan, N., Khalil, I., Moustafa, N., Rahman, M.S.: Privacy-preserving microservices in Industrial Internet of Things-driven smart applications. IEEE Internet of Things Journal 10(4), 2821–2831 (2023).

[3] Qi, M., Wang, Z., Han, Q.-L., Zhang, J., Chen, S., Xiang, Y.: Privacy protection for blockchain-based healthcare IoT systems: A survey. IEEE/CAA Journal of Automatica Sinica, 1–20 (2022)

[4] Maruseac, M., Ghinita, G.: Precision-enhanced differentially private mining of high-confidence association rules. IEEE Transactions on Dependable & Secure Computing 17(6), 1297–1309 (2020).

[5] Hameed, M., Yang, F., Ghafoor, M. I., Jaskani, F. H., Islam, U., Fayaz, M., & Mehmood, G. (2022). IOTA-based Mobile crowd sensing: Detection of fake sensing using logit-boosted machine learning algorithms. Wireless Communications & Mobile Computing, 2022, 1-15.

[6] Yang, X., Yang, X., Yi, X., Khalil, I., Zhou, X., He, D., Huang, X., Nepal, S.: Blockchain-based secure & lightweight authentication for Internet of Things. IEEE Internet of Things Journal 9(5), 3321–3332 (2022).

[7] Tahir, B., Jolfaei, A., Tariq, M.: A novel experience-driven and federated intelligent threat-defense framework in IoMT. IEEE Journal of Biomedical & Health Informatics, 1–8 (2023).

[8] Zhang, Y.-T., Poon, C.C.Y.: Health informatics: Unobtrusive physiological measurement technologies. IEEE Journal of Biomedical & Health Informatics 17(5), 893–893 (2013).

[9] Liu, Y., Yu, J., Fan, J., Vijayakumar, P., Chang, V.: Achieving privacy-preserving DSSE for intelligent IoT healthcare system. IEEE Transactions on Industrial Informatics 18(3), 2010–2020 (2022).

[10] Bi, H., Liu, J., Kato, N.: Deep learning-based privacy preservation & data analytics for IoT-enabled healthcare. IEEE Transactions on Industrial Informatics 18(7), 4798–4807 (2022).

[11] Peisong Li, Xinheng Wang, Kaizhu Huang , Yi Huang , Shancang Li, Muddesar Iqbal, Multi-Model Running Latency Optimization in an Edge Computing Paradigm , Sensor, 22(16), 6097-6107, (2022).

[12] W. El-Shafai, F. Khallaf, E. M. El-Rabaie, et al., "A multi-stage security solution for medical color images in healthcare applications," Computer Systems Science and Engineering, vol. 46, no.3, pp. 3599–3618, (2023).

[13] Akter, M., Moustafa, N., Lynar, T. : Edge intelligence: Federated learning-based privacy protection framework for smart healthcare systems. IEEE Journal of Biomedical & Health Informatics 26(12), 5805–5816 (2022).

[14] Meng, L., & Li, D. (2023). Novel Edge Computing-Based Privacy-Preserving Approach for Smart Healthcare Systems in the Internet of Medical Things. Journal of Grid Computing, 21(4), 66.

[15] Mehmood, G., Khan, M. Z., Bashir, A. K., Al-Otaibi, Y. D., & Khan, S. (2023). An Efficient QoS-Based Multi-Path Routing Scheme for Smart Healthcare Monitoring in Wireless Body Area Networks. Computers & Electrical Engineering, 109, 108517.

[16] Asif-Ur-Rahman, M., Afsana, F., Mahmud, M., Kaiser, M.S., Ahmed, M.R., Kaiwartya, O., James-Taylor, A.: Toward a heterogeneous mist, fog, & cloud-based framework for the Internet of Healthcare Things. IEEE Internet of Things Journal 6(3), 4049–4062 (2019).

