Minimax Theory and its Applications

Volume 11 (2026), No. 4s, 194 to 204

A Probabilistic Machine Learning Framework for Early Disease Prediction and Clinical Risk Stratification Using Statistical Inference and Artificial Intelligence

Dr. Jain Khusboo Ramesh[1], Dr G Naga Chandrika[2], Jeevan Krishna Paruchuri[3], Dr. Shaik Munawar[4], U.Udayakumar[5], Elangovan Muniyandy[6], Dr. K. Aruna Bhaskar[7], Kokila M M[8]

[1]Senior Resident, Department of Obstetrics and Gynaecology, Saveetha Medical College and Hospital, SIMATS, Thandalam, Chennai, Tamilnadu, India Email: khusboo2796@gmail.com
[2]Sr.Asst Professor, Department of IT, VNRVJIET, Bachupally, Hyderabad, 500118, Email: nchandrika_g@vnrvjiet.in Orcid: 0000000289915930
[3]Senior Software Engineer, United States of America. Email: paruchuri.g1@gmail.com
[4]Associate Professor, Department of computer science and engineering (AI&ML), CMR Engineering College, kandlakoya, Medchal, Hyderabad, Telangana , India-501401. Email: shaikmunawar215@gmail.com, ORCID URL: https://orcid.org/0000-0002-3229-5726
[5]Assistant Professor, Department of Computer Science, Faculty of Liberal Arts and Business Studies, SRM Institute of Science and Technology Ramapuram, Chennai. Email: phd.udayakumar1993@gmail.com
[6]Department of Biosciences, Saveetha School of Engineering. Saveetha Institute of Medical and Technical Sciences,Chennai, 602105, Tamil Nadu, India . Email: muniyandy.e@gmail.com
[7]Associate Professor, KLEF, KL University, Vaddeswaram, Guntur, Andhra Pradesh, India. Email: arunabhaskar@kluniversity.in Orchid Id: https://orcid.org/my-orcid?orcid=0000-0002-8279-9500
[8]Department of Artificial Intelligence and Data Science, Nandha Engineering College, Erode, Tamilnadu, India. Email: kokilamohan88@gmail.com

Abstract
This paper introduces a probabilistic machine-learning approach to nonlinear minimax optimization, statistical inference, and clinical risk stratification of early adverse clinical outcome prediction. Diabetes 130-US Hospitals dataset containing 101,764 encounters of 71,518 patients was used to assess the framework to predict 30-day hospital readmission. A neural-network model minimizing the smooth-max using the minimax-optimization was compared to logistic regression, random forest, gradient boosting, and Bayesian logistic regression. The assessment of performance was based on the use of the following measures of performance: AUROC, AUPRC, Brier score, worst-group loss, bootstrap inference, and sensitivity analysis of age-distribution. The proposed model achieved an AUROC of 0.6380, AUPRC of 0.1850, Brier score of 0.0927, and worst-group loss of 0.3441. Gradient boosting had a slightly higher discrimination (AUROC = 0.6416), with no statistically significant higher discrimination proven by the proposed model. Clinical risk stratification was used to define low-, moderate-, and high-risk groups with reported readmission rates of 5.76%, 10.40%, and 15.96% respectively. The results show that subgroup-oriented minimax estimation is feasible to predict probabilistic clinical risks with a focus on the necessity of additional external and clinical validation.

Keywords: Probabilistic Machine Learning, Minimax Estimation, Statistical Inference, Nonlinear Optimization, Artificial Intelligence, Clinical Risk Stratification, Hospital Readmission.

1. Introduction

Anticipated adverse clinical outcomes should be early predicted to enhance patient monitoring, clinical decision-making, and allocation of healthcare resources. Electronic health records can be used to make probabilistic predictions using statistical inference and artificial intelligence, and mathematical methods to control uncertainty and adverse predictive results can be found in robust optimization (Bertsimas et al., 2011). Random forests are used to learn nonlinear relationships between clinical variables (Breiman, 2001), but federated learning provides privacy-preserving methods of distributed medical data analysis (Muralidharan, 2025). The results of big-data analytics on the Diabetes 130-US Hospitals data set have proven that

