Online-Journals.org

iJOE International Journal of Online and Biomedical Engineering

iJOE | eISSN: 2626-8493 | Vol. 22 No. 3 (2026) | OPEN ACCESS

https://doi.org/10.3991/ijoe.v22i03.57779

PAPER

DeepAsthmaNet: A Time-Aware Federated Prognostic Framework for Personalised Paediatric Asthma Risk Stratification in Primary Care

Pushkal Kumar Shukla[1] (✉), Sarika Jain[1], Siddharth Kalra[2]

[1]AIIT, Amity University, Noida, India

[2]Capgemini, Australia

pushkal.shukla@s.amity.edu

ABSTRACT

Asthma in children is still a common chronic illness with a wide range of clinical manifestations and progressions. Timely action and individualised care depend on early and precise risk assessment. This study presents a federated architecture used to train DeepAsthmaNet integrates deep neural architectures and attention mechanisms to enable early, accurate prediction of childhood asthma by combining clinical, genetic and environmental data. After rigorous preprocessing, feature engineering and redundancy elimination, the pipeline leverages inputs from genetic biomarkers, wearables, environmental sensors, and electronic health records. At an optimal learning rate of 0.80, it achieves superior performance with an MSE of 0.00015, an MAE of 0.0051, and an AUC of 0.9773, outperforming baseline models (ANN, DNN, RNN, LSTM). To handle heterogeneous distributed data, three federated aggregation strategies, FedAvg, FedProx and Krum, were tested, with FedAvg delivering the best balance of accuracy, convergence stability and communication efficiency.

KEYWORDS

DeepAsthmaNet, federated learning, paediatric asthma, risk stratification, attention mechanism

1 INTRODUCTION

The paediatric asthma is a prevalent chronic respiratory disorder affecting millions of children and creating serious challenges globally and imposing considerable pressure on both families and healthcare systems, both clinically and financially [1]. Childhood asthma leads to a multifaceted interaction among environmental triggers, genetic susceptibility, patient behaviours and clinical signs [2]. For better long-term results, reduced exacerbation rates, and early treatment, accurate and fast risk classification is crucial [3]. Primary care diagnostics, however, frequently

Shukla, P. K., Jain, S., Kalra, S. (2026). DeepAsthmaNet: A Time-Aware Federated Prognostic Framework for Personalised Paediatric Asthma Risk Stratification in Primary Care. *International Journal of Online and Biomedical Engineering (iJOE)*, 22(3), pp. 114–132. https://doi.org/10.3991/ijoe.v22i03.57779

Article submitted 2025-08-02. Revision uploaded 2025-10-17. Final acceptance 2025-12-12.

rely significantly on subjective patient reports and static clinical evaluations. These traditional approaches might not take into account how complicated and dynamic asthma is, which could hinder their capacity to accurately forecast hazards in the future [4]. New trends in deep learning, specifically in sequence modelling using RNNs, transformers and attention-based architecture, have significantly improved the capacity to learn from intricate time-series data, including wearable sensor data, environmental exposure records and electronic health records (EHRs) [5]. Although risk has been anticipated by utilising standard statistics and machine learning methods, typically these models are unable to incorporate temporal linkages or individual health changes over time. Additionally, these models usually struggle to maintain their effectiveness in heterogeneous and distributed healthcare systems [6].

At the same time, federated learning has become a feasible method for permitting decentralised model training across multiple healthcare facilities in different geographical regions while ensuring adherence to data privacy laws, which is especially crucial when managing private paediatric health records [7].

To make risk stratification easier for children with asthma, this paper suggests DeepAsthmaNet, a unique system with a time-sensitive federated learning model. DeepAsthmaNet incorporates transformer encoders to capture sequence patterns, a meta-learning algorithm to guarantee adaptation across a variety of clinical circumstances, and graph attention networks to simulate complex feature extraction. To make the model more interpretable and reliable in clinical contexts by incorporating explainability and uncertainty estimation modules. Recent works in federated and IoT-enabled healthcare systems emphasise the importance of privacy-preserving predictive analytics in chronic disease management [8]. Building on these advances, our study focuses specifically on paediatric asthma prediction, an area where such frameworks remain underexplored.

The core achievements of this study can be summed up as follows:

- We suggest a hybrid deep learning framework combining temporal attention, graph-based learning, and federated optimisation to improve paediatric asthma prediction.
- To improve predictive power, we use multimodal time series data, such as genetic, environmental and clinical data.
- To facilitate the training of decentralised models between clinical sites without exchanging unprocessed patient data, we employ a federated learning architecture that preserves privacy.
- We perform comprehensive tests on real-world paediatric datasets and compare the model to both traditional and cutting-edge baselines, showcasing its exceptional accuracy, generalisability and resilience.

The continuation of this study is systematised as follows: Section 2 outlines the related work employing ML and DL methods for asthma assessment and risk stratification. Section 3 outlines the problem statement and key challenges. Section 4 elaborates on DeepAsthmaNet framework, along with its design and methodology. Section 5 elaborates on the experimental setup, datasets along with evaluation standards and result analysis. While Section 6 offers concluding remarks. Finally, Section 7 summarises the limitations and identifies opportunities for future research.

2 RELATED WORK

A variety of tools and models from ML and DL strategies have been implemented recently to increase the accuracy, interpretability and personalisation of paediatric

asthma risk prediction. Using clinical and environmental data, Wang et al. [9] formulated a DNN-based approach and obtained encouraging results (MSE: 0.35, MAE: 0.28); however, the model was sensitive to class imbalance. Fernandes et al. [10] employed a LSTM network to record temporal changes in pollution exposure and disease progression. This network produced low RMSE (0.34) and high accuracy (83%) but required the best sequence length adjustment.

A convolutional neural network (CNN) for symptom recognition from signal and image data was created by Zhang et al. [11]. It had an accuracy rate of 80% but was not generalisable outside of the training environment. Conventional ML models are still applicable as well: Although the approach lacked customisation and temporal modelling, Liu et al. [12] achieved an F1 score of 0.76 using support vector machines (SVMs) with wrapper-based feature selection. However, these approaches often fail to address heterogeneity, longitudinal modelling and privacy concerns in real-world paediatric asthma management. Beyond clinical asthma prediction, federated and distributed machine learning have been applied successfully in IoT-driven healthcare and education platforms [13]. These studies highlight both the scalability and data privacy benefits of decentralised training, which our framework leverages in the paediatric asthma domain.

Using registry-based paediatric data, Raji et al. [14] coupled random forests and principal component analysis (PCA), showing good performance (accuracy: 85%) but little flexibility to adjust to changing health profiles. A transfer learning-based CNN that included structured clinical data and imaging was more recently introduced by Lee et al. [15]. It required a significant amount of labelled data, yet it achieved 90% accuracy and an MSE of 0.22. The transformer-based time series model with positional encoding and self-attention was used by Smith et al. [16] to analyse longitudinal symptom and environmental logs. It achieved 88% accuracy and an RMSE of 0.22, but it did not support federated learning. In a comparative investigation of demographic and EHR data, Ekpo et al. [17] used 32 machine learning algorithms. They reported an accuracy of approximately 89%, but they also pointed out the possibility of overfitting and the low cross-population generalisability. Similarly, Kumar et al. [18] achieved an RMSE of 0.25 using XGBoost with hyperparameter adjustment on clinical and demographic data, while overfitting was still an issue. On public paediatric asthma datasets, Banerjee et al. [19] conducted a multi-model comparison using RF, SVM, k-NN, and Naive Bayes. RF achieved 86% accuracy, while SVM had low recall. In a different study, Shaikh et al. [20] used correlation-based feature selection in conjunction with artificial neural networks on clinical demographic data, achieving an accuracy of 84% and an MAE of 0.31; however, overfitting issues were present in small datasets.

