http://dx.doi.org/10.62762/TISC.2025.335076

RESEARCH ARTICLE

Privacy-Preserving Artificial Intelligence for Diabetes Prediction: A Comparison of Centralised and Federated Learning

Aimie Grant[1,*], **Amir Hussain**[1], **Mandar Gogate**[1], **Nasir Saleem**[1] **and Kia Dashtipour**[1]

[1] School of Computing, Engineering and the Built Environment, Edinburgh Napier University, Edinburgh EH11 4BN, United Kingdom

Abstract

Artificial intelligence (AI) is increasingly used in healthcare to support disease prediction and clinical decision-making. Traditional centralised machine learning approaches often require the aggregation of sensitive patient data into a single repository, which raises substantial privacy, ethical, and regulatory concerns. Federated learning has emerged as a privacy-preserving alternative that enables collaborative model training across distributed data sources without sharing raw patient data. In this study, we investigate whether federated learning can achieve predictive performance comparable to that of centralised machine learning when applied to structured healthcare data. Using the PIMA Indians Diabetes dataset, several centralised machine learning models are evaluated and compared with a federated logistic regression model implemented using the Flower framework. Model performance is assessed using standard classification evaluation metrics. The results show that the federated approach achieves performance comparable to the centralised baselines, with a small reduction in predictive performance. These findings indicate that federated learning is a practical and effective solution for privacy-preserving predictive modelling in healthcare.

Keywords: federated learning, privacy-preserving AI, healthcare, diabetes prediction.

1 Introduction

AI in Healthcare Prediction

Artificial intelligence (AI) and machine learning (ML) techniques are increasingly used within healthcare to support disease prediction, risk assessment, and clinical decision-making. Rather than replacing clinicians, these systems are commonly deployed as clinical decision support tools, assisting healthcare professionals by identifying patterns within patient data that may not be immediately apparent through manual analysis [1, 2, 13]. By analysing large volumes of clinical information, ML models can support early risk stratification and preventative intervention, which is particularly valuable for chronic conditions where timely treatment can reduce long-term complications and healthcare costs. Many healthcare prediction

Submitted: 20 December 2025
Accepted: 12 February 2026
Published: 10 April 2026

Vol. 2, **No.** 2, 2026.
10.62762/TISC.2025.335076

***Corresponding author:**
✉ Aimie Grant
40601584@live.napier.ac.uk

Citation
Grant, A., Hussain, A., Gogate, M., Saleem, N., & Dashtipour, K. (2026). Privacy-Preserving Artificial Intelligence for Diabetes Prediction: A Comparison of Centralised and Federated Learning. *ICCK Transactions on Information Security and Cryptography*, 2(2), 73–81.

tasks rely on structured clinical data, including laboratory measurements, demographic attributes, and physiological indicators, due to their widespread availability, interpretability, and compatibility with traditional predictive models.

Diabetes as a Benchmark Task

Diabetes prediction represents a widely studied healthcare application of machine learning techniques due to the prevalence of the condition and the availability of well-established clinical risk factors. Diabetes is a chronic disease associated with serious long-term complications, including cardiovascular disease, kidney failure, and neuropathy. Predictive modelling aims to identify individuals at an increased risk prior to a formal diagnosis, enabling earlier intervention and overall improved disease management [3]. As a result, diabetes prediction is frequently used as a benchmark task for evaluating the effectiveness of machine learning approaches applied to structured healthcare data [4].

Privacy, Governance, and Regulation

Despite the demonstrated benefits of AI-driven healthcare systems, the use of patient data introduces significant privacy and governance challenges. Healthcare data is inherently sensitive and subject to strict confidentiality requirements. Traditional machine learning approaches typically rely on centralised data collection, where data from multiple sources is aggregated into a single repository for model training. In practice, this introduces a single point of failure, where a security breach can expose large volumes of sensitive patient information. Healthcare institutions are often reluctant or legally unable to share raw patient data across boundaries set by the organisation, which limits the feasibility of large-scale collaborative model development. These challenges can undermine patient trust and also restrict the deployment of data-driven healthcare technologies.

Regulatory frameworks further constrain the deployment of centralised AI systems in healthcare. The General Data Protection Regulation (GDPR) enforces principles such as data minimisation, purpose limitation, and accountability, which directly affect how personal health data can be processed and shared [5]. Large-scale data centralisation can conflict with these principles, particularly in multi-institution settings. Additionally, emerging regulations such as the proposed EU AI Act classify many healthcare AI systems as high-risk, introducing additional requirements around transparency, governance, and risk management. These regulatory constraints make traditional data sharing impractical, prompting the use of alternative machine learning approaches.