computational approaches are applicable in the prediction of hospital readmission (Kanu & Khanal, 2023). The developments in the field of regenerative medicine underscore the drawbacks of healthcare technology transfer to practice in general (Quinby & Yannas, 2025). Distributionally robust optimization offers a conceptual framework to study predictive performance in terms of heterogeneous population distribution (Duchi and Namkoong, 2021), and embedded multi-sensor monitoring systems show complementary innovations in automated data acquisition (Laine, 2026). Gradient boosting helps predict nonlinearly by improving the model step by step (Friedman, 2001) and probability calibration helps remove differences between the forecasted risks and the observed frequency of events (Guo et al., 2017). Another example of the use of computational technologies in healthcare is medical embedded systems (Madugalla & Perera, 2024). Explainable artificial intelligence has been explored to predict complex medical images (Haruna and Taraba, 2025), such as to analyze the decision made by deep-learning through visualization (Bates, 2024). Statistical evaluation models focus on the aspects of discrimination, calibration, and predictive accuracy (Steyerberg et al., 2010), and past studies on the diabetic hospital records prove the importance of readmission as one of the clinical prediction endpoints (Strack et al., 2014). Multidisciplinary approaches based on systems focused on emphasize the combination of mathematical models with practical problem-solving (Rimada & Mrinh, 2025), and decision-curve analysis offers a framework of a potential assessment of the clinical utility of the predictive models (Vickers & Elkin, 2006). Although this has happened, traditional statistical and machine-learning models tend to reduce the average empirical prediction loss, which can mask the relatively large prediction errors within patient subgroups with heterogeneity. Specifically, maximizing overall discrimination may not be the best way to maximize the upper subgroup-specific binary cross-entropy loss or to ensure reliable probability estimates across age groups. Direct minimax optimization poses extra mathematical problems since the maximum-loss operator may be nonsmooth and nonconvex neural-network objectives are typically nonconvex. Moreover, training-group minimax loss improvements do not necessarily determine better test-set robustness, probability calibration and clinical utility. These constraints inspire a differentiable subgroup-based minimax model with log-sum-exp approximation, with approximation inequalities, entropy-regularized variational analysis, and numerical analysis. The current research thus comes up with a probabilistic machine-learning model that incorporates nonlinear age-group minimax estimation, statistical inference, probability calibration, and clinical risk stratification. Based on 101,766 hospital encounters of patients with diabetes, the framework predicts 30-day hospital readmission instead of new-onset disease. The minimax neural network proposed is contrasted with logistic regression, random forest, gradient boosting, and Bayesian logistic regression based on the means of the AUROC, AUPRC, Brier score, worst-group loss, patient-cluster bootstrap inference and age-distribution sensitivity analysis. The main contribution is a mathematically defined smooth-max minimax framework of learning that provides more weight to higher-loss patient subgroups and allows the assessment of risk probabilistically. The research assesses its quantitative viability and predictive validity without presupposing that it is globally optimal, statistically significantly better, or that it is clinically proven to be effective.

2. Methodology

2.1 Clinical Dataset, Preprocessing, and Statistical Inference

This research used a retrospective computational design based on Diabetes 130-US Hospitals data to forecast 30-day readmission of patients with preexisting diabetes. The data consisted of 101,764 hospital visits of 71,518 patients, out of which 11.16 were the prevalence of positive outcomes. The clinical records were data cleaned, categorically encoded, and prepared features and patient-separated into training, validation and test sets. The output of this was 160 coded features. Classifications were made of patients as younger than 50, 50 to 69 and 70 years and above. The characteristics of the dataset and the experimental setup are summarized in Table 1.

Table 1. Clinical Dataset Characteristics and Experimental Configuration

Parameter	**Reported value**
Dataset	Diabetes 130-US Hospitals
Hospital encounters	101,766
Unique patients	71,518
Training encounters	61,215
Validation encounters	20,398

Testing encounters	20,153
Encoded input features	160
Positive-outcome prevalence	11.16%
Prediction endpoint	30-day hospital readmission
Data partitioning	Patient-separated
Age groups	Below 50, 50–69, ≥70 years
Bootstrap replications	300

Assume that the clinical dataset is modeled by

$$\mathcal{D} = \{(x_i, y_i, g_i)\}_{i=1}^{N}, \quad x_i \in \mathbb{R}^{160}, \quad y_i \in \{0,1\}, \quad g_i \in \{1,2,3\}. \quad (1)$$

The probability of readmission is estimated as

$$p_\theta(x_i) = P(Y_i = 1 \mid x_i; \theta). \quad (2)$$

In the case of logistic regression, this probability is given as follows:

$$p_\theta(x_i) = \frac{1}{1 + \exp[-(\beta_0 + \beta^\top x_i)]}. \quad (3)$$

The binary cross-entropy objective is used as conventional empirical risk minimization.

$$\mathscr{L}_{\mathrm{ERM}}(\theta) = -\frac{1}{N}\sum_{i=1}^{N}[y_i \log p_\theta(x_i) + (1 - y_i)\log(1 - p_\theta(x_i)))]. \quad (4)$$

Though Equation (4) minimizes the average prediction loss, it does not explicitly minimize the maximum subgroup-specific error, which drives the proposed formulation of minimax.