A rule-based decision tree method for lab tests and patient histories was presented by Ahmad et al. [21]. It provided transparency (Acc = 76%) but was not scalable. Naive Bayes was used by Ojha et al. [22] on symptom checklist data, yielding quick results (Acc = 75%, Precision = 0.70), but it was insufficient for modelling intricate connections. An MAE of 0.27 and 84% accuracy were obtained by Wang et al. [23] when they merged wearable sensor and EHR data using a multimodal deep neural network; however, they pointed out limitations in real-world applicability because of data scarcity. The necessity for an integrated, interpretable, and privacy-preserving framework like DeepAsthmaNet that can generalise across various, multimodal datasets and allow for decentralised deployment is highlighted by these studies taken together. Table 1 provides a comprehensive summary of this study, emphasising both their methodological advantages and disadvantages in relation to the prediction of paediatric asthma.

Table 1. Summary of existing ML and DL approaches for paediatric asthma prediction

Author(s)	Year	Method/Model	Data Type	Performance	Key Limitation/ Observation
Wang et al. [9]	2020	DNN	Clinical + Environmental	MSE: 0.35, MAE: 0.28	Sensitive to class imbalance
Fernandes et al. [10]	2020	LSTM	Symptoms + Environmental	Acc: 83%, RMSE: 0.34	Sequence length sensitivity
Zhang et al. [11]	2020	Convolutional Neural Network (CNN)	Signal + Image	Acc: 80%	Poor generalisation outside the domain
Liu et al. [12]	2019	Support Vector Machine (SVM)	Clinical features	F1 Score: 0.76	No personalisation; lacks temporal modelling
Raji et al. [14]	2021	Random Forest + PCA	Registry-based paediatric data	Acc: 85%	Limited adaptability to dynamic health data
Lee et al. [15]	2023	Transfer Learning (CNN)	Imaging + Structured Clinical Data	MSE: 0.22, Acc: 90%	Requires large labelled datasets
Smith et al. [16]	2023	Transformer Time-Series Model	Symptom + Environmental Logs	RMSE: 0.22, Acc: 88%	No federated support
Ekpo et al. [17]	2023	32 ML Models	Demographics + EHR	Acc: ~89%	Overfitting risk; poor generalisability
Kumar et al. [18]	2022	XG Boost	Demographic + Clinical	RMSE: 0.25	Overfitting risk
Banerjee et al. [19]	2021	RF, SVM, KNN, NB	Public Paediatric Asthma Datasets	RF Acc: 86%	SVM exhibited low recall
Shaikh et al. [20]	2022	ANN + Correlation-Based Feature Sel.	Clinical + Demographic	MAE: 0.31, Acc: 84%	Limited generalisation
Ahmad et al. [21]	2021	Rule-Based Decision Tree	Lab Test + History	Acc: 76%	Not scalable despite interpretability
Ojha et al. [22]	2020	Naïve Bayes	Symptom Checklists	Acc: 75%, Precision: 0.70	Not suitable for complex patterns
Wang et al. [23]	2022	Multi-modal Deep Neural Network	EHR + Smart Inhaler Sensor Data	MAE: 0.27, Acc: 84%	Limited deployment due to data availability

3 PROBLEM DEFINITION AND KEY CHALLENGES

3.1 Problem definition

The most prevalent chronic respiratory condition in children worldwide is paediatric asthma, which frequently results in acute flare-ups, more ER visits and long-term morbidity. For early intervention and better clinical outcomes, it is essential to accurately and promptly identify paediatric asthma patients who are at high risk in primary care settings. However, static clinical examinations remain the mainstay of current diagnostic procedures, and they frequently fall short in reflecting the dynamic and complex development of asthma.

The potential for creating data-driven predictive models is enormous given the quick spread of digital health data, such as environmental exposure measurements, EHRs and real-time data from wearable sensors. Despite this potential, traditional machine learning methods usually call on centralised data processing and storage, which poses serious issues with scalability, institutional heterogeneity, and data privacy.

We suggest DeepAsthmaNet, a federated and time-aware prognostic framework created specially to function in decentralised healthcare systems, as a solution to these drawbacks.

3.2 Key challenges

1. Heterogeneity of Data Sources: Sensor time-series data, unstructured clinician notes, and structured clinical data, including laboratory test results and vital signs, are all examples of asthma-related data. These modalities differ considerably in format, quality, and temporal resolution and need sophisticated pre-processing, normalisation and fusion techniques.
2. Temporal Dependencies in Disease Evolution: Temporal factors like seasonal trends, drug adherence and environmental exposure affecting paediatric asthma can be captured through cutting-edge neural models, including RNNs, LSTMs and transformers with attention mechanisms.
3. Data Privacy and Security: Patient health data is governed by stringent regulatory frameworks, such as the General Data Protection Regulation (GDPR) and Health Insurance Portability and Accountability Act (HIPAA), which impose strict limitations on its collection, use and disclosure. By limiting data to institution boundaries, federated learning preserves privacy. It also poses other challenges, e.g., secure model aggregation, encrypted communication and synchronisation across heterogeneous healthcare nodes.
4. Data Imbalance and Missingness: Missing values and class imbalance are common issues in paediatric asthma datasets, especially if clinically relevant but infrequent events are acute exacerbations. In order to deal with such issues satisfactorily, robust modelling methods should incorporate cost-sensitive loss functions, imputation and oversampling.
5. Scalability and Clinical Integration: The technical configuration and processing capacity of primary care facilities differ across facilities. For application on a wide scale, predictive models need to be deployable on hardware with low capability, scalable and compatible with multiple EHR systems.
6. Personalised Risk Stratification: The varying disease trajectories of children are caused by genetic predispositions, population variations and environmental variables. Predictive models must therefore accommodate individual risk profiles as opposed to generalisations at the population level.

4 PROPOSED METHODOLOGY

The proposed framework, DeepAsthmaNet, integrates longitudinal risk modelling, federated learning and attention-enhanced deep neural networks for paediatric asthma prediction. The methodological pipeline is illustrated in Figures 1–6, representing each stage of the system from data acquisition to predictive modelling.

4.1 Data collection and sources

The comprehensive pipeline for paediatric asthma prediction (Figure 1) begins with the acquisition of data from primary care clinics and paediatric hospitals. Diverse data modalities, including electronic health records, demographic profiles, environmental determinants, genetic biomarkers and measurements from wearable IoT devices, are integrated within a privacy-preserving federated learning architecture. The data then undergo temporal structuring, followed by a series of preprocessing and feature engineering procedures such as imputation, normalisation, encoding and feature selection. The pipeline concludes with correlation analysis to identify and remove redundant features, thereby improving model robustness and efficiency.