Federated Learning as an Alternative

Federated learning has emerged as a potential solution to these challenges. Rather than requiring raw data to be shared centrally, federated learning enables models to be trained locally across multiple clients, with only model updates exchanged for aggregation. This approach reduces the exposure of sensitive patient data while still enabling collaborative model development across institutions. Federated learning has shown promising results in healthcare contexts, particularly in medical imaging applications [6, 9]. However, structured clinical data presents distinct challenges, including feature heterogeneity, class imbalance, and variations in data distributions across institutions. As a result, the effectiveness of federated learning for structured tabular healthcare data remains less explored.

This paper investigates whether federated learning can maintain predictive performance comparable to traditional centralised machine learning when applied to structured healthcare data. Using diabetes prediction as a case study, centralised and federated learning approaches are evaluated and compared in terms of predictive performance and practical trade-offs. The main contributions of this work are a comparative evaluation of centralised and federated learning for diabetes prediction using structured tabular data, and an analysis of the resulting performance, privacy, and regulatory implications for real-world healthcare deployment.

2 Related Work

Centralised Machine Learning in Healthcare

Machine learning has been widely applied to healthcare prediction tasks, with numerous studies demonstrating strong performance using centralised datasets. Supervised learning models such as Logistic Regression, Support Vector Machines, and ensemble-based approaches have been extensively used to predict conditions including diabetes, cardiovascular disease, and cancer [1, 10]. These models are particularly well suited to structured clinical data due to their interpretability and ability to handle numerical features commonly found in electronic health records. Despite their strong

predictive performance, these approaches depend on unrestricted access to large, multi-source patient datasets, which is increasingly unrealistic in regulated healthcare environments [7].

Privacy-Preserving Machine Learning

Growing concerns surrounding privacy, trust, and data governance have motivated significant research into privacy-preserving machine learning techniques. Centralised data storage increases the potential consequences of data breaches and raises ethical concerns related to patient consent, secondary data use, and long-term data retention. In response, researchers have explored alternative approaches that aim to reduce privacy risks while maintaining acceptable levels of predictive performance. These approaches include data anonymisation, secure computation, and distributed learning paradigms, each of them offering different trade-offs between privacy protection as well as model utility [8].

Federated Learning in Healthcare

Federated learning has gained particular attention as a promising privacy-preserving approach that enables collaborative model training without requiring raw data to be shared. In this paradigm, models are trained locally at participating institutions, with only model updates exchanged for aggregation. In healthcare, federated learning has been successfully applied to medical imaging tasks such as tumour classification, organ segmentation, and radiological analysis [9]. Existing studies suggest that federated learning can achieve performance comparable to centralised models under suitable conditions, while significantly reducing the need for direct data sharing and centralised data storage.

Limitations of Existing Work

Despite these successes, much of the existing federated learning literature in healthcare focuses on image-based data, which differs substantially from structured clinical datasets. Structured tabular data presents distinct challenges, including feature heterogeneity, class imbalance, and variations in data distributions across institutions, all of which can negatively impact model convergence and performance in federated settings. Diabetes prediction typically relies on structured clinical attributes rather than imaging data, yet relatively few studies have systematically evaluated federated learning approaches in this context. This gap in the literature motivates the present work, which investigates the suitability of federated learning for diabetes prediction using structured tabular healthcare data and compares its performance against traditional centralised machine learning approaches.

3 Methodology

In this section we present the proposed methodology for privacy-preserving approach for diabetes prediction.

3.1 Dataset

The PIMA Indians Diabetes dataset [4] was used in this study as a benchmark dataset for diabetes prediction using structured healthcare data. The dataset consists of 768 patient records, each described by eight numerical clinical features, including plasma glucose concentration, blood pressure, body mass index, insulin levels, and age. The target variable is binary, indicating the presence or absence of diabetes. It is worth mentioning that the PIMA Indians Diabetes dataset is a relatively small dataset that does not fully capture the scale and heterogeneity of real-world federated healthcare environments.

This dataset is widely used in healthcare machine learning research due to its well-defined clinical features and public availability, making it suitable for reproducible experimentation. As expected and common in medical datasets, the class distribution is imbalanced, with fewer positive diabetes cases than negative cases, reflecting real-world screening scenarios.