2.2 Nonlinear Minimax Estimation and Mathematical Properties

The suggested framework reduces the highest binary cross-entropy loss in three age groups of patients. The group-specific loss of subgroup$\mathcal{D}_g$containing N_g observations is:

$$\mathscr{L}_g(\theta) = -\frac{1}{N_g}\sum_{i \in \mathcal{D}_g}[y_i \log p_\theta(x_i) + (1 - y_i)\log(1 - p_\theta(x_i)))]. \quad (5)$$

The respective minimax goal is

$$\min_{\theta \in \Theta} \max_{g \in \{1,2,3\}} \mathscr{L}_g(\theta). \quad (6)$$

To get a differentiable objective, maximum operator is approximated by the log-sum-exp function:

$$\Psi_\eta(\theta) = \frac{1}{\eta}\log\left(\sum_{g=1}^{G}\exp[\eta \mathscr{L}_g(\theta)]\right), \quad \eta > 0. \quad (7)$$

Proposition 1 (Smooth-max approximation bound). The approximation satisfies

$$\max_g \mathscr{L}_g(\theta) \le \Psi_\eta(\theta) \le \max_g \mathscr{L}_g(\theta) + \frac{\log G}{\eta}. \quad (8)$$

This is based on the bounding of the exponential sum of $e^{\eta M}$and $Ge^{\eta M}$, where $M = \max_g \mathscr{L}_g(\theta)$. In the case of G=3 and $\eta = 10$the theoretical approximation error is bounded by

$$\frac{\log 3}{10} \approx 0.1099. \quad (9)$$

Proposition 2 (Entropy-regularized representation). The variational form of the smooth-max objective is admissible.

$$\Psi_\eta(\theta) = \max_{q \in \Delta_G} \left[\sum_{g=1}^{G} q_g \, \mathscr{L}_g(\theta) + \frac{H(q)}{\eta} \right], \quad (10)$$

where

$$H(q) = -\sum_{g=1}^{G} q_g \log q_g . \quad (11)$$

The maximizing subgroup weights are

$$q_g^*(\theta) = \frac{\exp\left[\eta \mathscr{L}_g(\theta)\right]}{\sum_{j=1}^{G} \exp\left[\eta \mathscr{L}_j(\theta)\right]}. \quad (12)$$

This identity follows from the nonnegativity of the Kullback–Leibler divergence between q and q^*.Accordingly, the objective gradient is.

$$\nabla_\theta \Psi_\eta(\theta) = \sum_{g=1}^{G} q_g^* \, (\theta) \nabla_\theta \mathscr{L}_g(\theta). \quad (13)$$

Therefore, age groups with higher losses are better optimized. The objective with optional regularization is

$$\min_{\theta \in \Theta} \left[\Psi_\eta(\theta) + \lambda R(\theta) \right]. \quad (14)$$

Proposition 3 (Existence of a minimizer). When the set of θ, denoted by or is non empty and compact, and the group losses and regularization function are continuous, the Weierstrass theorem implies that Equation (14) has a global minimum. In the case of differentiable unconstrained optimization, a condition of necessary stationarity is.

$$\nabla_\theta \left[\Psi_\eta(\theta^*) + \lambda R(\theta^*) \right] = 0. \quad (15)$$

Nonconvexity in a neural-network, however, does not imply that stationarity is a sufficient condition of global optimality.

2.3 Probabilistic Machine Learning and Clinical Risk Stratification Framework

The computational model is a combination of clinical preprocessing, statistical inference, probabilistic prediction, nonlinear minimax optimization, probability calibration, and clinical risk classification. Five models were tested: logistic regression, random forest, gradient boosting, Bayesian logistic regression and the proposed minimax neural network. The proposed neural network was the only neural network on which the minimax objective was applied, with 128 hidden neurons, dropout=0.20, and $\eta = 10$. Validation data were used to test four hyperparameter settings. Output of the model was transformed into readmission probabilities and the calibration of the sigmoid and isotonic was tested and sigmoid calibration was chosen. The probability of the calibration is denoted as.

Model outputs were converted into readmission probabilities, with sigmoid and isotonic calibration evaluated and sigmoid calibration selected. The calibrated probability is represented by

$$\tilde{p}_i = \frac{1}{1 + \exp(A s_i + B)}, \quad (16)$$

In which s_i is the model prediction score and A and B are calibration parameters that are estimated on validation data. Thresholds based on validation were used to classify the calibrated probabilities:

$$S_i = \begin{cases} \text{Low risk}, & \tilde{p}_i < 0.08165, \\ \text{Moderate risk}, & 0.08165 \le \tilde{p}_i < 0.11560, \\ \text{High risk}, & \tilde{p}_i \ge 0.11560. \end{cases} \quad (17)$$