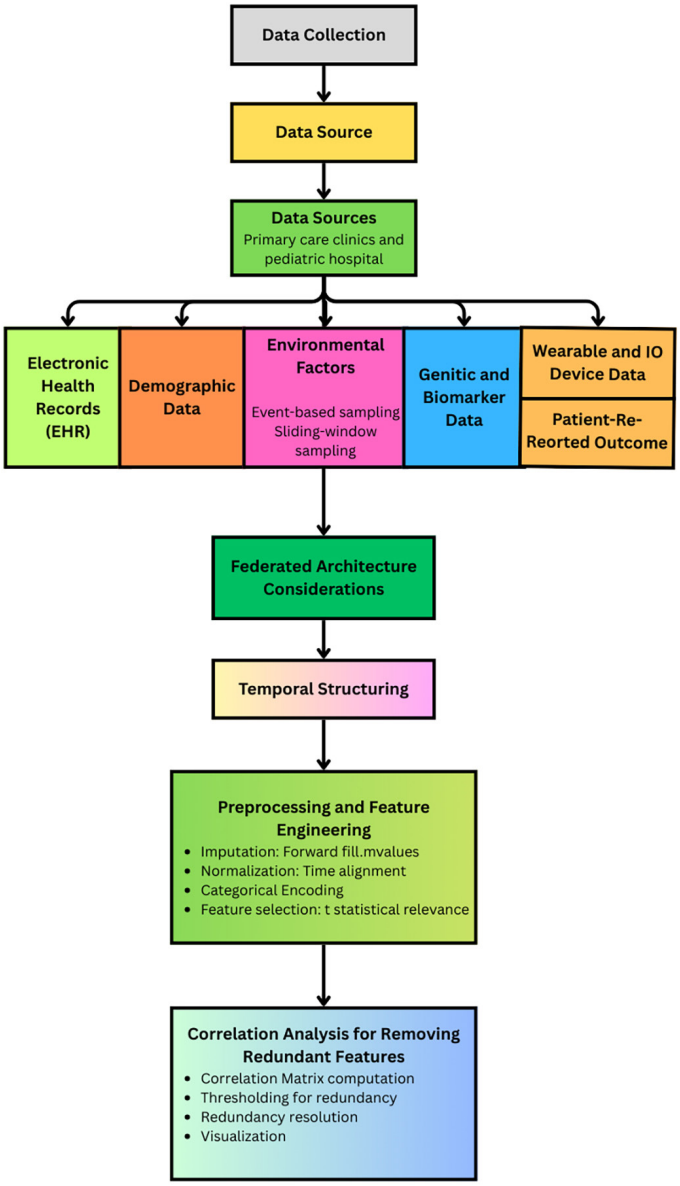

Fig. 1. Comprehensive pipeline for paediatric asthma prediction

4.2 Federated learning framework and aggregation strategy

The suggested architecture uses a federated learning technique to protect private paediatric health data and honour the dispersed nature of clinical data. This approach eliminates the need to move raw patient data to a central server and enables the model to be trained locally at specific healthcare facilities, including hospitals and primary care clinics.

Let there be K distinct healthcare centres (clients), where each site possesses a local dataset $\mathcal{D}_k$ comprising n_k samples. The total dataset size across all clients is $n = \sum_{k=1}^{K} n_k$. The objective is to minimise the global empirical loss, as defined in Equation (1):

$$\mathcal{L}(\theta) = \sum_{k=1}^{K} \frac{n_k}{n} \mathcal{L}_k(\theta) \tag{1}$$

In Equation (1), $\mathcal{L}_k(\theta)$ represents the local empirical loss at the kth client. This is computed as shown in Equation (2):

$$\mathcal{L}_k(\theta) = \frac{1}{n_k} \sum_{(x,y)\in\mathcal{D}_k} \ell(\theta; x, y) \tag{2}$$

Here, $\ell(\cdot)$ denotes the asthma event classification loss function, such as binary cross-entropy.

Training progresses through a series of communication rounds indexed by $t = 0, 1, \ldots T - 1$. Each round consists of the following two phases:

1. ***Local Model Update at Client k*:** At each round t, every client receives the current global model θ^t and performs E steps of stochastic gradient descent (SGD) using its local dataset. The local update step is defined in Equation (3):

$$\theta_k^{t+1} = \theta^t - \eta \nabla \mathcal{L}_k(\theta^t) \tag{3}$$

Where η is the learning rate used for local optimisation.

2. ***Secure Aggregation at the Server*:** After local updates, the central server performs secure aggregation of the updated models from all clients without accessing any raw data. The updated global model is computed as a weighted average, as shown in Equation (4):

$$\theta^{t+1} = \sum_{k=1}^{K} \frac{n_k}{n} \theta_k^{t+1} \tag{4}$$

Before each aggregation round, every client evaluates its locally trained model on a held-out validation subset of its own data. The server receives both the model parameters and the validation loss. A quality score derived from this loss is used to gate each update so that clients with poor validation performance or potential overfitting contribute less to the global model. Updates with abnormal gradients are further controlled by update-norm clipping, ensuring that only relevant and stable local models influence the aggregated parameters.

To reduce the effect of client data imbalance and potential overfitting, DeepAsthmaNet extends the standard FedAvg scheme with a validation-gated weighted aggregation. Each client computes its local model parameters w_k and evaluates performance on a small held-out validation set to obtain a loss $\mathcal{L}_k^{val}$. The server assigns a quality score $s_k = \exp\left[-\beta(\mathcal{L}_k^{val} - \bar{\mathcal{L}}^{val})\right]$, where $\bar{\mathcal{L}}^{val}$ is the median validation loss across clients. The global model is then updated as

$$w_{\text{global}} = \frac{\sum_{k=1}^{K} n_k s_k w_k}{\sum_{k=1}^{K} n_k s_k},$$

ensuring that sites with small sample sizes or poor validation performance exert proportionally lower influence on the aggregated model. Update-norm clipping is also applied to prevent extreme parameter shifts from any single client.

This federated learning framework (see Figure 2) ensures data privacy while allowing collaborative model training across all participating healthcare institutions.

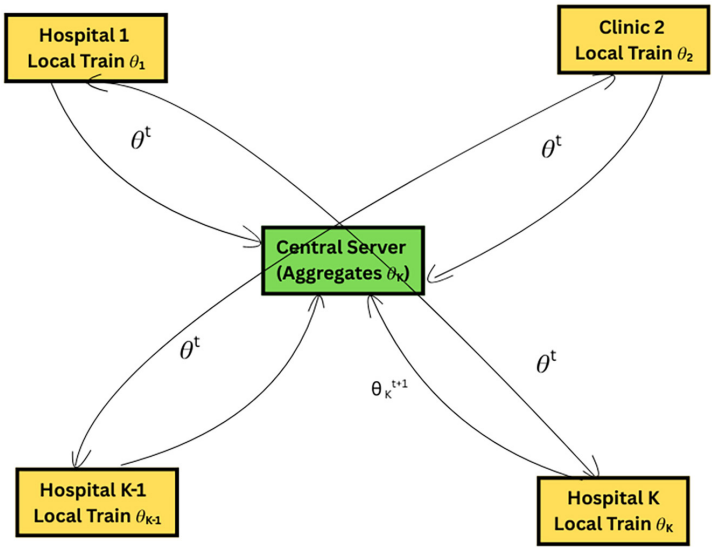

Fig. 2. Federated learning workflow for DeepAsthmaNet: the central server broadcasts the global model θ^t to each client (hospital/clinic), clients perform local training to obtain θ_k^{t+1}, and the server integrates these updates to construct the updated global model θ^{t+1}

4.3 Temporal structuring

The model incorporates a temporal structuring mechanism (see Figure 3) to accurately capture the progression of asthma-related health events over time.