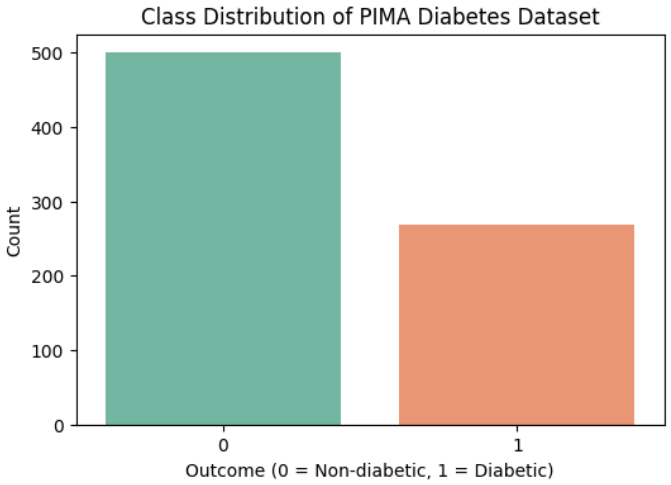

Figure 1. Class distribution of the PIMA Indians Diabetes dataset.

Figure 1 illustrates the class imbalance present in the dataset. This imbalance motivates the use of evaluation metrics beyond accuracy, as high accuracy

may obscure poor performance in identifying positive diabetes cases.

3.2 Data Preprocessing

Several preprocessing steps were applied prior to model training to improve data quality and model performance. Certain clinical features contain zero values that are medically implausible, particularly for attributes such as glucose concentration, blood pressure, body mass index, and insulin levels. These values were treated as missing data.

Missing values were handled using median imputation due to its robustness to outliers and suitability for skewed medical data distributions. Following imputation, feature standardisation was applied to transform all input variables to a common scale with zero mean and unit variance. This step is particularly important for distance-based and gradient-based models.

Identical preprocessing pipelines were applied across both centralised and federated learning settings. In federated learning, inconsistencies in preprocessing between clients can negatively affect convergence as well as global model performance, which makes uniform preprocessing a critical requirement.

3.3 Learning Models

3.3.1 Centralised Models

Three centralised machine learning models were implemented to establish baseline performance: Logistic Regression, Support Vector Machine with a radial basis function kernel, and Random Forest. These models were selected due to their widespread use in healthcare prediction tasks, interpretability, and suitability for structured tabular data.

Logistic Regression provides a transparent and interpretable baseline suitable for supporting clinical decision-making. The Support Vector Machine was included to capture potential non-linear relationships, while the Random Forest model represents ensemble-based approaches capable of modelling complex feature interactions.

All centralised models were trained using an 80:20 stratified train-test split to preserve class balance.

3.3.2 Federated Learning Model

Federated learning was implemented using the Flower framework to simulate a distributed, multi-institution healthcare environment. Logistic Regression was selected for the federated setting due to its computational efficiency, interpretability, and suitability for distributed optimisation.

Each client trained a local model using its own data, and only model parameters were then shared with a central server for aggregation.

3.4 Federated Learning Design

As illustrated in Figure 2, the federated learning architecture consists of multiple clients (healthcare institutions) that train local models on their respective data. Only the model parameters are transmitted to a central server for aggregation, while raw patient data remains on-premises.

3.4.1 Client Partitioning

To simulate institutional separation, the dataset was partitioned evenly across three clients, each representing an independent healthcare institution. Each client trained models locally on its subset of data.

While this partitioning does not fully capture real-world institutional heterogeneity, it introduces statistical variation between clients while maintaining experimental control and reproducibility.

3.4.2 Threat Model and Privacy Assumptions

A standard cross-silo federated learning threat model is used, where clients are honest-but-curious and may attempt to infer information from shared model updates. The aggregation server does not have access to raw patient data.

While federated learning reduces direct data exposure, it does not guarantee complete privacy protection. This study focuses on privacy benefits at the architectural level rather than providing formal privacy guarantees. The integration of secure aggregation and differential privacy is considered future work [11].

3.5 Experimental Setup and Reproducibility

To ensure reproducibility, fixed random seeds and consistent data splits were used across all experiments. Identical preprocessing steps, model configurations, and evaluation procedures were applied to both centralised and federated learning settings.

This design isolates the effect of the learning architecture, so any performance differences are mainly due to the use of distributed training rather than differences in configuration.

