Minimax Theory and its Applications

Volume 11 (2026), No. 4s, 214 to 224

# An AI-Driven Statistical Learning Model for Medical Diagnosis: Integrating Bayesian Probability, Multivariate Analysis, and Predictive Machine Learning

**Mr V Nivedan[1], S.Sindhu[2], Dr.T. Rameshkumar[3], Masarath Saba[4], Mithun M S[5], G.Sivarama Subramanium[6], Dr. V. Yamuna[7], Dr. K. Aruna Bhaskar[8]**

[1]Assistant Professor, Department of Electronics and Communication Systems, Sri Ramakrishna College of Arts & Science, Coimbatore, tamilnadu, India. Email: nivedan@srcas.ac.in orchid id: 0009-0007-3454-6852
[2]Department of Computer Applications(BCA &BCA Gen AI), SRM Institute of Science and Technology, Ramapuram, Chennai, India. Email: Sindhuu89@gmail.com, https://orcid.org/0000-0002-2133-6377
[3]Associate Professor in Mathematics, Department of Mathematics, V. S. B. Engineering College (Autonomous), Karur, Tamilnadu, India -639 111. Email: drtrameshkumar@gmail.com
[4]Assistant Professor, Department of CSE (AI&ML), CVR College of engineering, Ibrahimpatnam (Mandal, Pocharam Rd, Vastunagar, Mangalpalli (Village, Telangana 501510. Email: gswarnalatha79@gmail.com
[5]Department of ECE, Presidency University, Yelahanka, Bengaluru, Karnataka - 560119, India. Email: mithun.ms@presidencyuniversity.in
[6]Department of Electronics and Communication Engineering, Sasurie College of Engineering, Vijayamangalam-638056, Tamilnadu, India. Email: gsivaram47@gmail.com
[7]Assistant Professor, Department of Mathematics, Kongu Engineering College, Perundurai,Tamilnadu, India. Email: vyamuna.1986@gmail.com Orcid ID: 0000-0003-2379-6825
[8]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

**Abstract**
Medical diagnosis mediated by artificial intelligence demands dependable predictive models with the capacity for handling intricate diagnostic indicators and uncertainty in the form of probabilities. This study presents an integrated Bayesian–Multivariate–Machine Learning (BMML) framework which integrates the Bayesian inference, principal component analysis and classification of predictive models. It uses 1,000 observations, 10 diagnostic indicators, 2 diagnostic classes and 100 repeated evaluation runs in the computational design. Feature redundancy is dealt with by multivariate analysis, and logistic regression and random forest as well as gradient boosting aid in the diagnostic classification. Weighted probability fusion: A fusion of bayes and machine learning prediction. Here, the assumption of a likelihood ratio of 4.0 results in an illustrative Bayesian calculation that increases the posterior diagnostic probability from 30% to 63.16%. The evaluation framework also takes into account the metric of accuracy, sensitivity, specificity, F1-score, ROC-AUC and calibration reliability. The proposed approach will lay a structured framework for reviewing diagnostic discrimination and uncertainty, and it is important to remember that its effectiveness needs to be proven clinically in the future.

*Keywords:* Medical Diagnosis, Bayesian Probability, Multivariate Analysis, Predictive Machine Learning, Principal Component Analysis, Statistical Learning.

## 1. Introduction

Medical diagnosis, particularly when it is used to make accurate predictions, systematic analysis, and clinical decision making, is increasingly making use of artificial intelligence (AI) and statistical learning. Traditional diagnostic process involves the use of single statistical models which are not necessarily able to describe nonlinear relationships, linked clinical markers, or diagnostic insecurities. The mathematical basis for using Bayesian probability is for estimating the posterior diagnostic probability, while multivariate statistical analysis is utilized to find the relationships among multiple indicators and to reduce the redundancy of features. In addition to statistical and nonlinear learning classification methods like logistic regression, random forest and gradient boosting are used to enhance diagnostic classification using predictive machine learning techniques.

All these advances have been made, but the state-of-the-art diagnostic prediction methods often fail to effectively integrate probabilistic reasoning, dimensionality reduction and machine learning classification. Many methodological challenges still are shared, such as multicollinearity, lack of probability calibration and differing sensitivity and specificity. Furthermore, limited robustness assessment under different diagnostic situations confines the trustworthiness of the computational predictions.

These drawbacks are overcome in this research by proposing an integrated Bayesian–Multivariate–Machine Learning (BMML) framework which integrates principal component analysis (PCA), Bayesian posterior probability estimation and weighted predictive probability fusion. The proposed computational design involves 100 repeated evaluation runs, 10 normalized diagnostic indicators, two diagnostic classes and 1,000 observations. Logistic regression, random forest, and gradient boosting are used to explore the predictive performance of the different approaches. Different metrics for assessing accuracy, precision, sensitivity, specificity, F1-score, ROC-AUC and probability calibration are used for diagnostic discrimination and reliability. The main contribution is the integration within a unified framework of multivariate feature transformation, probabilistic inference and predictive classification, which is statistical learning and is reproducible. The framework provides a structured framework for assessing diagnostic uncertainty, classification performance and computational robustness, and acknowledges the need for clinical validation in the future.

