Journal of Artificial Intelligence in Governance and Public Policy (JAIGPP)
Volume 1 | Issue 3 | July- September 2026 | Pages 01-07
DOI: https://doi.org/xxxxxxxxxxxxx
e-ISSN: 3139-8693
Research Article OPEN ACCESS
Available online: https://jaigpp.org/index.php/j

Artificial Intelligence-Based Clinical Risk Prediction Using Ensemble Machine Learning Algorithms

Ms. Shivani Jain

Assistant Professor, Department of Computer Science and Applications , Mandsaur

madhav.sharma@meu.edu.in

Corresponding Author: *Shivani Jain, (madhav.sharma@meu.edu.in)

Received: 10 July, 2026 | **Revised**: 1 August, 2026 | **Accepted**: 5 August, 2026

***Abstract*—Risk prediction tools are becoming increasingly popular tools to assist in making clinical decisions. The models, however, are typically trained on data from general patient cohorts and may not be representative of and applicable to targeted patient cohorts when used in practice. This study overcame these obstacles by developing and evaluating a clinical risk prediction model using the MIMIC-III clinical dataset and an Artificial Neural Network (ANN). The suggested method accounts for all possible preprocessing steps—including normalization, encoding, managing missing values, and data balancing using SMOTE-ENN—to increase the dependability of the predictions. With an F1-Score (F1) of 96.9%, an accuracy (ACC) of 98.6%), a precision (PRE) of 97%, and a recall (REC) of 96.5%, the ANN model has high predictive capacity and can capture the complicated interaction between the clinical variables. The outcomes demonstrate that the suggested ANN architecture can accurately and consistently assess clinical risk, which in turn allows for the early identification of high-risk patients and aids healthcare providers in making data-informed therapeutic decisions. Because it can provide trustworthy prediction models from complex health data, the proposed method has great potential for clinical real-time application. It can also help develop smarter healthcare decision-support systems and enhance patient monitoring methods.**

***Keywords*—Clinical Risk Prediction, Artificial Intelligence, Explainable AI, Model Deployment, Clinical Decision Support Systems.**

I. INTRODUCTION

The enormous amounts of patient data collected by healthcare systems come from a variety of sources, including electronic health records, laboratory tests, clinical encounters, medical imaging, and even wearable devices. These medical data offer helpful information on disease progression, high-risk patients and clinical decision-making [1][2]. Hence, clinical risk prediction has become a pivotal element of contemporary healthcare, facilitating early detection of unfavorable clinical outcomes, including disease progression, complication rates, hospital re-admissions, and mortality rates. Healthcare data are, however, often heterogeneous, incomplete, noisy and imbalanced, which makes accurate and reliable clinical risk prediction challenging [3][4].

Clinical Risk Prediction Models help clinicians by calculating the probability that medical events occur in the future from demographic, physiological, laboratory and clinical data. Traditional prediction approaches have been useful for evidence-based medicine, but they can't handle the high dimensionality, nonlinear correlations, missing data, and temporal dependencies that are present in real-world clinical data [5]. Also, most current prediction models are designed for particular diseases or healthcare scenarios, limiting adaptability and scalability across various healthcare institutions and patient groups [6][7].

The advent of digital healthcare has made previously inaccessible quantities of clinical data suitable for predictive modeling [8][9][10]. While these advancements are encouraging, there are still several obstacles that need to be overcome. These include issues with data quality, missing clinical information, class imbalance, limited generalizability across healthcare settings, lack of external validation, and openness in clinical decision-making [11]. The restrictions underlined the need for building strong and reliable clinical risk prediction tools fit to be applied to the day-to-day practice of healthcare.

Clinical risk prediction has been greatly impacted by the rise of AI, which allows for smarter analysis of complicated healthcare data [12][13]. In the context of large-scale and diverse healthcare data, ML approaches can identify internal patterns in clinical data to enable accurate and rapid risk assessment, while DL can enhance prediction by capturing complex data representations. There are still problems with clinical integration, interpretability, fairness, and external validation, even if these technologies have improved prediction performance [14][15]. Therefore, developing reliable AI-based clinical risk prediction frameworks is

DOI: https://doi.org/xxxxxxxxxxxxx

essential for supporting effective healthcare decision-making and improving patient outcomes.

A. Motivation and Contribution

The main goal of this research is to create a clinical risk prediction system that can reliably identify high-risk people early on. Large-scale healthcare data, like electronic health records and clinical data, also offers opportunities to advance predictive analysis, although some obstacles, like data heterogenicity, missing values, noise, and class imbalance, hinder the ability of existing predictive analysis techniques. Therefore, this study aims to utilize Artificial Intelligence-based methods to analyze complex clinical patterns, enhance prediction ACC, and provide a robust decision-support solution for improving healthcare outcomes. The study's most significant contributions are as follows:

- Analyzed patient clinical records and found significant risk-related patterns for prediction using the MIMIC-III clinical dataset, a sizable de-identified ICU database.
- Developed an artificial intelligence-based clinical risk prediction framework for identifying high-risk patients using healthcare data.
- Designed a clinical risk prediction model using ANNs to accurately assess patients' risks by capturing the complicated nonlinear interactions among their clinical parameters.
- Evaluated the proposed clinical risk prediction framework using REC, ACC, PRE, and F1, which are widely used performance metrics in the industry.
- Assisted in creating healthcare decision-support systems that use artificial intelligence to do individualized risk assessments for patients.