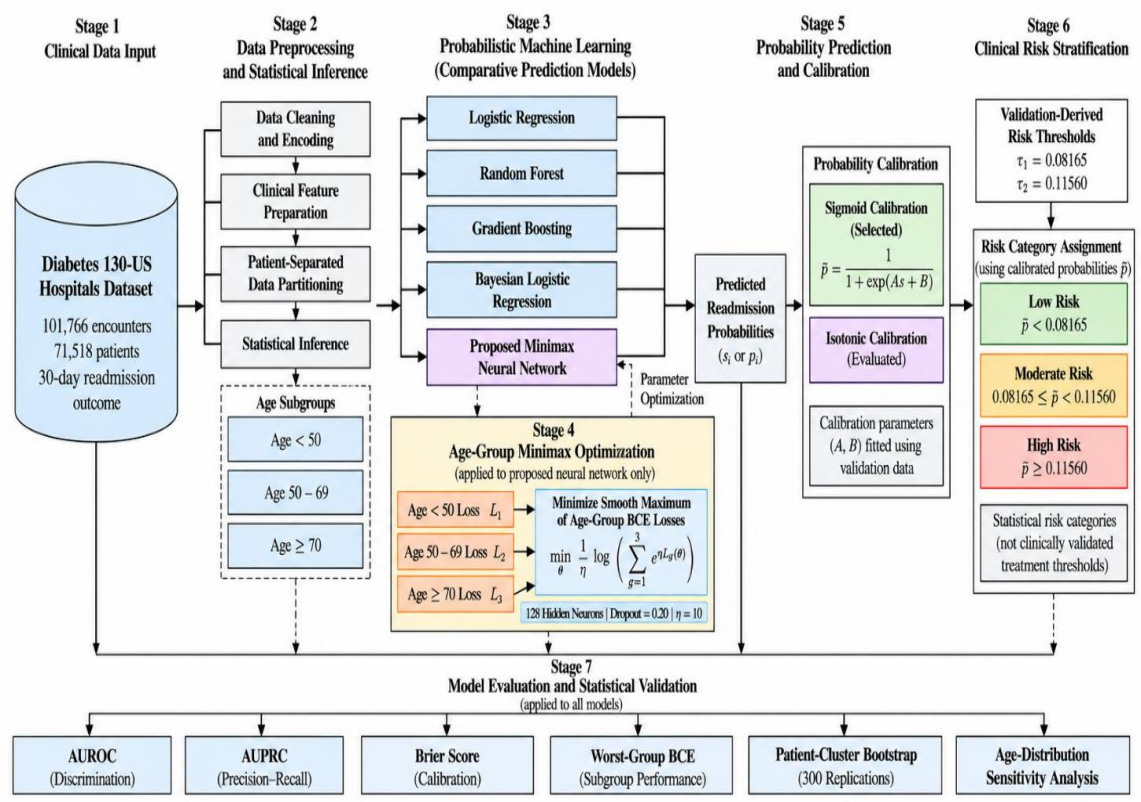


Figure 1. Proposed Probabilistic Machine Learning and Minimax-Based Clinical Risk Stratification Framework

Figure 1 shows a suggested seven-step computational model, which incorporates clinical data preprocessing, comparative probabilistic prediction, age-group minimax optimization, probability calibration using validation, clinical risk stratification, and statistical assessment. The minimax goal is implemented on the proposed neural-network model only, whereas the five models are evaluated on comparative performance.

2.4 Numerical Optimization and Statistical Evaluation

To optimize the proposed minimax model, the differentiable smooth-max objective was optimized numerically. An overall gradient-based update of parameters is

$$\theta^{(k+1)} = \theta^{(k)} - \alpha_k \nabla_\theta \Psi_\eta\left(\theta^{(k)}\right), \quad (18)$$

where α_k is the rate of learning. The predictive performance was evaluated by the use of the AUC, AUPRC, Brier score and worst-group binary cross-entropy. The Brier score is

$$BS = \frac{1}{N}\sum_{i=1}^{N}(\tilde{p}_i - y_i)^2, \quad (19)$$

and worst-group prediction loss is

$$\mathscr{L}_{\text{worst}} = \max_{g=1,\ldots,G} \mathscr{L}_g. \quad (20)$$

The statistical uncertainty was determined by 300 bootstrap replications of patients clusters.Age-group composition sensitivity was measured using scenario-weighted losses in prediction:

$$\mathscr{L}_{\text{shift}}(\pi) = \sum_{g=1}^{G} \pi_g \mathscr{L}_g, \quad \pi_g \ge 0, \quad \sum_g \pi_g = 1. \quad (21)$$

The consideration included original, younger, older and strong older-population distributions in the analysis. This is a process that quantifies age-mixture change sensitivity in the current test data instead of extrapolation to the population. Table 2 indicates the optimization and statistical assessment setup.

Table 2. Proposed Minimax Model and Statistical Evaluation Configuration

Parameter	Reported configuration
Proposed architecture	Nonlinear minimax neural network
Optimization objective	Smooth maximum of age-group BCE losses
Number of age groups	3
Selected hidden-layer width	128
Dropout rate	0.20
Smooth-max parameter (η)	10
Hyperparameter configurations	4
Best validation worst-group loss	0.35016
Selected probability calibration	Sigmoid
Risk thresholds	0.08165 and 0.11560
Statistical validation	300 patient-cluster bootstrap replications
Evaluation metrics	AUROC, AUPRC, Brier score, worst-group BCE
Distribution-shift scenarios	Original, younger, older, strong older-population shift

Table 2 indicates the experimental set-up reported. The provided summary does not define all the optimizer settings, learning-rate schedule, training epochs, or random seeds; all of the information would be required to achieve full computational reproducibility.