All experiments were conducted using Python-based machine learning libraries, including `scikit-learn`

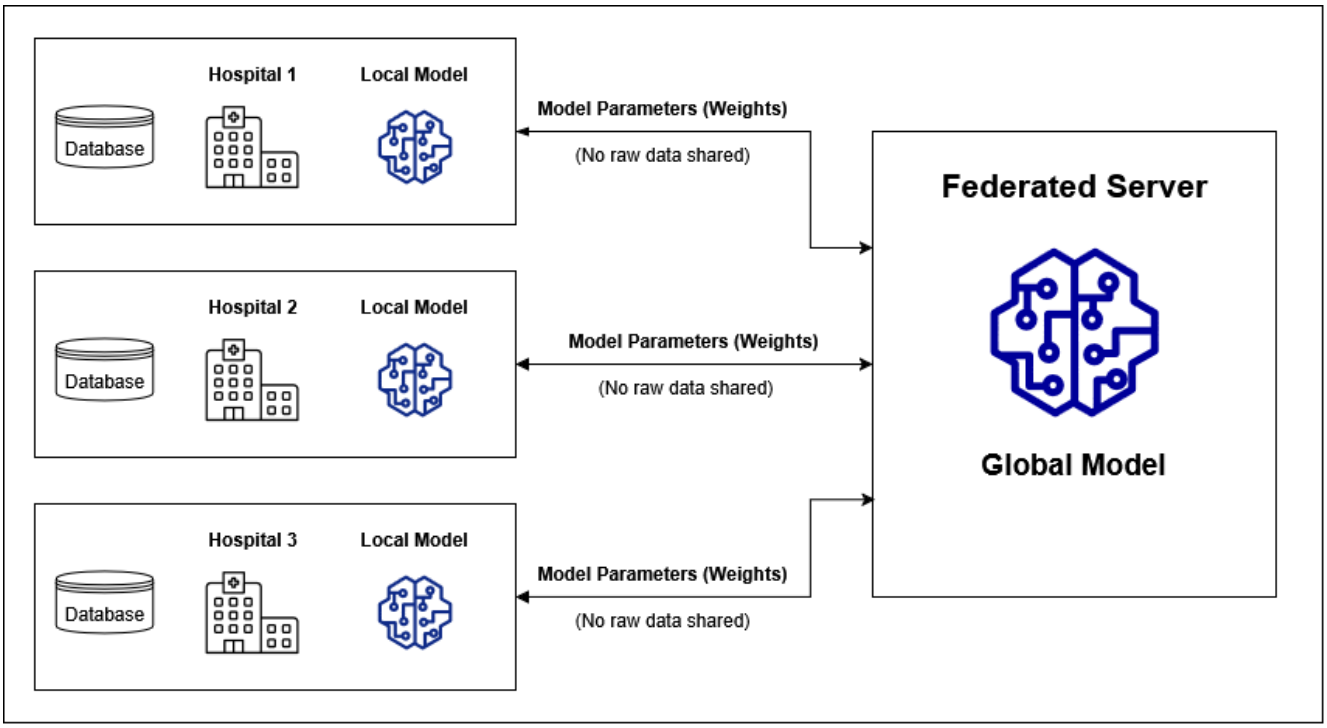

Figure 2. Federated learning architecture illustrating local training and centralised aggregation.

for centralised models and the Flower framework for federated learning.

3.6 Evaluation Metrics

Model performance was evaluated using accuracy, precision, recall, F1-score, and the area under the receiver operating characteristic curve (ROC-AUC). In healthcare prediction tasks, accuracy alone can be misleading due to class imbalance and asymmetric error costs.

Recall and ROC-AUC were therefore emphasised, as they better reflect a model's ability to identify positive diabetes cases across varying decision thresholds, which is critical in screening-oriented clinical applications.

4 Results and Discussion

4.1 Overall Model Performance

The centralised Logistic Regression model achieved strong baseline performance on the diabetes prediction task, with an ROC-AUC of approximately 0.78. This result is consistent with previously reported performance for Logistic Regression models trained on the PIMA Indians Diabetes dataset and confirms that the preprocessing and experimental setup were appropriate.

The federated Logistic Regression model achieved a comparable ROC-AUC of approximately 0.77, representing only a marginal reduction in predictive performance. This small performance gap suggests that distributing the training process across multiple clients does not substantially degrade the model's ability to distinguish between diabetic and non-diabetic cases, despite the absence of centralised access to raw patient data.

Table 1 summarises the evaluation metrics for both approaches. While overall accuracy remains similar, small reductions are observed in precision, recall, and F1-score for the federated model, particularly for the positive class, which is of greater clinical importance in diabetes screening scenarios.

Table 1. Performance comparison between centralised and federated Logistic Regression models.