B. Organization of the Paper

The paper is organized as follows: Section II reviews the literature on clinical risk prediction and AI-based methods. Section III, goes over the dataset, data pretreatment methods, and the steps to build the model. In Section IV, offers the experimental findings, assesses the performance, and compares them to existing approaches. Section V concludes by reviewing the study's key findings and discussing possible avenues for further investigation.

II. Literature Review

Existing research works on Clinical Risk Prediction were analyzed in detail to find out current advancements, methodologies, and limitations that shaped the proposed framework. Table I summarizes recent studies of Clinical Risk Prediction, including the models used, the datasets on which they were applied, the findings, and the challenges.

S. Gupta and S. Ahmed (2026) explained a clinical decision support system that uses interpretable ML to risk-stratify patients with early cardiac disease based on structured clinical information. With an ACC of 93.2, PRE of 93, REC of 92, F1 of 92, and AUC of 0.96, the experimental results show that the XGBoost model has very good performance. This confusion matrix has low false-negative rates, according to the analysis, which makes it useful for clinical screening [16].

N. G. Kamdi et al. (2026) provided a methodology for early CKD identification and risk assessment utilizing extensive clinical data that is based on ML. A collection of real-world CKD data was used to train and test a battery of prediction models, including LR, SVM, RF, GB, and DNNs. With ROC-AUC values surpassing 0.96, ACC levels more than 94%, and high PRE and REC, several ensemble models performed better, with RF and GB achieving the best results. Results from feature significance and SHAP analyzes showed that the most important factors in determining CKD are blood pressure, hemoglobin, serum creatinine, and urine protein levels. These findings provide clear and practical information about the factors that put people at risk for CKD [17].

A. Kumar et al. (2025) evaluated XGBoost, RF, SVM, and LR, four ML algorithms, making use of ROC-AUC, F1, REC, ACC, and PRE criteria. With interpretability based on SHAP revealing important factors including age, systolic blood pressure, blood glucose, and blood pressure, XGBoost got the best ACC (90.3%) and AUC (0.937). The paper highlights the need for privacy protection and fairness-aware learning in clinical decision assistance, and also highlights the promise of transparent ML models. The next phase involves real-time integration of multimodal data and federated learning [18].

A. Raghav et al. (2025) prove that many classification methods and computational learning are used for CVD prediction. Model training and testing were conducted using the heart and blood vessel datasets. The following measures were used to assess performance: AUC, F1, REC, ACC, and PRE. The ensemble model achieved the best results compared to the other models because of its high area under the curve (AUC) of 0.87 and ACC level of 80.2%. The findings show that better and more reliable tools for assessing the risk of cardiovascular disease may be achieved by combining multiple learning algorithms. Increased diagnostic ACC and prediction reliability are outcomes of this approach [19].

F. Shahidi et al. (2024) built an EnH-DL, an ensemble hybrid DL model, to outperform existing DL models and decrease mistakes on unseen data. Results demonstrated a 2.5% improvement in AUC and a sensitivity improvement of more than 1.5% on average. Unseen data is better represented by a more reliable model when the error reduction is more than 0.22. The model was evaluated on a dataset that was highly imbalanced and attained an AUC of 83% and a sensitivity of 79% for FHE-H, all while mimicking real-world settings with a clinical demand window. Overall, when comparing current models for various MH outcomes utilizing Admin-HD, EnH-DL proved to be the superior choice [20].

C. Kaur and R. Madaan (2024) gather patient data, purge it of unnecessary information to reduce its dimensionality and clean it up to standardize its attributes before dividing it into a training set and a testing set. Breast cancer recurrence can be detected using a variety of diagnostic procedures, including physical exams, ultrasounds, blood tests, and mammography. To improve breast cancer prediction, the planned study uses both the extracted characteristics and the identified risk variables. An evaluation and validation of the recommended work is conducted using the Breast Cancer Surveillance Consortium Dataset. At 75.2% ACC, Random Forest produced the greatest outcomes [21].

D. Rai et al. (2023) looked into how well ensemble learning and conventional ML worked for forecasting cardiac problems. Experimentally evaluate these models' performance on a broad dataset including clinical, demographic, and lifestyle characteristics using a number of performance measures such cross-validation mean ACC, REC, PRE, F1,

DOI: https://doi.org/xxxxxxxxxxxxx

and ROC - AUC Score. At 89.71% REC and 88.69% total cross-validation ACC, Gradient Boost is the clear winner. Improving the management of cardiovascular disease and lowering the mortality rates that are connected with it are the overarching goals of this study [22].

Existing Work Gap: The use of AI and ensemble ML algorithms for clinical risk prediction utilizing varied healthcare datasets has been shown beneficial in recent research. These methods have led to good disease prediction, risk assessment, and clinical decision support, but they are limited by being inaccurate and/or difficult to understand. They have been used in existing methods to make predictions, assess risk, and support clinical decisions with good ACC and interpretability. However, challenges related to data imbalance, complexity of clinical features, generalization, and real-time applicability still remain. Therefore, developing robust AI-based clinical risk prediction frameworks is essential for improving prediction reliability and supporting effective healthcare decision-making.