3. Results and Discussion

3.1 Comparative Prediction Performance and Statistical Validation

The proposed minimax model was contrasted with logistic regression, random forest, gradient boosting and Bayesian logistic regression on 20,153 independent test encounters. The comparative predictive performance is given in Table 3. Gradient boosting had the largest AUROC (0.6416), AUPRC (0.1855) whilst the proposed minimax model had an AUROC of 0.6380, AUPRC of 0.1850, Brier score of 0.0927 and worst-group loss of 0.3441. Figure 2 compares the reported values of the AUROC and shows that the proposed model was similar to gradient boosting without surpassing its total discrimination.

Table 3. Comparative Performance of Statistical, Machine Learning, and Proposed Minimax Models

Prediction model	AUROC	AUPRC	Brier score	Worst-group loss
Logistic regression	0.6367	0.1820	0.0928	0.3446
Random forest	0.6343	0.1837	0.0927	0.3442
Gradient boosting	0.6416	0.1855	0.0927	0.3434
Bayesian logistic regression	0.6367	0.1818	0.0928	0.3447
Proposed minimax model	0.6380	0.1850	0.0927	0.3441

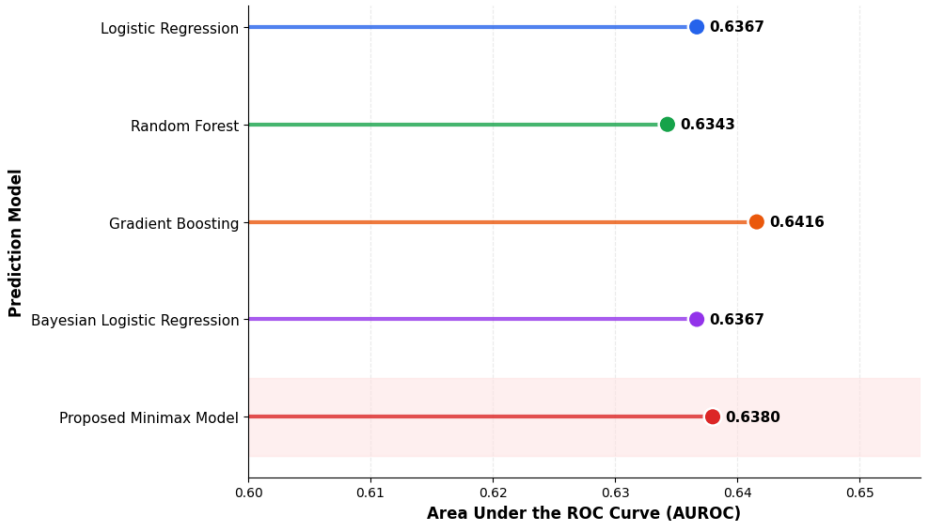


Figure 2. Comparative AUROC Performance of Statistical, Machine Learning, and Proposed Minimax Models.

The statistical reliability was determined at 300 patient-cluster bootstrap replications. The proposed model had a 95% confidence interval of 0.62400.6510 in terms of AUROC, as shown in Table 4. Its paired AUROC difference relative to logistic regression was 0.00127 (95% CI: −0.00424 to 0.00723), whereas the difference relative to gradient boosting was −0.00370 (95% CI: −0.00955 to 0.00193). Both confidence intervals contained zero, which means that there was no statistically significant higher superiority established.

Table 4. Bootstrap Confidence Intervals for Comparative Prediction Performance

Prediction model	**AUROC 95% CI**	**AUPRC 95% CI**
Logistic regression	0.6225–0.6498	0.1658–0.1992
Random forest	0.6199–0.6498	0.1669–0.2017
Gradient boosting	0.6267–0.6565	0.1659–0.2058
Bayesian logistic regression	0.6226–0.6498	0.1655–0.1989
Proposed minimax model	0.6240–0.6510	0.1658–0.2032

3.2 Minimax Optimization and Age-Group Robustness

The proposed neural network used smooth-max minimax optimization, 128 hidden neurons, dropout of 0.20 and a smoothing parameter of 10. The best reported validation worst-group loss was 0.35016. The theoretical smooth-max approximation bound is represented in Figure 3 and declines with an increase in the smoothing parameter. The upper bound of 0.1099 is approximate when age is grouped into three groups and $\eta = 10$. This mathematical finding describes accuracy of approximation, as opposed to convergence of empirical training.