## 2. Related Work

Statistical learning approaches have been extensively investigated for medical diagnostic prediction, particularly through conventional classification techniques and logistic regression (Steyerberg et al., 2010). These methods provide interpretable relationships between clinical indicators and diagnostic outcomes (Rajkomar et al., 2019). However, traditional statistical models often exhibit limitations when handling nonlinear interactions, correlated variables, and complex diagnostic patterns (Beam & Kohane, 2018). Bayesian probability offers an alternative approach by incorporating prior knowledge, likelihood functions, and posterior probability estimation to quantify diagnostic uncertainty (Fouladvand et al., 2023). Although Bayesian inference supports probabilistic decision-making, its effectiveness depends on appropriate prior assumptions and accurate likelihood specification (Wynants et al., 2020).

Multivariate statistical analysis provides additional capabilities for examining relationships among diagnostic indicators (Maddy et al., 2019). Correlation analysis and covariance-based transformations identify statistical dependencies, while principal component analysis (PCA) reduces dimensionality and addresses feature redundancy (Jolliffe & Cadima, 2016). Nevertheless, excessive dimensionality reduction may eliminate diagnostically relevant information (Miotto et al., 2018). Predictive machine learning algorithms, including logistic regression, random forest, and gradient boosting, have been widely explored for binary classification (Kuhn et al., 2013). Random forest captures nonlinear relationships through ensemble decision trees, whereas gradient boosting improves predictive learning through sequential error correction (Topol, 2019). Their effectiveness is commonly assessed using accuracy, sensitivity, specificity, F1-score, ROC-AUC, and probability calibration (Geetha, 2026).

Despite these methodological advances, existing approaches do not consistently integrate Bayesian probability estimation, multivariate feature transformation, and predictive machine learning within a unified diagnostic framework (Freire, 2025). Furthermore, probability calibration, model robustness, and comparative evaluation under controlled computational conditions require greater attention (Kavitha, 2026). Addressing these gaps, the present study proposes a Bayesian–Multivariate–Machine Learning (BMML) framework integrating PCA, Bayesian posterior estimation, and weighted predictive probability fusion (Ambuli, 2026). The framework supports systematic evaluation of diagnostic discrimination, uncertainty, and computational stability across repeated numerical assessments (Surendar, 2025).

## 3. Proposed Bayesian–Multivariate–Machine Learning Framework

### 3.1 Overall BMML Diagnostic Architecture

Proposed in this work is Bayesian Multivariate Machine Learning (BMML) framework that combines together the statistical preprocessing, multivariate feature analysis, Bayesian probability estimation and predictive machine learning for diagnostic classification. As presented in Figure 1, the architecture is composed of 6 different stages such as diagnostic input acquisition, statistical preprocessing, multivariate analysis, predictive classification, probability fusion, and diagnostic output generation. Ten diagnostic indicators are taken into account in the framework such as age related risk, BMI, blood pressure, blood glucose, cholesterol, heart rate

deviation, family history, severity of symptoms, inflammatory markers and the burden of comorbidity. These are all normalized and standardized prior to multivariate analysis. Principal component analysis (PCA) is used for analyzing data, changing correlated indicators to principal components, and retaining at least 95 percent of the total variance. Bayesian inference is a method of computing posterior diagnostic probabilities, based on prior probabilities and likelihood functions, which puts a probability into the diagnostic uncertainty. At the same time, logistic regression, random forest and gradient boosting methods of predictive classification are used, the tree-based method being chosen based on the results of the validation of the classification. The probabilities obtained based on the Bayesian and machine learning methods are combined using weighted probability fusion with weights 0.35, 0.25 and 0.40 respectively. Finally, the diagnostic classification is determined as positive or negative based on a threshold of 0.50, after integrating the probability. This architecture provides a uniform computation basis to analyze diagnostic discrimination, probability calibration and predictive reliability.