TABLE I. REVIEW SUMMARY OF EXISTING CLINICAL RISK PREDICTION APPROACHES

Author	Proposed Work	Dataset	Results	Limitations & Future Work
S. Gupta and S. Ahmed (2026)	Developed an interpretable ML-based clinical decision support system for early heart disease risk prediction using XGBoost.	Structured clinical heart disease dataset	XGBoost achieved 93.2% accuracy, 93% precision, 92% recall, 92% F1-score, and AUC = 0.96 with a low false-negative rate.	Limited to structured clinical data. Future work should incorporate multimodal patient data and external clinical validation.
N. G. Kamdi et al. (2026)	Proposed a CKD prediction framework using Logistic Regression, SVM, Random Forest, Gradient Boosting, and DNN with SHAP-based explainability.	Real-world Chronic Kidney Disease (CKD) dataset	Random Forest and Gradient Boosting achieved >94% accuracy and ROC-AUC >0.96. SHAP identified key CKD biomarkers.	Requires validation on larger multi-center datasets. Future work includes real-time clinical deployment and integration with EHR systems.
A. Kumar *et al.*, (2025)	Compared Logistic Regression, Random Forest, SVM, and XGBoost for clinical risk prediction with SHAP interpretability.	Clinical healthcare dataset	XGBoost achieved 90.3% accuracy and AUC = 0.937. Important predictors included glucose, BMI, blood pressure, and age.	Limited by centralized data. Future work focuses on federated learning, fairness-aware AI, and multimodal healthcare data integration.
A. Raghav *et al.*, (2025)	Evaluated multiple ML classifiers and ensemble learning for cardiovascular disease prediction.	Heart and blood vessel datasets	Ensemble model achieved 80.2% accuracy and AUC = 0.87, outperforming individual models.	Moderate predictive performance. Future work should investigate advanced ensemble and deep learning techniques with larger datasets.
Shahidi et al. (2024)	Proposed an Ensemble Hybrid Deep Learning (EnH-DL) model for mental health outcome prediction.	Admin-HD dataset	Improved AUC by 2.5%, sensitivity by >1.5%, and reduced prediction error by 0.22. Achieved 83% AUC and 79% sensitivity on imbalanced data.	Requires validation across different healthcare populations. Future work includes improving generalization and handling highly imbalanced datasets.
C. Kaur and R. Madaan (2024)	Proposed breast cancer recurrence prediction using feature extraction and multiple ML algorithms.	Breast Cancer Surveillance Consortium dataset	Random Forest achieved the highest 75.2% accuracy.	Prediction accuracy remains relatively low. Future work should explore deep learning models and incorporate additional clinical biomarkers.
D. Rai *et al* (2023)	Compared classical ML and ensemble learning techniques for heart disease prediction using clinical and lifestyle attributes.	Heart disease dataset containing clinical, demographic, and lifestyle features	Gradient Boost achieved 88.69% cross-validation accuracy and 89.71% recall.	Limited by traditional ML models. Future work should investigate explainable AI, larger datasets, and hybrid deep learning approaches for improved prediction performance.

III. METHODOLOGY

Data pre-processing procedures, including noise reduction, encoding, normalization, and SMOTE-ENN data balance, were implemented to guarantee the accuracy of the prediction. Data collection from the MIMIC-III clinical dataset was the subsequent approach. A training set and a test set were created from the processed dataset using stratified sampling. With this method, a model of an artificial neural network (ANN) may learn intricate clinical patterns capable of risk prediction. It is possible to determine the ACC, PRE, REC, and F1 by analyzing the confusion matrix, which allows one to assess the performance of the suggested ANN model. The comprehensive model for clinical risk prediction is displayed in Figure 1.A detailed description of the suggested structure is given in the next section:

A. Data Collection and Exploratory Analysis

The MIMIC-III80 database contains de-identified critical care information and complete clinical records for more than 40,000 ICU patients. Vital signs, medical history, lab results, medical records, and discharge summaries are all part of it. Age, gender, pulse, respiration rate, blood pressure (both systolic and diastolic), oxygen saturation, glucose level, and markers from other lab tests are all structured elements that are extracted to make sure this study is useful for risk assessment and monitoring. To gain a better understanding of the dataset's properties, exploratory data analysis made use of bar graphs, histograms, and correlation heatmaps. Class distributions, feature distributions, and relationships among clinical characteristics were all analyzed with the use of these graphics. The outcomes may be found below:

DOI: https://doi.org/xxxxxxxxxxxxx

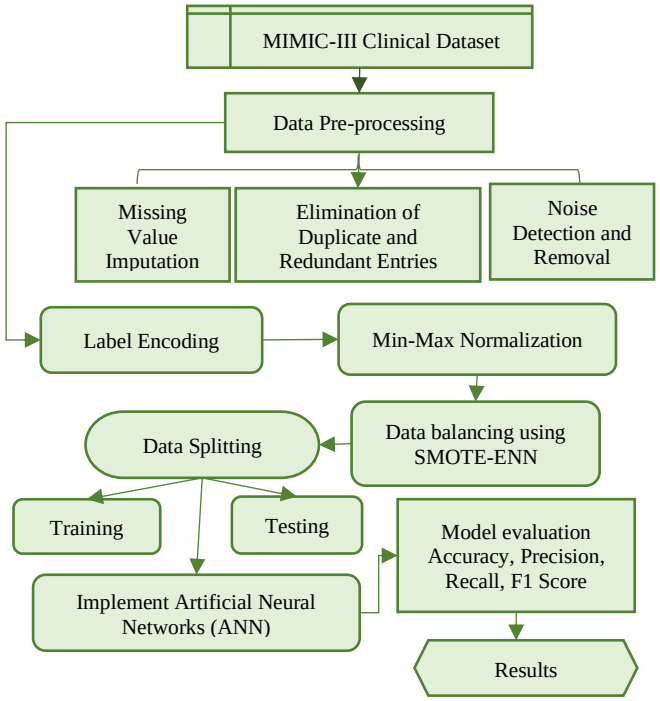

Fig. 1. Proposed Flowchart for Clinical Risk Prediction

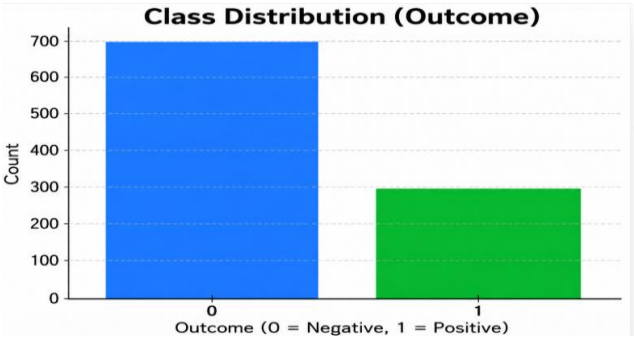

Fig. 2. Bar Graph of Class Distribution

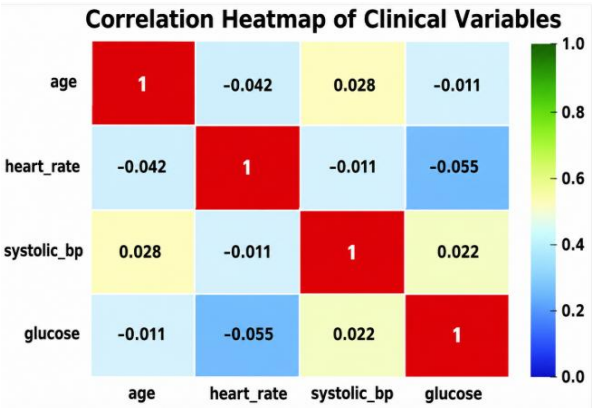

Fig. 3. Correlation Heatmap of Clinical Variables

Figure 2 is a bar graph showing the distribution of the dataset's two result classes; Class 0 (Negative) has a significantly larger sample size than Class 1 (Positive). This means there is a moderate class imbalance, requiring consideration when training the model to ensure dependable and fair predictive performance. The correlation heatmap (Figure 3) shows the correlation between the most important clinical variables, of which most of the variables show a weak linear correlation, with correlation levels close to zero. This means that there are no duplicate variables that give information, and this is important for developing an effective predictive model.

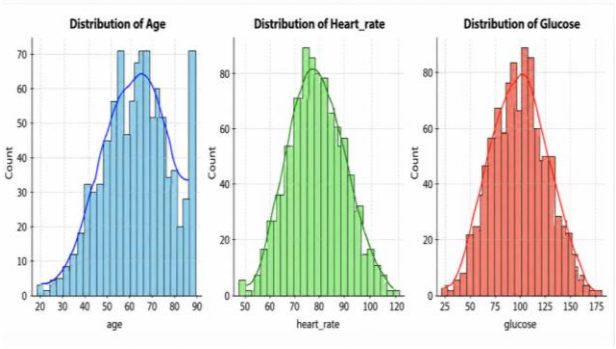

Fig. 4. Distribution of Key Clinical Features: Age, Heart Rate, and Glucose

The histograms in Figure 4 show the distribution of the data for each of these three features, which are all approximately normally distributed around the central range. The distributions suggest that the clinical variables are well captured and appropriate for robust predictive modeling and statistical analysis.

B. Data Pre-Processing

There is a thorough data preparation step for the MIMIC-III dataset that ensures the data is good and makes clinical risk prediction more reliable. As part of the data compilation process, check for inconsistencies and missing values, use statistical methods to fill them in, and eliminate any duplicate or redundant information. Moreover, noise and inconsistencies are identified and removed to ensure data quality, providing a clean and reliable set of data for proper ML model training and assessment.

C. Label Encoding

Label encoding is a method for transforming categorical data into numeric values that may be utilized in ML. Data consistency is maintained and each distinct category is assigned a unique integer label, which allows for fast model training and better prediction.

D. Normalization Using StandardScaler

Each feature in the dataset has a different scale, thus the dataset is scaled using the StandardScaler approach. A perfect distribution, with a mean of zero and a standard deviation of one, is the desired outcome for all characteristics. Equation (1) states that in order to accomplish standardization, the average value is subtracted from each feature's value and then divided by the feature's standard deviation.

$$z = \frac{x-\mu}{\sigma} \quad (1)$$

the transformed value (z) of the feature, the initial values (x) of every trait in the dataset, the mean (μ), and the dispersion (σ) of every trait

E. Data Balancing Using SMOTE-ENN

Data balancing is a pre-processing technique that improves model performance by reducing classification bias caused by an imbalance in the data between majority and minority classes. The generated minority class dataset was balanced using SMOTE-ENN to generate synthetic samples which reduced the noisy and misclassified instances and achieved better robustness, generalization, and predictability.