[17] Lv, Z., Qiao, L.: Analysis of healthcare big data. Future Generation Computer Systems 109, 103–110 (2020)

[18] Malina, L., Srivastava, G., Dzurenda, P., Hajny, J., Ricci, S.: A privacy-enhancing framework for Internet of Things services. In: Network & System Security (NSS 2019), Japan, Dec 15–18, vol. 13, pp. 77–97 (2019). Springer

[19] Mosaiyebzadeh, F., Pouriyeh, S., Parizi, et al..: Privacy-enhancing technologies in federated learning for the Internet of Healthcare Things: A survey. arXiv preprint arXiv:2303.14544 (2023)

[20] Guo, X., Lin, H., Wu, Y., Peng, M.: A new data clustering strategy for enhancing mutual privacy in healthcare IoT systems. Future Generation Computer Systems 113, 407–417 (2020)

[21] Zhang, M., Chen, Y., Susilo, W.: Ppo-cpq: A privacy-preserving optimization of clinical pathway query for e-healthcare systems. IEEE Internet of Things Journal 7(10), 10660–10672 (2020).

[22] A. A. Siddique et al., Privacy-Enhanced Pneumonia Diagnosis: IoT-Enabled Federated Multi-Party Computation in Industry 5.0, IEEE Transactions on Consumer Electronics, in press, (2024).

[23] Ni, C., Cang, L., Gope, P., Min, G.: Data anonymization evaluation for big data & IoT environment. Information Sciences 605, 381–392 (2022)

[24] Li, T., Wang, H., He, D., Yu, J.: Blockchain-based privacy-preserving and rewarding private data sharing for IoT. IEEE Internet of Things Journal 9(16), 15138–15149 (2022)

[25] Fan, Y., Bai, J., Lei, X., Lin, W., Hu, Q., Wu, G., Guo, J., Tan, G.: Ppmck: Privacy-preserving multi-party computing for k-means clustering. Journal of Parallel & Distributed Computing 154, 54–63 (2021)

[26] Samet, S., Miri, A., Orozco-Barbosa, L.: Privacy preserving k-means clustering in multi-party environment. In: SECRYPT, pp. 381–385 (2007)

[27] Kirienko, M., Sollini, M., Ninatti, G., Loiacono, D., Giacomello, E., Gozzi, N., Amigoni, F., Mainardi, L., Lanzi, P.L., Chiti, A.: Distributed learning: A reliable privacy-preserving strategy to change multicenter collaborations using AI. European Journal of Nuclear Medicine & Molecular Imaging 48, 3791–3804 (2021)

[28] Andrew, J., Eunice, R.J., Karthikeyan, J.: An anonymization-based privacy-preserving data collection protocol for digital health data. Frontiers in Public Health 11 (2023)

[29] Sowmiya, B., Poovammal, E.: A heuristic k-anonymity-based privacy preserving for student management hyperledger fabric blockchain. Wireless Personal Communications 127(2), 1359–1376 (2022)

[30] Liu, Y.-l., Huang, L., Yan, W., Wang, X., Zhang, R.: Privacy in AI and IoT: The privacy concerns of smart speaker users & personal information protection laws in China. Telecommunications Policy 46(7), 102334 (2022)

[31] Mhlanga, D.: The role of artificial intelligence & machine learning amid the Covid-19 pandemic: What lessons are we learning. International Journal of Environmental Research & Public Health 19(3), 1879 (2022)

[32] Carey, C., Dick, T., Epasto, A., Javanmard, et al., G.H., Vassilvitskii, S., et al.: Measuring re-identification risk. Proceedings of the ACM on Management of Data 1(2), 1–26 (2023)

[33] Atkinson, D.: Geometry in medical imaging: DICOM & NIfTI formats (2022)

[34] Mujumdar, A., Vaidehi, V.: Diabetes prediction using machine learning algorithms. Procedia Computer Science 165, 292–299 (2019)

[35] Zhao, Z., Sun, Y., Bashir, A.K., Lin, Z.: Adadpfed: A differentially private federated learning algorithm with adaptive noise on non-iid data. IEEE Transactions on Consumer Electronics (2023)

[36] Arqub, O.A., Abo-Hammour, Z.: Numerical solution of systems of second-order boundary value problems using continuous genetic algorithm. Information Sciences 279, 396–415 (2014)

[37] Abo-Hammour, Z., Abu Arqub, O., Momani, S., Shawagfeh, N., et al.: Optimization solution of Troesch’s and Bratu’s problems of ordinary type using novel continuous genetic algorithm. Discrete Dynamics in Nature & Society 2014 (2014)