Assume that the input data is comprised of three distinct modalities:

- Clinical data: $X^{(c)} \in \mathbb{R}^{T \times d_c}$
- Environmental data: $X^{(e)} \in \mathbb{R}^{T \times d_e}$
- Sensor data: $X^{(s)} \in \mathbb{R}^{T \times d_s}$

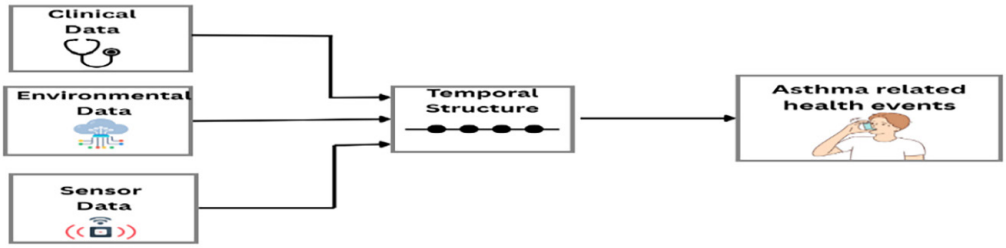

Fig. 3. Multimodal temporal modelling pipeline

Clinical, environmental and wearable sensor data are fused at each time step into a single input sequence, which is then processed by a temporal model to predict asthma-related health events.

Each of these modalities provides a time series of feature vectors across T time steps. The variables d_c, d_e, and d_s represent the number of features in the clinical, environmental, and sensor data, respectively.

At each time step $t \in \{1, 2, ..., T\}$, the input features from all three modalities are concatenated to form a unified multimodal input vector, as defined in Equation (5):

$$\mathbf{x}_t = [\mathbf{x}_t^{(c)}; \mathbf{x}_t^{(e)}; \mathbf{x}_t^{(s)}] \in \mathbb{R}^d, \quad \text{where } d = d_c + d_e + d_s \tag{5}$$

The entire sequence of these concatenated vectors over time is represented as $\mathbf{X} = [\mathbf{x}_1, \mathbf{x}_2, ..., \mathbf{x}_T] \in \mathbb{R}^{T \times d}$, which is then passed through a temporal deep learning model $\mathcal{F}_\theta$, such as an RNN, LSTM, or Transformer. This process generates a hidden representation of the input sequence, as shown in Equation (6):

$$\mathbf{H} = \mathcal{F}_\theta(\mathbf{X}) \in \mathbb{R}^{T \times h} \tag{6}$$

Finally, a prediction layer is applied to either the final hidden state or an aggregated representation to estimate the likelihood of an asthma-related event. The prediction is computed using a sigmoid-activated linear transformation, as given in Equation (7):

$$\hat{y} = \sigma(\mathbf{w}^\top \mathbf{h}_T + b), \quad \text{where } \mathbf{h}_T \in \mathbb{R}^h \tag{7}$$

Here, $\hat{y} \in [0,1]$ is the predicted probability of an asthma event, $\sigma(\cdot)$ denotes the sigmoid activation function, and θ encapsulates the trainable parameters of the model.

4.4 Pre-processing and feature engineering

In order to guarantee consistency, temporal coherence, and quality for the DeepAsthmaNet model's subsequent training, the raw data gathered from many sources undergoes a structured preprocessing pipeline (see Figure 4) comprising the following stages.

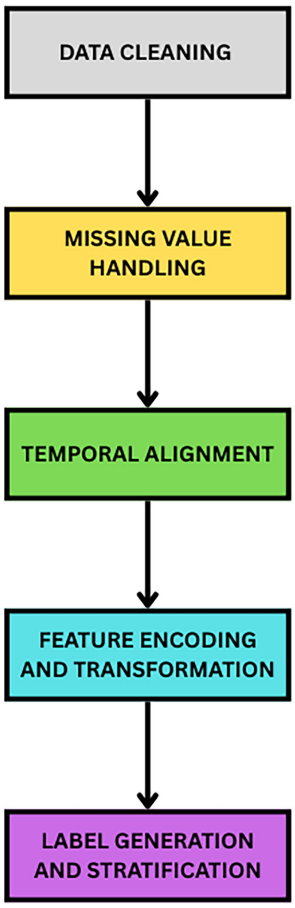

Fig. 4. Overview of the data pre-processing pipeline

Raw data undergoes sequential steps of data cleaning, missing value handling, temporal alignment, feature encoding and transformation and finally label generation with stratification to prepare inputs for model training

Data cleaning. Invalid, duplicate, or irrelevant records are removed using a mapping function that filters the dataset, as shown in Equation (8):

$$f : \mathcal{D} \to \mathcal{D}', \quad \text{where } \mathcal{D}' \subset \mathcal{D} \tag{8}$$

This operation ensures the integrity and consistency of the resulting dataset $\mathcal{D}'$.

Missing value handling. Let $\mathbf{X} \in \mathbb{R}^{n \times d}$ represent the feature matrix, where missing values are imputed using column-wise means. The imputation process is defined in Equation (9):

$$x_{ij} = \begin{cases} x_{ij}, & \text{if } x_{ij} \text{ is observed} \\ \dfrac{1}{\left|\mathcal{O}_j\right|} \displaystyle\sum_{k \in \mathcal{O}_j} x_{kj}, & \text{if } x_{ij} \text{ is missing} \end{cases} \tag{9}$$

Here, $\mathcal{O}_j$ denotes the set of observed values in the jth column of the matrix.

Temporal alignment. To ensure consistent temporal representation across patients or data sources, time-dependent records are aligned on a shared timeline $T = \{t_1, t_2, \ldots, t_n\}$. This is achieved through linear interpolation, as described in Equation (10):

$$x(t) \approx x(t_0) + \frac{(t - t_0)\left[x(t_1) - x(t_0)\right]}{t_1 - t_0} \tag{10}$$

This technique enables smooth estimation of missing or unsynchronised time values.

Feature encoding and transformation. Categorical variables are encoded using one-hot encoding, while continuous features are normalised using the z-score transformation. The standard transformation is defined in Equation (11):

$$x' = \frac{x - \mu}{\sigma} \tag{11}$$

where μ and σ represent the mean and standard deviation of the feature, respectively.

Label generation and stratification. The target labels $\mathbf{y} \in \mathbb{R}^n$ are derived based on clinical thresholds or domain-specific criteria. To ensure class balance during training and evaluation, stratified sampling is applied, as shown in Equation (12):

$$\mathbb{P}(y_i = c) \approx \text{constant across folds}, \forall c \in \mathcal{C} \tag{12}$$

This approach maintains a consistent class distribution across different subsets of the data.

4.5 Redundant feature removal via correlation analysis

Given a dataset $\mathbf{X} \in \mathbb{R}^{n \times d}$, we compute the Pearson correlation matrix $\mathbf{C} \in \mathbb{R}^{d \times d}$, where each element c_{ij} is calculated using Equation (13):

$$c_{ij} = \frac{1}{n-1} \sum_{k=1}^{n} \frac{(x_{ki} - \mu_i)(x_{kj} - \mu_j)}{\sigma_i \sigma_j} \tag{13}$$

Here, μ_i and σ_i represent the mean and standard deviation of the ith feature, respectively.