DOI: https://doi.org/xxxxxxxxxxxxx

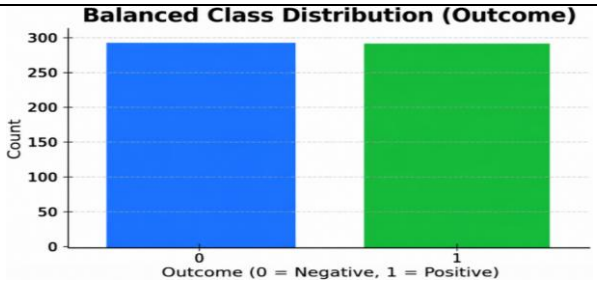

Fig. 5. Distribution Graph After SMOTE-ENN

Figure 5 shows the balanced dataset, which was achieved by applying the SMOTE-ENN approach. Almost equal numbers of samples were found in each class. A more accurate and trustworthy classification model may be achieved with this evenly distributed dataset by reducing the influence of class bias.

F. Data Splitting

A training subset comprising 70% of the dataset and a testing subset comprising 30% of the dataset were created using stratified sampling. In order to train and assess the model without bias, method made sure that the distribution of classes in the training set and testing set matched that of the original dataset.

G. Proposed Artificial Neural Networks (ANN) Model

ANNs are a type of computer architecture that aims to mimic the functions and structure of the human brain. In data, ANNs may uncover intricate patterns and correlations. Input, hidden, and output layers—all comprised of linked networks of neurons—make up a common design for such a network. The network is able to learn from data using training techniques like backpropagation because each neuron analyzes incoming data using weighted connections and activation functions. The ANN is a popular model that mimics the way the human brain uses synapses to store and retrieve information, allowing it to carry out tasks like pattern recognition, classification, and acquiring new knowledge from its surroundings. Equation. (2) gives the output y of neuron k.

$$y_k = \varphi(u_k + b_k) \text{ And } u_k = \sum_{j=0}^{m} w_{kj} x_j \quad (2)$$

This bias influences the activation function's input φ, and w_{kj} is the synaptic weight from input xj to k.

Equation (3) defines the error function that, for the n-th iteration on neuron i, is used to train the weights and biases of a multilayer perceptron. The hidden neurons in a multilayer perceptron are highly connected to one another through synapses, which allows the network to recognize patterns as an essential part of solving complex problems.

$$e_i = d_i(n) - y_i(n) \quad (3)$$

where the neuron's output is denoted by y and d is the intended output. Equation (4) shows the revised error energy function for an output layer of size O.

$$E_{avg} = \frac{1}{N}\sum_{n=1}^{N} E(n) \text{ where } E(n) = \frac{1}{2}\sum_{i=1}^{o} e_i^2(n) \quad (4)$$

The dataset size is represented by N. Optimization seeks to minimize the error energy function as much as possible as it iteratively traverses the network's layers. The recommended ANN model was trained for 30 epochs using the Adam optimizer, with a learning rate of 0.001, batch size of 32, and a binary cross-entropy loss function. There is an input layer, a hidden layer with 64 neurons, an activation function called ReLU, a dropout rate of 0.5, and an output layer that employs sigmoid activation for clinical risk assessment. The architecture of the network further includes these components.

H. Evaluation Metrics

The proposed ANN model is evaluated using a confusion matrix. Discuss the key metrics used to evaluate the model's classification performance, such as ACC, PRE, REC, and F1, that were generated from these data sets below:

Accuracy: A confusion matrix is used to assess the suggested ANN model. This metric, which is given by Equation (5), evaluates the trained model's prediction performance as a whole.

$$Accuracy = \frac{\text{TP+TN}}{\text{TP+Fp+TN+FN}} \quad (5)$$

Precision: The ACC rate, often known as the PRE, is the proportion of correctly predicted positive cases relative to the total number of positive occurrences. It is a way to evaluate the model's prediction accuracy and is represented by Equation (6).

$$Precision = \frac{\text{TP}}{\text{TP+FP}} \quad (6)$$

Recall: REC measures how many positive occurrences were accurately predicted relative to how many positive occurrences actually occurred. This represents the model's ability to find all relevant positive examples, and it's written as Equation (7).

$$Recall = \frac{\text{TP}}{TP+FN} \quad (7)$$

F1 score: This metric represents the harmonic mean of the two variables PRE and REC. You may find it in Equation (8); it ranges from 0 to 1, with higher values indicating better categorization performance.

$$F1 - score = 2 \times \frac{Precision \times Recall}{Precision + Recall} \quad (8)$$

IV. Results and Discussion

This section details the experimental setting and compares the suggested model's performance on a number of performance criteria, drawing attention to its efficacy and efficiency.

A. Experimental Setup

The research was conducted on a workstation equipped with an Intel Core i7 CPU, 32 GB of RAM, and an NVIDIA RTX 3080 GPU. NumPy and Pandas were utilized for data processing; SHAP was employed to include explainability. The software environment also included the DL framework TensorFlow 2.11, Scikit-learn 1.2, and Python 3.9.