Metric	Centralised LR	Federated LR
Accuracy	0.77	0.75
Precision	0.72	0.70
Recall	0.67	0.64
F1-score	0.69	0.66
ROC-AUC	0.78	0.77

4.2 Comparison of Centralised Models

In addition to Logistic Regression, Support Vector Machine (SVM) and Random Forest models were evaluated in the centralised setting to provide broader performance context. These models represent

commonly adopted approaches for structured healthcare prediction tasks, ranging from linear to non-linear and ensemble-based methods.

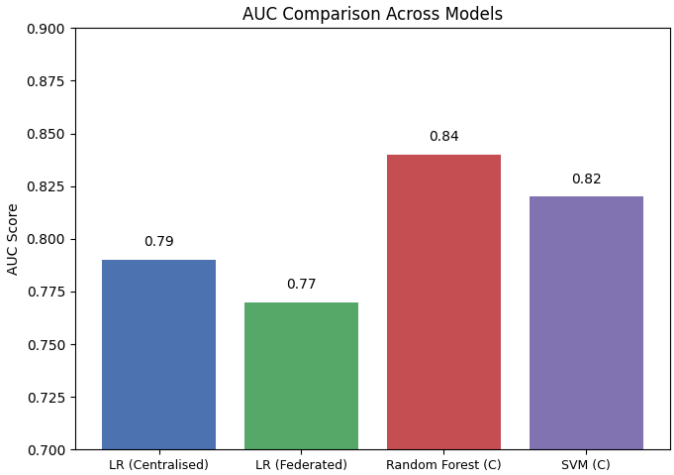

Figure 3. ROC-AUC comparison across centralised and federated machine learning models.

As shown in Figure 3, ensemble and kernel-based models achieve strong discrimination in the centralised setting. However, Logistic Regression provides a favourable balance between performance and interpretability, which motivates its selection as the baseline for federated learning.

4.3 Confusion Matrix Analysis

To provide deeper insight into classification behaviour beyond aggregate metrics, confusion matrices were analysed for the centralised models. Confusion matrices reveal class-specific errors, which is particularly important in healthcare screening tasks where false negatives can delay clinical intervention, which may be dangerous.

Figures 4, 5, and 6 show that Logistic Regression and SVM models achieve higher sensitivity for diabetic cases, while Random Forest demonstrates stronger specificity but a higher number of false negatives. This trade-off is critical in diabetes screening, where failing to identify at-risk individuals can delay early intervention as mentioned earlier.

From a clinical perspective, models that prioritise recall while maintaining interpretability are often preferred in screening contexts. Logistic Regression therefore remains an attractive candidate for deployment in regulated healthcare environments.

4.4 Federated Versus Centralised Learning Analysis

The receiver operating characteristic (ROC) curves for the centralised and federated Logistic Regression models are shown in Figure 7.

The close alignment of the curves across a wide range of classification thresholds indicates that the federated model maintains comparable discriminatory capability to the centralised baseline. The marginal reduction in ROC-AUC is likely attributable to decentralised optimisation, where each client trains on locally available data only.

4.5 Effects of Data Distribution and Training Dynamics

The slight reduction in federated performance can be explained by factors that are inherent to federated learning. Partitioning the dataset across multiple clients introduces non-identically distributed (non-IID) data characteristics, even when data is split evenly. Such statistical heterogeneity can negatively impact convergence as well as global optimisation.

Additionally, federated learning relies on iterative communication rounds and parameter aggregation, constraining information sharing compared to centralised training. Despite these limitations, the observed performance gap still remains small, which indicates that federated learning effectively mitigates privacy risks without substantially compromising predictive performance.

4.6 Practical and Clinical Implications

From a clinical perspective, the ability of the federated model to achieve performance close to the centralised baseline is significant. In diabetes screening scenarios, models are typically used to support early risk identification rather than definitive diagnosis.

Federated learning therefore offers a practical balance between predictive accuracy and privacy preservation. By avoiding the centralisation of sensitive patient data, it aligns with regulatory requirements while still enabling collaborative model development across healthcare institutions.

4.7 Summary of Key Findings

Overall, the experimental results demonstrate that federated learning can achieve predictive performance comparable to traditional centralised machine learning for structured healthcare data. While small reductions in precision and recall are observed in the federated setting, the overall discriminatory capability, as measured by ROC-AUC, remains largely intact.