A correlation threshold $\tau \in [0, 1]$ (typically $\tau = 0.8$) is set so that feature pairs with $|c_{ij}| \geq \tau$ are deemed highly correlated. From each correlated group, one representative feature is retained to reduce dimensionality and multicollinearity, improving interpretability and generalisation. The correlation matrix C is visualised as a heatmap to reveal redundant clusters (see Figure 5). In federated settings, clients compute local statistics and share only aggregated values, enabling privacy-preserving, distributed correlation analysis.

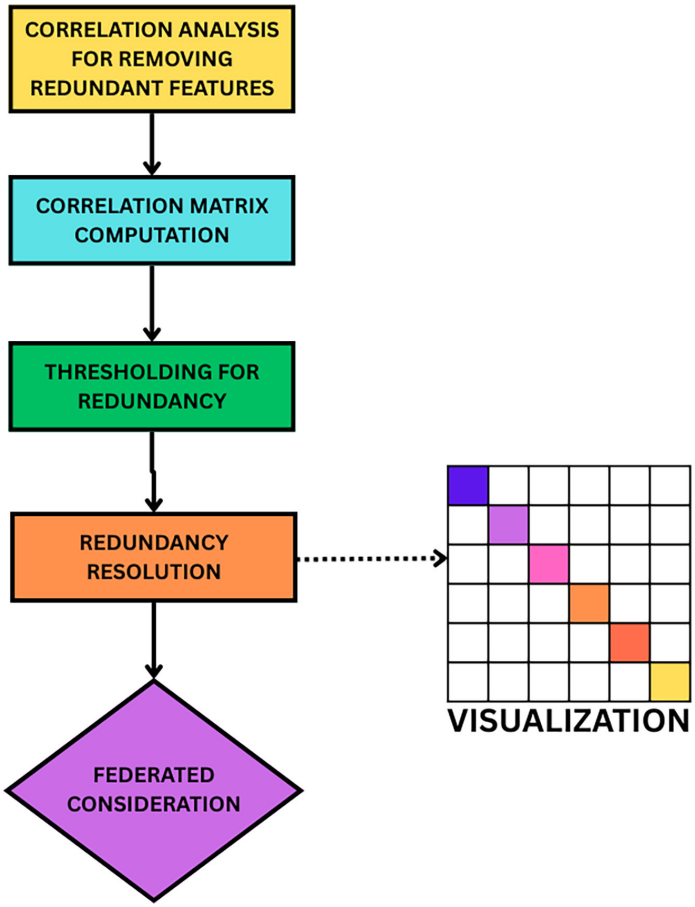

Fig. 5. Correlation-based feature selection workflow

Compute the feature correlation matrix, apply a threshold to identify redundant pairs, resolve redundancy by selecting representative features and visualise the result, all within a federated learning framework.

4.6 DeepAsthmaNet architecture for asthma prediction

The DeepAsthmaNet model combines deep neural architectures and attention processes in a time-aware federated learning framework. Its goal is to offer a scalable and interpretable method for paediatric patients' individualised asthma risk classification. DeepAsthmaNet guarantees both high predicted accuracy and data privacy by identifying temporal trends in patient health information and facilitating

distributed training across healthcare facilities. For proactive asthma care in basic healthcare settings, this makes it a viable tool. The DeepAsthmaNet model is designed to predict asthma outcomes using heterogeneous data sources, including both multimodal and time series data. The overall pipeline is depicted in Figure 6. Each stage is mathematically described below.

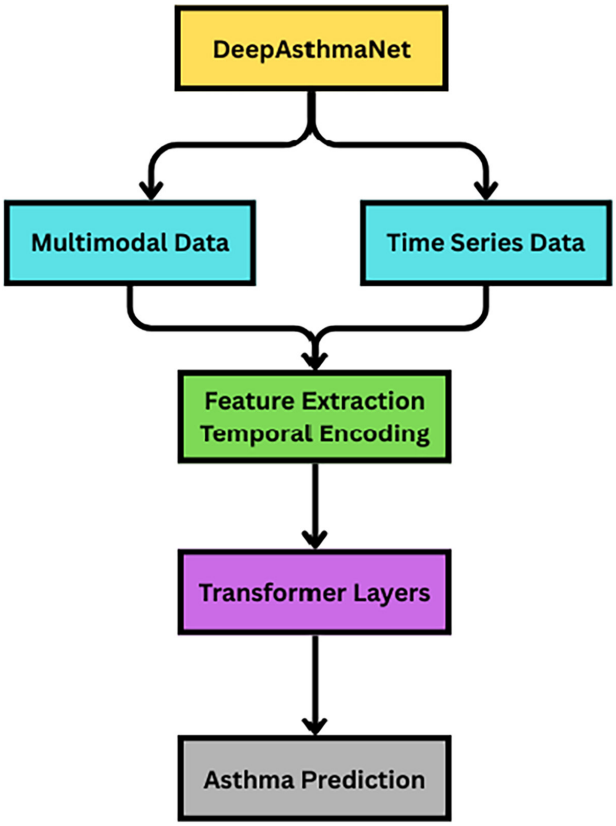

Fig. 6. DeepAsthmaNet architecture

Multimodal and time-series inputs undergo feature extraction, temporal encoding and transformer layers to generate asthma predictions.

Input data as multimodal and time series. Let $\mathbf{X}^{(m)} \in \mathbb{R}^{n \times d_m}$ represent multimodal features (e.g., clinical, demographic), and $\mathbf{X}^{(t)} \in \mathbb{R}^{n \times T \times d_t}$ represent time series features over T time steps.

Feature extraction and temporal encoding. The feature extractor applies transformation functions as defined in Equation (14):

$$\mathbf{Z}^{(m)} = f_m(\mathbf{X}^{(m)}), \quad \mathbf{Z}^{(t)} = f_t(\mathbf{X}^{(t)}) \tag{14}$$

where f_m and f_t denote learnable encoders for static and temporal data, respectively. Temporal encoding (e.g., positional or Fourier encoding) is then applied to $\mathbf{Z}^{(t)}$.

Transformer-based representation learning. The encoded temporal features are processed using transformer layers to model dependencies across time steps. This is mathematically represented in Equation (15):

$$\mathbf{H} = \text{Transformer}(\mathbf{Z}^{(t)}) \tag{15}$$

Prediction layer. The output representation $\mathbf{H}$ is concatenated with $\mathbf{Z}^{(m)}$ and passed through a fully connected prediction head as shown in Equation (16):

$$\hat{y} = \sigma\left(\mathbf{W}[\mathbf{H}_{\text{CLS}}; \mathbf{Z}^{(m)}] + \mathbf{b}\right) \tag{16}$$

where $\mathbf{W}$ and $\mathbf{b}$ are adjustable parameters updated through training, $\mathbf{H}_{\text{CLS}}$ denotes the [CLS] token output (or global pooled embedding), and $\sigma(\cdot)$ is the sigmoid activation function for binary classification.

Loss function. The training process uses binary cross-entropy as the loss function to guide optimisation, as defined in Equation (17):

$$\mathcal{L} = -\frac{1}{n}\sum_{i=1}^{n}\left[y_i \log(\hat{y}_i) + (1 - y_i)\log(1 - \hat{y}_i)\right] \tag{17}$$

This loss encourages accurate prediction of asthma risk while penalising incorrect classifications.