B. Experimental Results

The proposed model is validated on key performance metrics like ACC, PRE, REC etc., in the proposed paper by using the MIMIC-III Clinical dataset as summarized in Table II. The ANN model has an ACC of 98.6%, indicating its ability to accurately classify clinical risk instances. It also achieved a PRE of 97.0%, showing a good identification of positive cases, and a REC of 96.5%, meaning that the cases it detected were actual positive cases. Moreover, the model achieved an F1 of 96.9%, which shows a good balance between PRE and REC.

DOI: https://doi.org/xxxxxxxxxxxxx

TABLE II. PERFORMANCE METRICS OF THE PROPOSED CLINICAL RISK PREDICTION MODEL

Matrix	ANN Model
Accuracy	98.6
Precision	97
Recall	96.5
F1-score	96.9

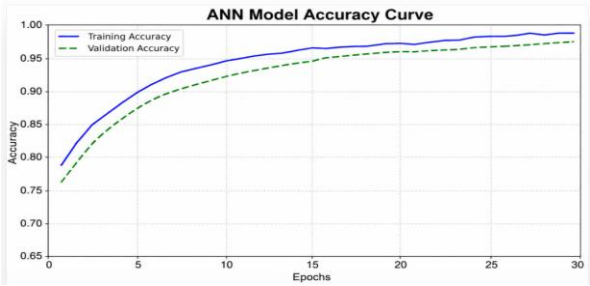

Fig. 6. Training and Validation Accuracy Curve of ANN

Figure 6 shows the ANN model's ACC curve, which shows that ACC becomes higher throughout validation and training. Since the training and validation curves converge smoothly, it means that the model learns well and can generalize well to new data. The curves are very close, which shows low overfitting and proves the stability, robustness, and reliability of the proposed ANN model.

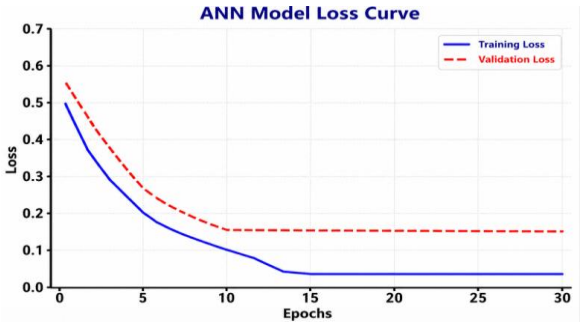

Fig. 7. Training and Validation Loss Curve of ANN model

Training and validation losses go downhill with increasing epoch counts (Figure. 7), suggesting that the model is learning well. The training loss approaches a low loss, and the validation loss doesn't fluctuate significantly. This action proves that the ANN model converges well and produces trustworthy generalization results. Figure 8 shows the suggested model's classification performance using both the actual and anticipated labels. Only five and six samples, respectively, were misclassified; the remaining 204 and 85 samples were correctly classified. The findings provide a high accuracy rate for predictions and a good dependability for classifications.

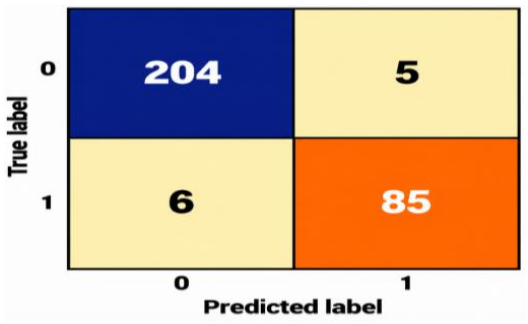

Fig. 8. Confusion Matrix of ANN model

C. Comparative Analysis

The performance of the proposed model is contrasted with that of several ML and DL models in Table III. The purpose of this is to determine whether the model is effective. While comparing the proposed ANN model to KNN, LR, and DNN, we use ACC, PRE, REC, and F1. This KNN model appears to produce subpar classification results with an ACC of 77%, PRE of 50%, REC of 12%, and F1 of 22%. The DNN model outperformed the LR model in terms of ACC, with a 92.1% success rate, 91% PRE, 90% REC, and 90% F1. The suggested ANN model outperformed the competition with impressive results: 98.6% ACC, 97.0% PRE, 96.5% REC, and 96.9% F1. Results show that proposed ANN model beats existing models in terms of clinical predictive performance and prediction accuracy.

TABLE III. COMPARISON OF MACHINE LEARNING AND DEEP LEARNING MODELS FOR CLINICAL RISK PREDICTION

Model	Accuracy	Precision	Recall	F1-score
KNN[23]	77	50	12	22
LR[24]	89	-	-	-
DNN[25]	92.1	91	90	90
ANN	98.6	97	96.5	96.9

The proposed ANN model showed better performance in clinical risk prediction with ACC of 98.6% which was better than the compared models. The model successfully reduces prediction errors while identifying clinical risk patients, as shown by the high PRE, REC, and F1. In order to enhance the algorithm's dependability and predictive ACC, the ANN design may learn intricate nonlinear patterns from clinical data. The highlighted advantages show that the suggested ANN model has promise for assisting healthcare decision-making and providing reliable clinical risk predictions.

D. Justification and Novelty

Health care providers must have access to credible clinical risk prediction tools in order to aid in the early identification of high-risk patients, which is why the proposed study is warranted. Most existing approaches have limitations when it comes to handling complicated clinical interactions, data variability, and the dependability of predictions. The novelty of this work is establishing an ANN-based clinical risk prediction model that can learn non-linear patterns from clinical data and enhance predictive performance by using full preprocessing and optimization process. The proposed solution is intelligent and data-driven, aiming to improve the clinical decision-making process and facilitate personalized healthcare management.