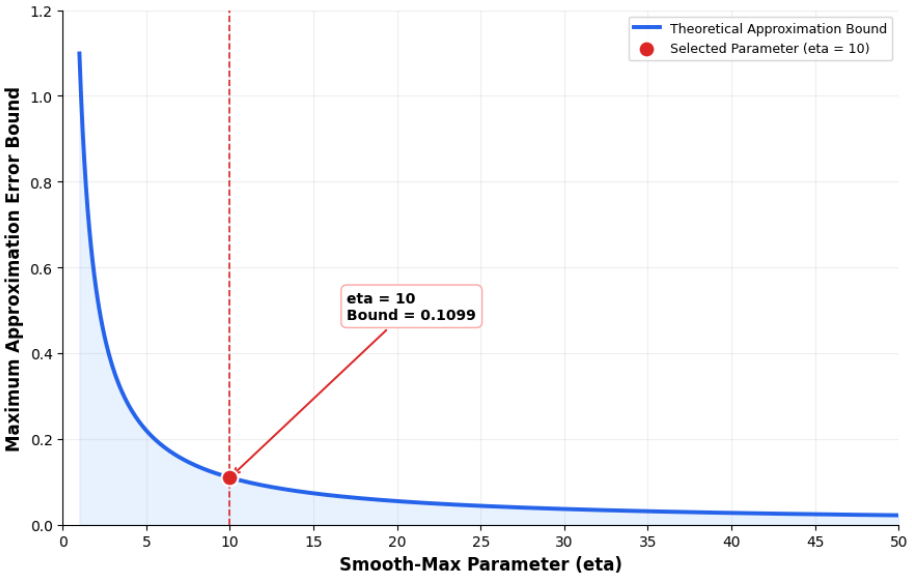


Figure 3. Smooth-Max Approximation Bound in Minimax Optimization.

The age-group prediction losses are summarized, Table 5 and visualized, Figure 4. The greatest prediction loss was observed in all models among patients with the age of 70 years and more. The minimax model proposed had losses of 0.3105, 0.3155 and 0.3441 in the three age groups. Gradient boosting had the lowest maximum loss of 0.3434 but the maximum loss was lower than that of logistic regression by a margin of 0.0005.

Table 5. Comparative Binary Cross-Entropy Loss Across Patient Age Groups

Prediction model	**Below 50**	**50–69**	**≥70**	**Maximum loss**
Logistic regression	0.3084	0.3164	0.3446	0.3446
Random forest	0.3110	0.3162	0.3442	0.3442
Gradient boosting	0.3097	0.3154	0.3434	0.3434
Bayesian logistic regression	0.3083	0.3164	0.3447	0.3447
Proposed minimax model	0.3105	0.3155	0.3441	0.3441

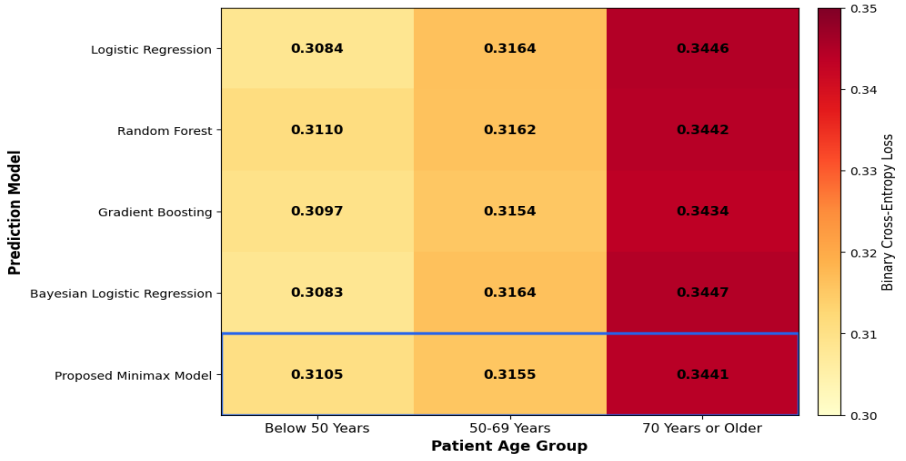


Figure 4. Age-Group-Specific Prediction Loss Comparison.

The sensitivity of age-distribution was studied in four population-compositions. Table 6 and Figure 5 indicate that the weighted prediction losses also rose with the proportion of older patients. In the strong older-population

shift, the proposed model experienced a loss of 0.3382, which is 0.3376 in gradient boosting. These results are sensitive to composition by age group, but are not robust to arbitrary changes in clinical distribution.

Table 6. Predictive Loss Under Age-Distribution Shift Scenarios

Prediction model	Original	Younger	Older	Strong older shift
Logistic regression	0.3277	0.3204	0.3343	0.3387
Random forest	0.3279	0.3211	0.3342	0.3384
Gradient boosting	0.3270	0.3202	0.3333	0.3376
Bayesian logistic regression	0.3277	0.3204	0.3344	0.3388
Proposed minimax model	0.3274	0.3206	0.3338	0.3382