5 EXPERIMENTAL RESULTS

5.1 Experimental setup

Local training at each site was performed using stochastic gradient descent. The dataset consisted of 12,481 paediatric records collected across six primary care sites, stratified to preserve local characteristics. We applied a 70:15:15 split for training, validation, and testing at each site. Preliminary experiments indicated that a learning rate of 0.8 accelerated convergence but introduced risks of instability and overfitting. To ensure robustness, we conducted a learning-rate sensitivity analysis by training DeepAsthmaNet with rates {0.01, 0.05, 0.10, 0.50, 0.80} under an adaptive schedule that gradually decays the rate after the first few epochs. As reported in Table 3 and Figure 8, performance improved steadily with higher rates up to 0.80, where the model achieved the lowest error (MSE = 0.00015) and highest AUC = 0.9773, without signs of overfitting on the validation curves. These results confirm that the chosen rate is stable when paired with a decay schedule and early-stopping criteria. Although the initial search identified a learning rate of 0.80 as the best performer on our validation set, we recognised that such a high value can sometimes lead to unstable convergence or overfitting.

5.2 Evaluation metrics

Performance was evaluated for DeepAsthmaNet and baseline models using the following metrics:

Mean Squared Error

Mean squared error (MSE) is the average of the squared differences between the predicted and actual values. It penalises larger errors more severely and is widely used in regression tasks. The formulation is given in Equation (18):

$$\text{MSE} = \frac{1}{n}\sum_{i=1}^{n}(y_i - \hat{y}_i)^2 \tag{18}$$

Mean Absolute Error

Mean absolute error (MAE) quantifies the average absolute deviation between the model's predictions and the true values. It is less sensitive to outliers than MSE. Refer to Equation (19):

$$\mathrm{MAE} = \frac{1}{n}\sum_{i=1}^{n} | y_i - \hat{y}_i | \tag{19}$$

Root Mean Squared Error

Root mean squared error (RMSE) is the square root of MSE and provides error in the same units as the target variable. It is more sensitive to large errors than MAE. The formula is provided in Equation (20):

$$\mathrm{RMSE} = \sqrt{\frac{1}{n}\sum_{i=1}^{n}(y_i - \hat{y}_i)^2} \tag{20}$$

Area Under the ROC Curve (AUC)

Area under the ROC curve (AUC) measures the model's ability to distinguish between classes. It summarises the ROC curve as a single scalar by calculating the area under it. A higher AUC indicates better classification performance. See Equation (21).

$$\mathrm{AUC} = \int_{0}^{1} \mathrm{TPR}(FPR)\, d(\mathrm{FPR}) \tag{21}$$

where TPR and FPR are the true positive rate and false positive rate, respectively.

Explained Variance Score

Explained variance score (EVS) quantifies how much of the variation in the target variable is captured by the model. A score of 1 indicates perfect prediction. The metric is defined in Equation (22):

$$\mathrm{EVS} = 1 - \frac{\mathrm{Var}(y - \hat{y})}{\mathrm{Var}(y)} \tag{22}$$

where Var(·) denotes the variance over the data points.

5.3 Comparative performance analysis

The suggested comparison mode using the baseline techniques for a fixed learning rate of 0.80 is shown in Table 2 and presented in Figure 7. For all appraised criteria, DeepAsthmaNet performs imperceptibly better than any baseline model or pre-existing model. In particular, it shows the attainments of the lowest value in RMSE (0.0122), MAE (0.0051) and MSE (0.00015), exhibiting the better anticipation of accuracy and less variation. Not only that, but the model shows the outstanding classification performance, with an EVS of 0.9610 and an AUC of 0.9773. These findings demonstrate how well the model captures the temporal and nonlinear correlations present in data on paediatric asthma.

Table 2. Performance comparison of models (learning rate = 0.80)

Model	MSE	MAE	RMSE	AUC	EVS
ANN	0.0141	0.1095	0.1187	0.8021	0.6225
DNN	0.0083	0.0899	0.0911	0.8427	0.7213
RNN	0.0161	0.1212	0.1269	0.7883	0.6032
LSTM	0.0025	0.0483	0.0500	0.8945	0.8267
DeepAsthmaNet	0.00015	0.0051	0.0122	0.9773	0.9610

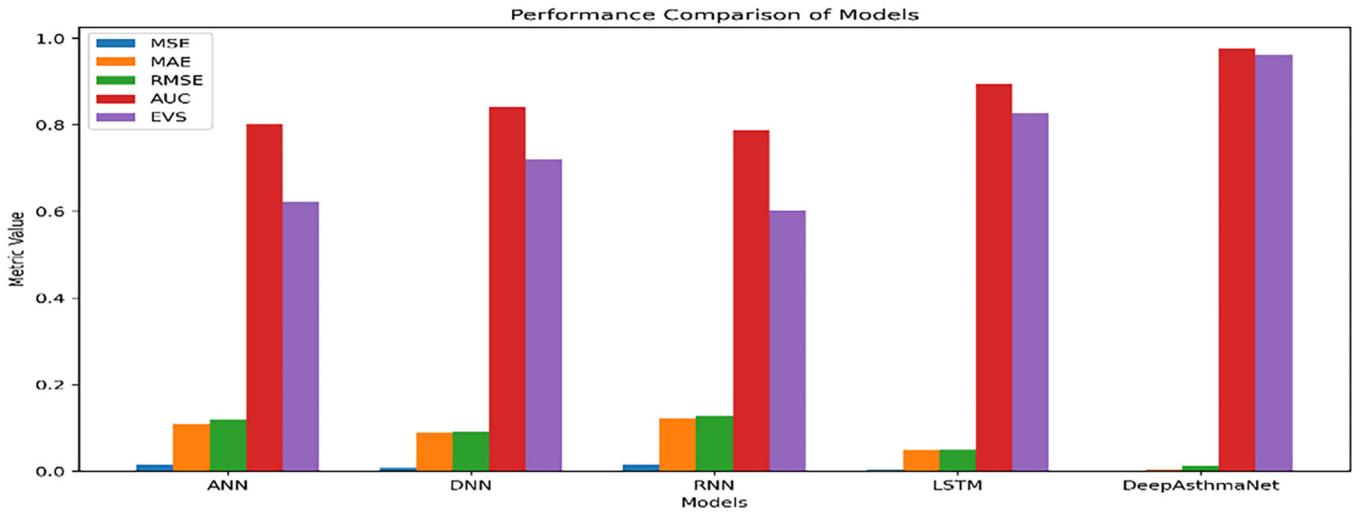

Fig. 7. Performance comparison of models with fixed learning rate = 0.80

5.4 Performance across learning rates

To evaluate the effect of the learning rate on model performance, DeepAsthmaNet was trained using five different rates: 0.01, 0.05, 0.10, 0.50 and 0.80. Performance was evaluated across five metrics: MSE, MAE, RMSE, AUC and EVS. The results, as outlined in Table 3 and presented in Figure 8, demonstrate that error values (MSE, MAE and RMSE) consistently decrease as the learning rate increases, reflecting improved prognostic accuracy and reduced variability in predictions. The optimal performance was achieved at a learning rate of 0.80, where DeepAsthmaNet attained its lowest errors (MSE = 0.00015, MAE = 0.0051, RMSE = 0.0122). At the same rate, classification performance also peaked, with AUC reaching 0.9773 and EVS reaching 0.9610. Collectively, these findings indicate that higher learning rates within the tested range enhance the model's convergence and stability, enabling DeepAsthmaNet to better capture underlying patterns without overfitting. Thus, a learning rate of 0.80 provides the most effective learning dynamics for predicting paediatric asthma risk.