V. CONCLUSION, LIMITATIONS AND FUTURE STUDY

Clinical risk prediction models are crucial in the healthcare sector because they allow doctors and nurses to identify patients who may be at risk and develop personalized care plans. The proposed Artificial Intelligence-based clinical risk prediction framework with ANN is an effective framework that can analyze complex clinical patterns and enhance the ACC of disease risk assessment of patients. The model aims to incorporate all major data preprocessing and learning mechanisms that can boost the reliability of prediction and classification performance. The testing findings demonstrated that the suggested ANN model outperformed competing models with ACCs of 98.6%, LR (89%) and DNN (92.1%). The results demonstrate that the suggested framework has the potential to accurately detect clinical risk factors and deliver healthcare choices. This study's results highlight the potential

of AI-driven clinical risk prediction systems to improve early risk detection, patient monitoring, and the development of smart healthcare aid apps. The study's limitations include using only one clinical dataset, which could restrict the model's wider clinical relevance in other patient populations and clinical settings. Future steps involve additional validation with multiple center datasets, expansion of clinical data sources and optimization for real-time deployment for clinical decision support.

Funding: This research received no external funding.
Institutional Review Board Statement: Not applicable.
Informed Consent Statement: Not applicable.
Data Availability Statement: The data supporting the findings ofthis study are available from the corresponding author uponreasonable request.
Conflicts of Interest: The author declares no conflict of interest.

How to cite this article
S, Jain, "Artificial Intelligence-Based Clinical Risk Prediction Using Ensemble Machine Learning Algorithm," *Journal of Artificial Intelligence in Governance and Public Policy*, vol. 1, no. 3, Jul.–Sep. 2026.Doi. XXXXXX

REFERENCES

[1] F. Wang *et al.*, "Combined assessment of stress hyperglycemia ratio and glycemic variability to predict all-cause mortality in critically ill patients with atherosclerotic cardiovascular diseases across different glucose metabolic states: an observational cohort study with," *Cardiovasc. Diabetol.*, vol. 24, no. 1, p. 199, May 2025, doi: 10.1186/s12933-025-02762-0.

[2] A. Warrier, "Real-Time Healthcare Event Processing: Stream Analytics for Clinical Decision Support," *Int. J. Emerg. Res. Eng. Technol.*, vol. 1, no. 4, pp. 47–54, 2020, doi: 10.63282/3050-922X.IJERET-V1I4P106.

[3] P. Shah, M. Shukla, N. H. Dholakia, and H. Gupta, "Predicting cardiovascular risk with hybrid ensemble learning and explainable AI," *Sci. Rep.*, vol. 15, no. 1, pp. 1–21, 2025, doi: 10.1038/s41598-025-01650-7.

[4] S. D, S. Dharaneesh P, P. K N, R. S, V. V, and V. EA, "A stacked ensemble machine learning framework for healthcare risk prediction," *Int. J. Sci. Res. Arch.*, vol. 19, no. 1, pp. 827–834, Apr. 2026, doi: 10.30574/ijsra.2026.19.1.0824.

[5] S. Mahmud, K. K. Mohammed, V. R. Jain, and S. A. Shah, "Artificial Intelligence-Driven Predictive Models for Identifying Risk Factors of Chronic Diseases," in *2026 14th International Symposium on Digital Forensics and Security (ISDFS)*, Boston, MA, USA: IEEE, Mar. 2026, pp. 01–06. doi: 10.1109/ISDFS69419.2026.11459003.

[6] J. K. Tan *et al.*, "Ensemble Machine Learning Models for Predicting Patients With High Usage: Model Validation and Economic Impact Analysis," *JMIR Med. Informatics*, vol. 14, p. e77202, Feb. 2026, doi: 10.2196/77202.

[7] S. Srivastava, D. Jain, K. Shah, V. D. P. Sambrow, M. Gautam, and S. Sharma, "Heart attack risk prediction using ensemble machine learning models on clinical data," in *Recent Advances in Computational Methods in Science and Technology*, London: CRC Press, 2025, pp. 290–296. doi: 10.1201/9781003662839-45.

[8] F. S. Alotaibi, "Implementation of Machine Learning Model to Predict Heart Failure Disease," *Int. J. Adv. Comput. Sci. Appl.*, vol. 10, no. 6, 2019, doi: 10.14569/IJACSA.2019.0100637.

[9] M. Mittal and V. A. J. Raja, "AI-Driven Clinical Decision Support Systems: Evaluating Impact On Diagnosis And Treatment Accuracy," *Int. J. Eng. Appl. Artif. Intell.*, vol. 3, no. 1, pp. 10–29, Jun. 2025, doi: 10.34218/IJEAAI_03_01_002.

[10] A. Warrier, "Multi-Modal AI Integration for Comprehensive Patient Risk Assessment: Combining Clinical, Imaging, and Genomic Data," *Int. J. Artif. Intell. Data Sci. Mach. Learn.*, vol. 5, no. 1, pp. 125–132, 2024, doi: 10.63282/3050-9262.IJAIDSML-V5I1P112.