These findings suggest that the performance cost of decentralised training is small compared to the

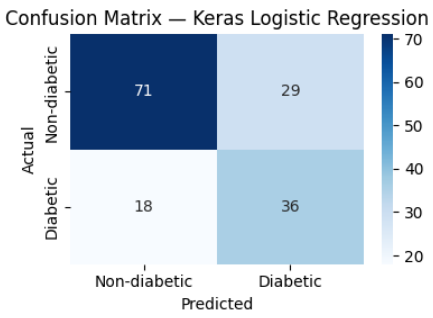

Figure 4. Logistic Regression

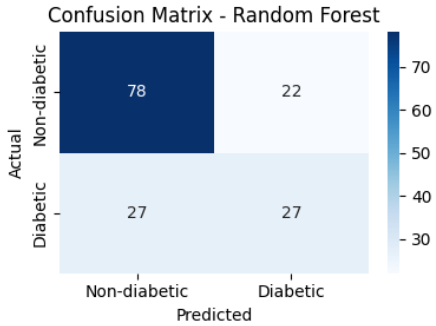

Figure 5. Random Forest

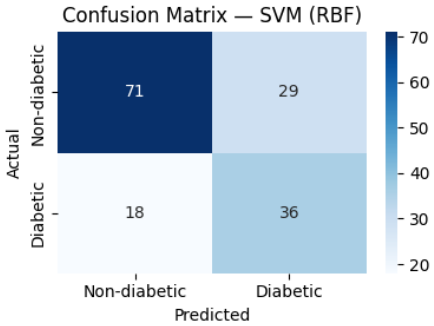

Figure 6. SVM (RBF)

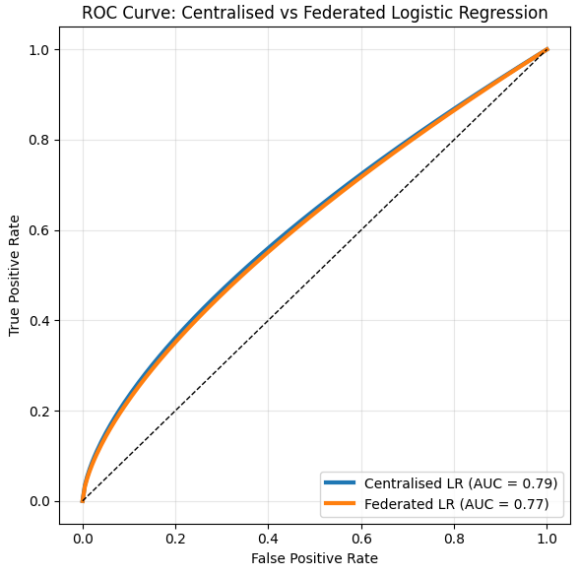

Figure 7. ROC curve comparison between centralised and federated Logistic Regression models.

privacy and governance benefits it provides. In particular, Logistic Regression emerges as a strong candidate for federated deployment due to its stability, interpretability, and consistent performance across both learning paradigms.

5 Regulatory and Ethical Considerations

5.1 Regulatory Compliance

Federated learning aligns closely with key regulatory principles governing healthcare data processing. Under the General Data Protection Regulation (GDPR), principles such as data minimisation, purpose limitation, and data security restrict unnecessary sharing and centralisation of personal health data. By retaining patient data locally and exchanging only model parameters, federated learning reduces the scope of personal data processing and limits any cross-institution data exposure.

In addition, emerging regulatory frameworks such as the proposed EU AI Act classify many healthcare AI systems as high-risk, which introduces enhanced requirements for governance, transparency, and risk management. Federated learning supports these requirements by decentralising data processing and reducing reliance on centralised data repositories, which is representative of high-value targets for cyberattacks.

5.2 Ethical Considerations

Beyond regulatory compliance, federated learning also addresses broader ethical issues related to patient trust, data handling, and institutional accountability. Centralised machine learning approaches require organisations to give up direct control of sensitive data, which raises concerns about consent, secondary use, and long-term data storage [14]. Federated learning mitigates these risks by allowing institutions to maintain local control over patient data while still participating in collaborative model development.

While federated learning alone does not guarantee full regulatory or ethical compliance, it provides a strong architectural foundation upon which additional safeguards, such as privacy-enhancing technologies.

6 Limitations

This study is subject to several limitations that should be considered when interpreting the results. First, the evaluation was conducted using a single public dataset, which may not fully capture the complexity, scale, and heterogeneity of real-world healthcare data. Clinical datasets collected across institutions often differ in demographic composition, measurement practices, and data quality, which may affect model performance in practice.