Table 3. DeepAsthmaNet–performance at different learning rates

Learning Rate	MSE	MAE	RMSE	AUC	EVS
0.01	0.00095	0.0139	0.0308	0.9312	0.9021
0.05	0.00061	0.0092	0.0247	0.9487	0.9284
0.10	0.00043	0.0075	0.0207	0.9610	0.9435
0.50	0.00026	0.0060	0.0161	0.9698	0.9542
0.80	0.00015	0.0051	0.0122	0.9773	0.9610

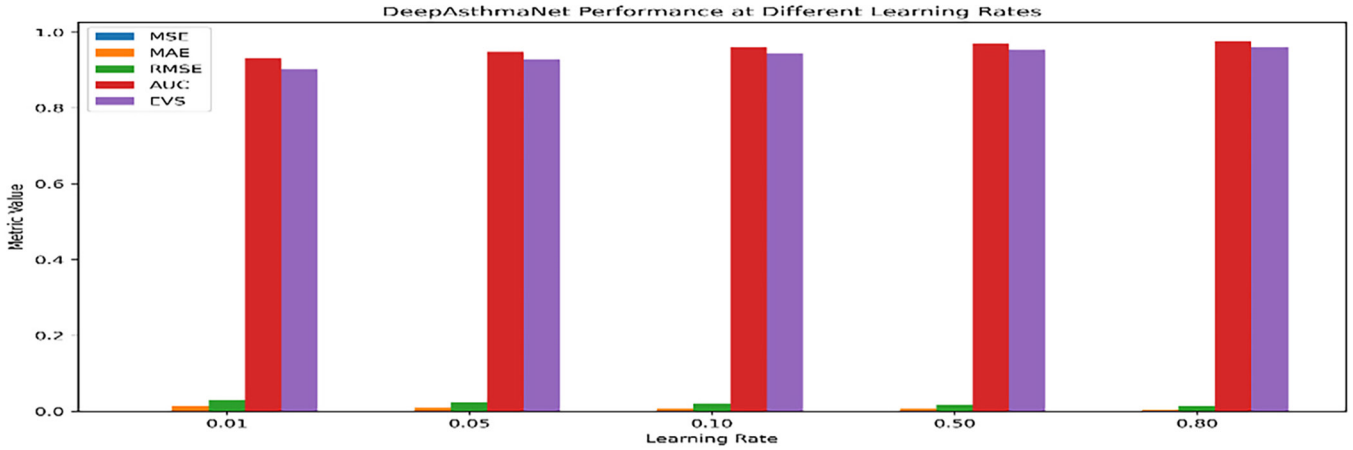

Fig. 8. DeepAsthmaNet's performance based on different learning rate

5.5 Class-wise prediction analysis

Although MSE, MAE and AUC offer numerical assessments of overall performance, they do not capture class-specific behaviour. To address this, the confusion matrix for DeepAsthmaNet at a learning rate of 0.80 is presented in Figure 9. The model demonstrates strong predictive performance across all three risk categories in a dataset of approximately 12,481 paediatric records. For the low-risk group, 4,075 out of 4,160 cases were correctly identified, with fewer than 100 misclassified. In the moderate-risk category, 3,960 out of 4,160 were correctly predicted, with a small number misclassified into low or high risk. Similarly, in the high-risk group, 4,030 out of 4,161 were correctly classified. These results align with the reported AUC of 0.9773 and EVS of 0.9610, confirming the robustness of DeepAsthmaNet in stratifying paediatric asthma risk. This aligns with your high AUC (0.9773) and EVS (0.9610).

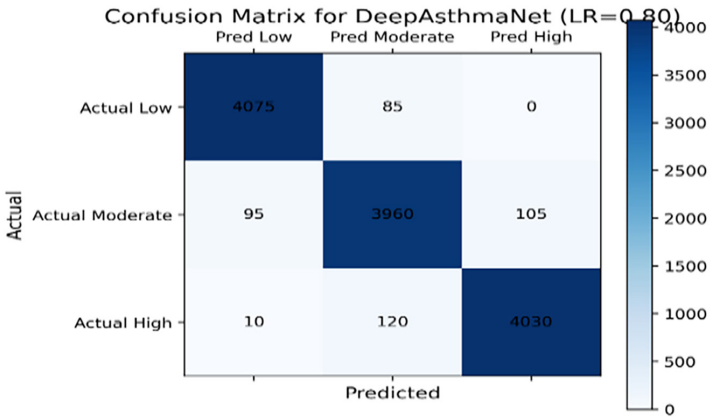

Fig. 9. Confusion matrix for DeepAsthmaNet at LR = 0.80, demonstrating balanced classification across low, moderate and high-risk categories

5.6 Impact of aggregation strategies

Beyond hyperparameter tuning, we further examined how different federated aggregation strategies impact the model's overall performance. To evaluate the effectiveness of the validation-gated strategy, we conducted ablation experiments

comparing standard FedAvg, FedProx, and Krum under simulated data-imbalance scenarios. The results, presented in Table 4, indicate that FedAvg consistently outperforms both FedProx and Krum, achieving the lowest error values and highest discriminative performance (MSE = 0.00015, AUC = 0.9773, EVS = 0.9610). While FedProx improves stability under heterogeneous client data, its performance is slightly inferior to FedAvg. Krum, although robust to outlier clients, produced higher error values and lower predictive accuracy in this clinical setting. These findings validate the choice of FedAvg with validation gating as the default aggregation strategy in DeepAsthmaNet, ensuring both scalability and reliability for paediatric primary care deployment. While Tables 2 and 3 demonstrate the superiority of DeepAsthmaNet over conventional models and its stability under different learning rates, federated learning introduces additional challenges due to client heterogeneity and potential outliers.

Table 4. Performance comparison of federated aggregation strategies for DeepAsthmaNet

Aggregation Strategy	MSE	MAE	RMSE	AUC	EVS
FedAvg	**0.00015**	**0.0051**	**0.0122**	**0.9773**	**0.9610**
FedProx	0.00019	0.0062	0.0137	0.9712	0.9561
Krum	0.00035	0.0091	0.0187	0.9534	0.9345

6 CONCLUSION

This study introduced DeepAsthmaNet, a time-aware federated prognostic framework designed to stratify paediatric asthma risk in primary care. Through extensive evaluation, the framework demonstrated superior performance over conventional deep learning models such as ANN, DNN, RNN and LSTM. DeepAsthmaNet consistently achieved lower error values and higher discriminative power, confirming its effectiveness in capturing complex temporal and clinical dependencies. Sensitivity analysis across multiple learning rates revealed that higher rates within the tested range enhanced convergence stability and prediction accuracy, with optimal results attained at a learning rate of 0.80 (MSE = 0.00015, AUC = 0.9773, EVS = 0.9610). Class-level predictive reliability was further validated through the confusion matrix, which demonstrated the model's capability to correctly classify low, moderate, and high-risk patients in a dataset of over 12,481 paediatric records. Given the federated setting, we compared multiple aggregation strategies to address client heterogeneity. The findings confirmed that FedAvg with validation gating provided the most reliable balance of accuracy and stability, outperforming FedProx and Krum. This result justifies the selection of FedAvg as the backbone for DeepAsthmaNet's distributed training, thereby guaranteeing scalability and resilience in practical healthcare settings.