[11] S. Appachu Devanira Poovaiah and P. Kumar Pareek, "AI-Optimized Multimodal Neuroimaging Pipelines for Early Alzheimer's Disease Detection: A System-Level Perspective," in *Alzheimer's Disease - Diagnostics, Pathogenesis, and Treatment [Working Title]*, IntechOpen, 2026. doi: 10.5772/intechopen.1015774.

[12] C. K. N and M. T. R, "Explainable Ensemble Learning for Clinical Risk-Based Cardiovascular Disease Prediction," in *2025 IEEE 17th International Conference on Computational Intelligence and Communication Networks (CICN)*, IEEE, Dec. 2025, pp. 1968–1972. doi: 10.1109/CICN67655.2025.11368125.

[13] S. Mukherjee, "Early Detection and Prediction of Depression Based on Data-Driven Machine Learning Techniques in Mental Healthcare," in *2026 IEEE 5th International Conference on AI in Cybersecurity (ICAIC)*, Houston, TX, USA: IEEE, Feb. 2026, pp. 1–6. doi: 10.1109/ICAIC67076.2026.11395670.

[14] R. V. S. S. B. R, A. M, S. Sharhrah, K. Keerthana, and S. Senthil, "A Personalized Federated Multi-Task Learning Framework for Multi-Modal COVID-19 Diagnosis at the Edge," in *2025 Third International Conference on Networks, Multimedia and Information Technology (NMITCON)*, IEEE, Aug. 2025, pp. 1–6. doi: 10.1109/NMITCON65824.2025.11188318.

[15] R. Snehamrutha, "Data-Driven Optimization of Pharmaceutical Manufacturing Processes using Quality by Design (QbD) Frameworks," *Int. J. Curr. Eng. Technol.*, vol. 14, no. 6, pp. 557–566, Nov/Dec, 2024.

[16] S. Gupta and S. Ahmed, "An Explainable Machine Learning-Based Clinical Decision Support Framework for Early Heart Disease Risk Stratification," in *2026 Innovations in Machine, Engineering, and Digital Conference (IMED)*, IEEE, Mar. 2026, pp. 1–6. doi: 10.1109/IMED68921.2026.11484172.

[17] N. G. Kamdi, S. Dagadkar, and K. Wanjari, "Machine Learning Based Early Detection and Risk Prediction of Chronic Kidney Disease Using Clinical Data," in *2026 5th International Conference on Sentiment Analysis and Deep Learning (ICSADL)*, IEEE, Feb. 2026, pp. 998–1002. doi: 10.1109/ICSADL67539.2026.11452119.

[18] A. Kumar, H. Sharma, V. Kumar, M. Kumar, and A. Baiswar, "Comparative Analysis of Machine Learning Models for Clinical Risk Prediction," in *2025 IEEE 1st International Conference on Recent Trends in Computing and Smart Mobility (RCSM)*, IEEE, Dec. 2025, pp. 1–5. doi: 10.1109/RCSM67767.2025.11506805.

[19] A. Raghav, A. Gupta, A. Shinghal, A. K. Pandey, and S. Jain, "Ensemble-Based Cardiovascular Heart Disease Detection System," in *2025 2nd International Conference on Advanced Computing and Emerging Technologies (ACET)*, IEEE, Nov. 2025, pp. 1–4. doi: 10.1109/ACET67282.2025.11430321.

[20] F. Shahidi, M. E. MacDonald, D. Seitz, R. Barry, and G. Messier, "Enhancing Risk Prediction in Mental Health Using Ensemble Hybrid Models and Administrative Healthcare Data with Irregular Intervals," in *2024 IEEE International Conference on Bioinformatics and Biomedicine (BIBM)*, IEEE, Dec. 2024, pp. 3673–3678. doi: 10.1109/BIBM62325.2024.10822734.

[21] C. Kaur and R. Madaan, "Breast Cancer Prediction from Risk Factors Using Ensemble Technique," in *2024 2nd International Conference on Advances in Computation, Communication and Information Technology (ICAICCIT)*, IEEE, Nov. 2024, pp. 353–358. doi: 10.1109/ICAICCIT64383.2024.10912245.

[22] D. Rai, J. Agrawal, D. Pandey, D. Pandey, and M. Jain, "Clinical Efficacy of Ensemble Learning for Cardiovascular Disease Prediction: A Comparative Analysis With Traditional Machine Learning Classifiers," in *2023 International Conference on Modeling, Simulation & Intelligent Computing (MoSICom)*, IEEE, Dec. 2023, pp. 565–569. doi: 10.1109/MoSICom59118.2023.10458781.

[23] N. Shafiabady *et al.*, "Explainable AI for mortality prediction: a comparative study using the MIMIC-III dataset," *BMJ Heal. Care Informatics*, vol. 33, no. 1, p. e101406, Feb. 2026, doi: 10.1136/bmjhci-2024-101406.

[24] K. Karboub and M. Tabaa, "A Machine Learning-Based Discharge Prediction of Cardiovascular Diseases Patients in Intensive Care Units," *Healthcare*, vol. 10, no. 6, p. 966, May 2022, doi: 10.3390/healthcare10060966.

[25] P. Saxena, M. Kumari, and L. Sharma, "Predictive Analysis in Healthcare Using Machine Learning Models : A Hybrid Deep Ensemble Framework for Disease Prediction and Early Diagnosis," no. 12, pp. 2477–2483, 2025.