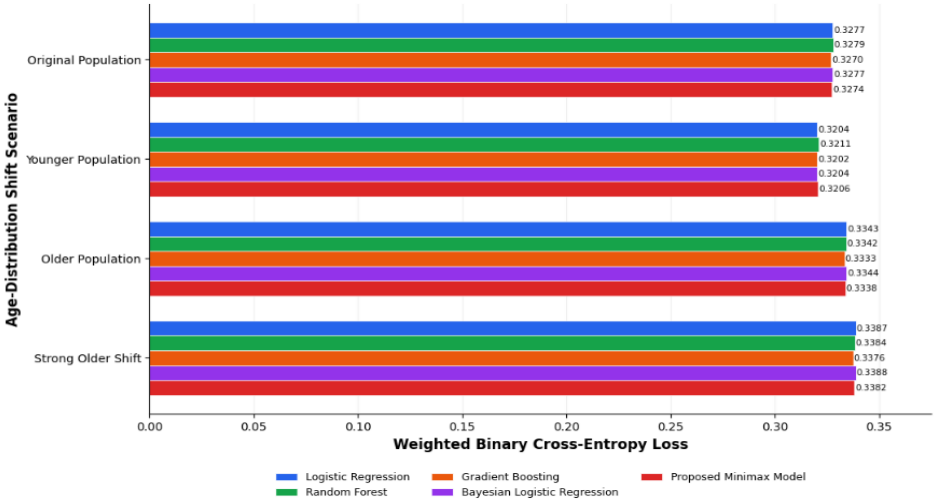


Figure 5. Prediction Loss Under Age-Distribution Shift.

3.3 Probability Calibration and Clinical Risk Stratification

Probability calibration was assessed based on sigmoid and isotonic approaches, with the choice of sigmoid calibration based on validation data. The minimax model proposed had a Brier score of 0.0927 that is similar to random forest and gradient boosting. The Brier-score comparison in Figure 6 shows a close match of overall probabilistic prediction accuracy between the models considered. Calibration-bin observations are not available in the provided summary, which means that the agreement of the calibration curves cannot be calculated separately.

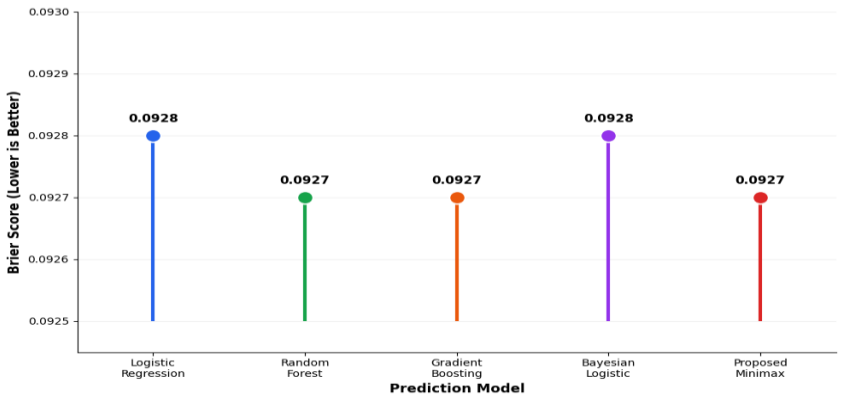


Figure 6. Comparison of Probabilistic Prediction Accuracy.

The probability thresholds 0.08165 and 0.11560 were derived with validation to do clinical risk stratification. As shown in Table 7 and Figure 7, the 30-day readmission rates were observed to rise between 5.76% in the low-risk group to 10.40% in the moderate-risk group and 15.96% in the high-risk group. The readmission rate was about 2.77 times higher in the high-risk group as compared to the low-risk group.

Table 7. Clinical Readmission Risk Stratification Using the Proposed Minimax Model

Risk category	Predicted probability range	Encounters	Observed readmission rate
Low risk	$p < 0.08165$	6,734	5.76%
Moderate risk	$0.08165 \leq p < 0.11560$	6,801	10.40%
High risk	$p \geq 0.11560$	6,618	15.96%
Total	—	20,153	—

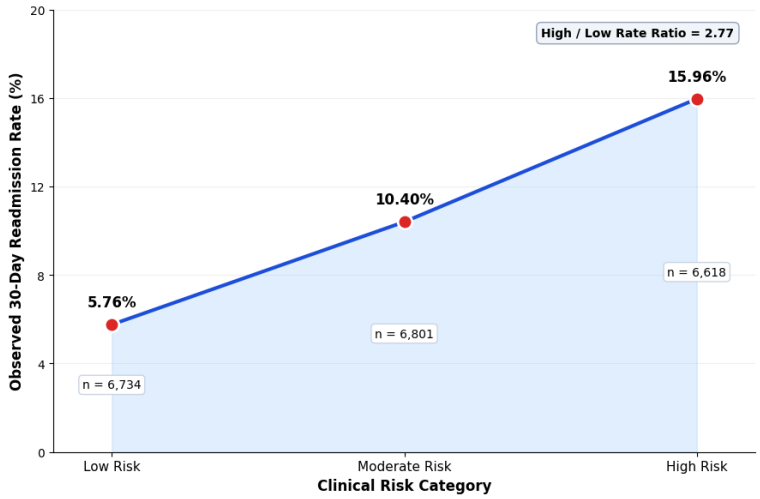


Figure 7. Observed Readmission Risk Across Clinical Categories.