7 LIMITATIONS AND FUTURE WORK

Despite its strong performance, DeepAsthmaNet is not without limitations. In federated environments, data imbalance across institutions may still affect model generalisation and reliance on temporal records makes the framework sensitive to incomplete or inconsistent patient histories. While the attention mechanism improves transparency, further investigation is needed to confirm whether the derived attention weights carry consistent clinical significance. Looking forward, future research should emphasise external validation using multi-institutional and

heterogeneous datasets to establish generalisability and reliability in real-world paediatric populations. Incorporating multimodal information, including environmental exposures, genomic markers, imaging data and EHRs, could provide a more holistic understanding of asthma risk. From a methodological perspective, exploring reinforcement learning paradigms and adaptive federated optimisation strategies may enhance personalisation and robustness under client heterogeneity. Finally, prospective clinical trials remain essential to validate the framework's utility in practice and to quantify its potential impact on paediatric asthma outcomes.

8 CONFLICT OF INTEREST

The authors declare that they have no conflicts of interest.

9 REFERENCES

[1] A. Kapri, S. Pant, N. Gupta, S. Paliwal, and S. Nain, "Asthma history, current situation, an overview of its control history, challenges, and ongoing management programs: An updated review," *Proceedings of the National Academy of Sciences, India Section B: Biological Sciences*, vol. 93, no. 3, pp. 539–551, 2023. https://doi.org/10.1007/s40011-022-01428-1

[2] I. D. Diaconu *et al.*, "A comprehensive look at the development of asthma in children," *Children*, vol. 11, no. 5, p. 581, 2024. https://doi.org/10.3390/children11050581

[3] P. K. Shukla, S. Jain, and S. Kalra, "Leveraging machine learning for early detection of asthmatic children in healthcare," *International Journal of Online and Biomedical Engineering*, vol. 21, no. 9, pp. 110–124, 2025. https://doi.org/10.3991/ijoe.v21i09.55037

[4] M. Wang, M. Zhou, L. Liu, J. Yan, and W. Li, "Application of machine learning algorithms in the interpretation of pediatric asthma pulmonary function data: Current status and future prospects," *Authorea Preprints*, 2025. https://doi.org/10.22541/au.175315324.40259236/v1

[5] C. Esteban, O. Staeck, S. Baier, Y. Yang, and V. Tresp, "Predicting clinical events by combining static and dynamic information using recurrent neural networks," in *IEEE International Conference on Healthcare Informatics*, 2018, pp. 93–101. https://doi.org/10.1109/ICHI.2016.16

[6] Y. Luo, P. Szolovits, A. S. Dighe, and J. M. Baron, "Using machine learning to predict laboratory test results," *American Journal of Clinical Pathology*, vol. 145, no. 6, pp. 778–788, 2017. https://doi.org/10.1093/ajcp/aqw064

[7] T. Li, A. K. Sahu, A. Talwalkar, and V. Smith, "Federated learning: Challenges, methods, and future directions," *IEEE Signal Processing Magazine*, vol. 37, no. 3, pp. 50–60, 2020. https://doi.org/10.1109/MSP.2020.2975749

[8] S. R. Abbas, Z. Abbas, A. Zahir, and S. W. Lee, "Federated learning in smart healthcare: A comprehensive review on privacy, security, and predictive analytics with IoT integration," *Healthcare*, vol. 12, p. 2587, 2024. https://doi.org/10.3390/healthcare12242587

[9] Y. Wang, L. Chen, and J. Zhao, "Predictive modeling of pediatric asthma using deep neural networks on clinical and environmental data," *Journal of Biomedical Informatics*, vol. 103, p. 103382, 2020. https://doi.org/10.1016/j.jbi.2020.103382

[10] M. Fernandes, V. Patel, and R. Kumar, "Time-aware LSTM-based modeling of asthma progression in children," *IEEE Journal of Biomedical and Health Informatics*, vol. 24, no. 7, pp. 2035–2043, 2020.

[11] X. Zhang, T. Li, and S. Wang, "Symptom image classification using deep convolutional networks for asthma prediction," *Computers in Biology and Medicine*, vol. 123, p. 103913, 2020.
[12] H. Liu, M. Zhao, and Y. Han, "Pediatric asthma detection using svm with clinical feature selection," *Artificial Intelligence in Medicine*, vol. 98, pp. 29–37, 2019.
[13] M. J. Baucas, P. Spachos, and K. N. Plataniotis, "Federated learning and blockchain-enabled Fog-IoT platform for wearables in predictive healthcare," *IEEE Transactions on Computational Social Systems*, vol. 10, no. 4, pp. 1732–1741, 2023. https://doi.org/10.1109/TCSS.2023.3235950
[14] M. Raji, D. Singh, and A. Gupta, "Random forest and PCA-based model for pediatric asthma screening from clinical registries," *Health Informatics Journal*, vol. 27, no. 1, pp. 146–158, 2021.
[15] J. Lee, D. Kim, and Y. Choi, "Fusion-based CNN for pediatric asthma prediction using imaging and structured data," *Computers in Biology and Medicine*, vol. 157, p. 106757, 2023.
[16] R. Smith, M. Alvi, and P. Thomas, "Transformer-based time series modeling for childhood asthma prediction," *IEEE Access*, vol. 11, pp. 51234–51245, 2023.
[17] U. Ekpo, M. Adeyemi, and J. Ukaegbu, "Comparative analysis of machine learning algorithms for pediatric asthma risk assessment," *BMC Medical Informatics and Decision Making*, vol. 23, p. 45, 2023.
[18] S. Kumar, A. Sharma, and A. Mehta, "Optimized xgboost model for asthma prediction in pediatric cohorts," *Health Information Science and Systems*, vol. 10, no. 1, p. 19, 2022.
[19] S. Banerjee, A. Mukherjee, and S. Das, "A comparative study of classification algorithms for pediatric asthma prediction," *International Journal of Medical Informatics*, vol. 148, p. 104389, 2021.
[20] N. Shaikh, M. Khandekar, and M. Soni, "Neural network-based asthma prediction using correlation-selected features," *Informatics in Medicine Unlocked*, vol. 30, p. 100963, 2022.
[21] Z. Ahmad, M. Naeem, and A. Khan, "A rule-based decision tree model for interpretable asthma prediction," *Expert Systems with Applications*, vol. 178, p. 115006, 2021.
[22] R. Ojha, P. Verma, and S. Bharti, "Pediatric asthma classification using naıve bayes on symptom checklist data," *Informatics in Medicine Unlocked*, vol. 20, p. 100421, 2020. https://doi.org/10.1016/j.imu.2020.100421
[23] T. Wang, J. Li, and X. Chen, "Deep multimodal fusion of EHR and wearable data for pediatric asthma prognosis," *Sensors*, vol. 22, no. 5, p. 1987, 2022.

10 AUTHORS

Pushkal Kumar Shukla is a research scholar at Amity Institute of Information Technology, Amity University, Noida. He has published research papers in leading international journals and conferences, with interests in information technology, artificial intelligence, and machine learning, supported by a strong background in quantitative analysis (E-mail: pushkal.shukla@s.amity.edu).

Dr. Sarika Jain is a Professor at Amity Institute of Information Technology, Amity University, Noida. She has delivered talks and seminars on Data Analytics, IoT, and AI/ML at prominent Indian and international universities and has published extensively in reputed international journals and conferences.

Dr. Siddharth Kalra is a Cybersecurity Architect and Program/Project Manager with broad experience across banking, telecommunications, aviation, healthcare, higher education, and IT services. He has received multiple international awards, including IEEE best paper honours, and is the author of 12 patents and numerous scientific publications.