Second, the federated learning environment was simulated using a limited number of clients. While this setup enables controlled experimentation and

fair comparison with centralised baselines, it may not reflect the communication constraints, client heterogeneity, and scalability challenges present in large-scale, real-world multi-institution deployments.

Finally, the federated learning configuration did not incorporate formal privacy guarantees such as differential privacy, homomorphic encryption or secure aggregation. As a result, privacy protection in this study is architectural rather than mathematically bounded, and model updates could still leak limited information under particular adversarial scenarios.

These limitations highlight opportunities for future research rather than deficiencies in the proposed approach, and also motivate further investigation into scalable private federated learning systems within healthcare [11, 12].

7 Conclusion

This paper investigated federated learning as a privacy-preserving alternative to traditional centralised machine learning for diabetes prediction using structured healthcare data. The primary objective was to assess whether federated learning can maintain predictive performance comparable to centralised approaches while reducing the need for raw patient data sharing. By comparing centralised and federated Logistic Regression models on the PIMA Indians Diabetes dataset, this study evaluated both predictive effectiveness and the practical trade-offs that were introduced by distributed training.

The experimental results demonstrate that federated learning achieves performance closely aligned with the centralised baseline, with only a marginal reduction in evaluation metrics such as ROC-AUC, precision, and recall. These findings indicate that distributing the training process across multiple clients does not substantially impair the model's ability to distinguish between diabetic and non-diabetic cases. Importantly, this suggests that strong predictive performance can be retained even when direct access to centrally aggregated patient data is restricted.

From a practical perspective, the results highlight the potential of federated learning to address key privacy and governance challenges in healthcare machine learning. By keeping sensitive patient data local to participating institutions and only sharing the model updates, federated learning reduces the risks associated with data breaches, unauthorised access, and regulatory non-compliance. This architectural approach aligns closely with data protection principles such as data minimisation and purpose limitation under the General Data Protection Regulation (GDPR). As many healthcare AI systems are classified as high-risk under emerging frameworks such as the EU AI Act, federated learning offers a promising foundation for enabling collaborative model development in regulated environments.

Despite these promising results, this study has several limitations. The evaluation was conducted using a single public dataset and a simulated federated environment with a limited number of clients, which may not fully capture the heterogeneity, scale, and operational constraints of real-world healthcare systems. In addition, the federated learning configuration did not incorporate formal privacy guarantees such as differential privacy, homomorphic encryption or secure aggregation, meaning that privacy protection in this study is architectural rather than mathematically bounded.

Future work will extend this investigation by evaluating federated learning on larger and more diverse healthcare datasets, including real multi-institution deployments with non-identically distributed data. Further research will also explore the integration of advanced privacy-preserving techniques, such as differential privacy and homomorphic encryption, to strengthen privacy guarantees while assessing their impact on model utility and training efficiency. Together, these directions aim to support the development of practical, privacy-preserving AI systems that can be responsibly deployed in regulated healthcare environments.

Data Availability Statement

Data will be made available on request.

Funding

This work was supported without any funding.

Conflicts of Interest

The authors declare no conflicts of interest.

AI Use Statement

The authors declare that no generative AI was used in the preparation of this manuscript.

Ethical Approval and Consent to Participate

Not applicable.

References

[1] Rajkomar, A., Dean, J., & Kohane, I. (2019). Machine learning in medicine. *New England Journal of Medicine, 380*(14), 1347-1358. [CrossRef]

[2] Topol, E. J. (2019). High-performance medicine: the convergence of human and artificial intelligence. *Nature medicine, 25*(1), 44-56. [CrossRef]

[3] World Health Organization. (2016). *Global report on diabetes: executive summary*. World Health Organization. Available at: https://iris.who.int/handle/10665/204874

[4] Smith, J. W., Everhart, J. E., Dickson, W. C., Knowler, W. C., & Johannes, R. S. (1988, November). Using the ADAP learning algorithm to forecast the onset of diabetes mellitus. In *Proceedings of the annual symposium on computer application in medical care* (p. 261).

[5] Cloud Security Alliance. (2018). *GDPR preparation & challenges survey report*. Available at: https://cloudsecurityalliance.org/press-releases/2018/04/17/gdpr-preparation-and-challenges-survey-report

[6] McMahan, B., Moore, E., Ramage, D., Hampson, S., & y Arcas, B. A. (2017, April). Communication-efficient learning of deep networks from decentralized data. In *Artificial intelligence and statistics* (pp. 1273-1282). Pmlr.