The rising trend in readmission rates indicates that the model can differentiate cases with varying amounts of statistical risk. But the thresholds have been determined based on validation data and have not been determined as clinical intervention criteria. On the whole, the suggested framework has shown viable subgroup-based minimax learning and significant statistical risk segregation, yet gradient boosting had a slight numerical edge. Before it can be used practically, further external validation, prospective analysis, and clinical decision-utility analysis are necessary.

4. Conclusion

This paper formulated a probabilistic machine-learning model, which combines nonlinear minimax optimization, statistical inferences, and probability calibration, and clinical risk stratification in predicting 30-day hospital readmission in patients with preexisting diabetes. The proposed minimax neural network with 101,766 hospital encounters had an AUROC of 0.6380, AUPRC of 0.1850, a Brier score of 0.0927 and the worst-group loss of 0.3441. The proposed approach offered a mathematically interpretable process of focusing on age groups that had more prediction losses (although the gradient boosting showed a little better overall performance). There was no statistically significant better result in patient-cluster bootstrap analysis as compared to the comparative models. The quantified framework distinguished low-, moderate-, and high-risk encounters with the readmission rates of 5.76%, 10.40%, and 15.96%, respectively. These results indicate the computational viability of subgroup-oriented minimax learning of probabilistic clinical risk assessment. Nevertheless, a small level of discrimination, constrained optimization warranties, and external clinical validation lack eliminate expedient practical application. Additional studies are needed in the future to explore a better nonlinear optimization, wider patient subgroups, external hospital data, and prospective clinical validation to determine generalizability and clinical utility.

References

1. Bertsimas, D., Brown, D. B., & Caramanis, C. (2011). Theory and applications of robust optimization. *SIAM review*, *53*(3), 464-501.

2. Breiman, L. (2001). Random forests. *Machine learning*, *45*(1), 5-32.

3. Muralidharan. J, "Federated Learning Models for Privacy-Preserving Medical Image Analysis", *National Journal of Ubiquitous Computing and Intelligent Environments*, vol. 2, no. 1, pp. 1–6, Dec. 2025.

4. Kanu, D. K. S., & Khanal, M. (2023). Implementation of Big Data Analytics on Diabetes 130-US Hospitals for year 1999–2008 for predicting patient readmission. *ResearchGate Preprint*.

5. Quinby, B., & Yannas, B. (2025). Future of tissue engineering in regenerative medicine: Challenges and opportunities. *Innovative Reviews in Engineering and Science*, *3*(2), 73-80.

6. Duchi, J. C., & Namkoong, H. (2021). Learning models with uniform performance via distributionally robust optimization. *The Annals of Statistics*, *49*(3), 1378-1406.

7. Laine, A. (2026). Design and implementation of a real-time edge-based IoT monitoring system using embedded multi-sensor networks. *Archives of Embedded and IoT Systems Engineering*, *2*(1), 29-34.

8. Friedman, J. H. (2001). Greedy function approximation: a gradient boosting machine. *Annals of statistics*, 1189-1232.

9. Guo, C., Pleiss, G., Sun, Y., & Weinberger, K. Q. (2017, July). On calibration of modern neural networks. In *International conference on machine learning* (pp. 1321-1330). PMLR.

10. Madugalla, A. K., & Perera, M. (2024). *Innovative uses of medical embedded systems in healthcare. Progress in Electronics and Communication Engineering, 2 (1), 48–59*.

11. Haruna, I. U., & Taraba, I. H. (2025). Visualizing deep learning decisions: Grad-CAM-based explainable AI for medical image analysis. *IIRJET*, *11*(2).

12. Bates, M. P. (2024). Visualizing deep learning decisions: grad-CAM-based explainable AI for medical image analysis. *Journal of Scalable Data Engineering and Intelligent Computing*, 28-35.

13. Steyerberg, E. W., Vickers, A. J., Cook, N. R., Gerds, T., Gonen, M., Obuchowski, N., ... & Kattan, M. W. (2010). Assessing the performance of prediction models: a framework for some traditional and novel measures. *Epidemiology (Cambridge, Mass.)*, *21*(1), 128.

14. Strack, B., DeShazo, J. P., Gennings, C., Olmo, J. L., Ventura, S., Cios, K. J., & Clore, J. N. (2014). Impact of HbA1c measurement on hospital readmission rates: analysis of 70,000 clinical database patient records. *BioMed research international*, *2014*(1), 781670.

15. Rimada, Y., & KL Mrinh, C. (2025). Bridging the Theory–Practice Gap: A Systems-Oriented Multidisciplinary Framework for Complex Real-World Problem Solving. *Bridge: Journal of Multidisciplinary Explorations*, *1*(3), 41-48.

16. Vickers, A. J., & Elkin, E. B. (2006). Decision curve analysis: a novel method for evaluating prediction models. *Medical Decision Making*, *26*(6), 565-574.