[7] Kavakiotis, I., Tsave, O., Salifoglou, A., Maglaveras, N., Vlahavas, I., & Chouvarda, I. (2017). Machine learning and data mining methods in diabetes research. *Computational and structural biotechnology journal, 15*, 104-116. [CrossRef]

[8] Kairouz, P., & McMahan, H. B. (2021). Advances and open problems in federated learning. *Foundations and trends in machine learning, 14*(1-2), 1-210. [CrossRef]

[9] Rieke, N., Hancox, J., Li, W., Milletari, F., Roth, H. R., Albarqouni, S., ... & Cardoso, M. J. (2020). The future of digital health with federated learning. *NPJ digital medicine, 3*(1), 119. [CrossRef]

[10] Sheller, M. J., Edwards, B., Reina, G. A., Martin, J., Pati, S., Kotrotsou, A., Milchenko, M., Xu, W., Marcus, D., Colen, R. R., & Bakas, S. (2020). Federated learning in medicine: facilitating multi-institutional collaborations without sharing patient data. *Scientific Reports, 10*(1), 12598. [CrossRef]

[11] Li, T., Sahu, A. K., Talwalkar, A., & Smith, V. (2020). Federated learning: Challenges, methods, and future directions. *IEEE signal processing magazine, 37*(3), 50-60. [CrossRef]

[12] Truex, S., Baracaldo, N., Anwar, A., Steinke, T., Ludwig, H., Zhang, R., & Zhou, Y. (2019, November). A hybrid approach to privacy-preserving federated learning. In *Proceedings of the 12th ACM workshop on artificial intelligence and security* (pp. 1-11). [CrossRef]

[13] Shortliffe, E. H., Shortliffe, E. H., Cimino, J. J., & Cimino, J. J. (2014). *Biomedical informatics: computer applications in health care and biomedicine*. Springer. [CrossRef]

[14] Mittelstadt, B. (2019). Principles alone cannot guarantee ethical AI. *Nature machine intelligence, 1*(11), 501-507. [CrossRef]

Aimie Grant received the B.S. degree in Cyber Security from Edinburgh Napier University, Edinburgh, U.K., in 2026. (Email: 40601584@live.napier.ac.uk)

Amir Hussain received the B.Eng. degree (First Class Honours with distinction) and the Ph.D. degree from the University of Strathclyde, Glasgow, U.K., in 1992 and 1997, respectively. He is a Professor of Computing Science at Edinburgh Napier University, U.K., and the founding Director of the Centre for Artificial Intelligence & Robotics (CAIR). His research interests span cognitive data science, trustworthy AI, machine learning, and cyber analytics for smart healthcare and industrial systems. He has co-authored hundreds of publications, including 350+ international journal papers and 25 books/monographs, and has supervised over 40 Ph.D. students. Professor Hussain serves as founding Editor-in-Chief of Springer's Cognitive Computation journal and holds editorial roles in several leading journals, including IEEE Transactions on Neural Networks and Learning Systems and Information Fusion. He has led major research projects funded by national and international agencies and regularly chairs flagship conferences in computational intelligence and AI. (Email: a.hussain@napier.ac.uk)

Kia Dashtipour received the B.Sc. degree in Computer Engineering and the Ph.D. degree in Artificial Intelligence and Machine Learning from the University of Stirling, U.K. He is currently a lecture at Edinburgh Napier University, U.K. His research interests include cybersecurity, machine learning, multimodal and cognitive computing, sentiment and opinion mining, speech enhancement, and intelligent systems. He has published extensively in leading international journals and conferences and actively contributes to interdisciplinary research in AI-driven security and human-centred computing. (Email: k.dashtipour@napier.ac.uk)

Nasir Saleem is a researcher and academic affiliated with Edinburgh Napier University and Gomal University. His research interests focus on areas including interactive multimedia, artificial intelligence, and related computer science disciplines. He has published in international journals such as the International Journal of Interactive Multimedia and Artificial Intelligence and contributes to interdisciplinary research and academic scholarship in intelligent systems and multimedia applications. (Email: n.saleem@napier.ac.uk)

Mandar Gogate is a Principal Research Fellow at Edinburgh Napier University, U.K., where he conducts research in speech enhancement, audio-visual speech separation, and machine learning. His work focuses on developing advanced signal processing and AI-driven methods for improving speech intelligibility and separation in complex acoustic environments. He has authored numerous publications in leading international journals and conferences and contributes actively to interdisciplinary research at the intersection of multimedia signal processing and machine learning. (Email: m.gogate@napier.ac.uk)