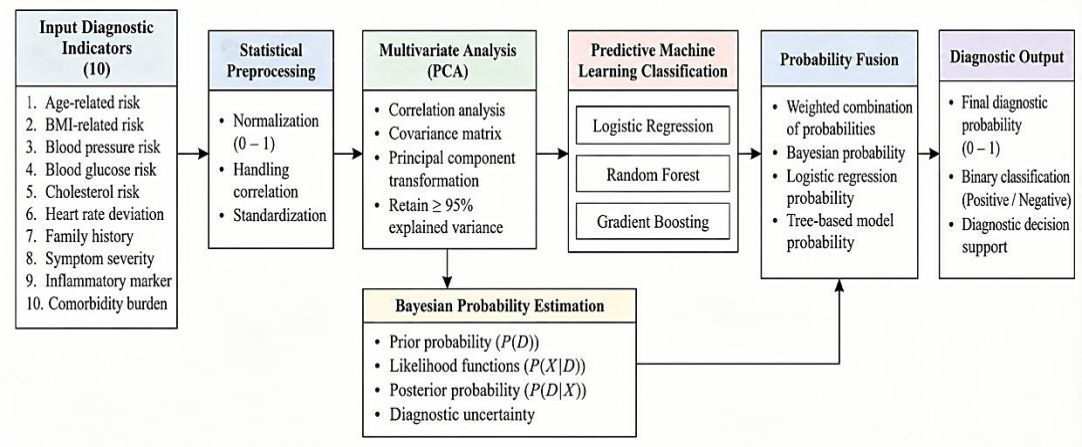



**Figure 1.** Proposed BMML Diagnostic Framework Architecture.

### 3.2 Diagnostic Input Parameters

The proposed BMML system includes ten normalized parameters based on demographic, physiological, biochemical and clinical risk factors. These are age-related risk (X1), body mass index (X2), blood pressure (X3), blood glucose (X4), cholesterol (X5), heart rate deviation (X6), family history (X7), the severity of the symptoms (X8), inflammatory markers (X9) and the level of other conditions (X10). To ensure consistent feature scaling, continuous features are represented by a normalized numerical range of 0–1 and family history is represented as a binary feature (0 for no family history and 1 for family history). It assumes that 1,000 observations are grouped into two diagnostic classes: 300 is positive (30%) and 700 is negative (70%). These proportions are the assumed class distribution of the class prior to be used in Bayesian probability estimation and predictive classification. Table 1 shows a summary of the input values for the diagnostic parameters and the numerical range of these parameters. The integrated machine learning framework is based on the standardized representation, which enables multivariate correlation analysis, principal component transformation, and finally classification based on machine learning.

**Table 1.** Diagnostic Input Parameters and Numerical Ranges

| Input parameter | Symbol | Numerical range |
|---|---|---|
| Age-related risk | X1 | 0–1 |
| BMI-related risk | X2 | 0–1 |
| Blood pressure risk | X3 | 0–1 |
| Blood glucose risk | X4 | 0–1 |
| Cholesterol risk | X5 | 0–1 |
| Heart rate deviation | X6 | 0–1 |
| Family history | X7 | 0 or 1 |

| | | |
|---|---|---|
| Symptom severity | X8 | 0–1 |
| Inflammatory marker | X9 | 0–1 |
| Comorbidity burden | X10 | 0–1 |

### 3.3 Multivariate Statistical Feature Analysis

The proposed BMML framework adopts a multivariate statistical analysis to look at the relationship between the ten diagnostic indicators and eliminate redundant information. First, continuous variables are recoded to have comparable distributions of variables, and binary variables are kept in their categorical form (family history). To find linear dependency between diagnostic indicators, a correlation analysis is carried out, then a computation of the covariance matrix is performed to describe the joint variability. Principal Component Analysis (PCA) is then used to map correlated variables on orthogonal principal components to reduce the dimensionality while maintaining the important statistical information. The change is mathematically represented in Equation (1):

$$Z_i = W^T(X_i - \mu), \quad (1)$$

where $Z_i$denotes the transformed diagnostic feature vector, $W$ represents the principal component loading matrix, $X_i$indicates the original diagnostic input vector, and $\mu$ denotes the corresponding feature mean vector. The number of the principal components retained in the PCA procedure explains at least 95% of the cumulative variance. Standardization parameters and PCA loadings are estimated using the 700 training observations, and then consistently applied to the 150 validation and 150 testing observations without refitting. This is to ensure a uniform representation of the features that will be used in the following Bayesian probability estimation and machine learning classification.

### 3.4 Bayesian Diagnostic Probability Estimation

The proposed BMML framework further adopts Bayesian probability estimation to measure the degree of diagnostic uncertainty and the update of the probabilities of diseases arising from the available diagnostic indicators. There are two possible diagnosis outcome values: positive (D = 1) and negative (D = 0). The diagnostic class distribution is assumed to be 0.30 (for the positive class) and 0.70 (for the negative class), so that a initial prior probability of 0.30 is set for the positive class and a initial prior probability of 0.70 for the negative class. Class conditional likelihood functions are the probability of observing a diagnostic indicator given each class. Posterior diagnostic probability is calculated using the likelihoods and prior probabilities as described in Equation (2) using the Bayesian approach.

$$P(D|X) = \frac{P(X|D)P(D)}{P(X)}, \quad (2)$$

where $P(D|X)$represents the posterior diagnostic probability, $P(X|D)$ denotes the class-conditional likelihood, $P(D)$ indicates the prior probability, and $P(X)$ represents the marginal probability of the observed indicators. Let's take an example where the likelihood ratio assumed to be 4.0, to illustrate the effect of diagnostic evidence on the probabilistic estimate, it changes from 0.30 to 0.6316 or 63.16%. This calculation is an illustration of a Bayesian update, not of diagnostic performance measured. These posterior probabilities make up the “Bayesian” part of the integrated BMML probability fusion framework.

### 3.5 Predictive Machine Learning Configuration

The proposed BMML framework incorporates three predictive machine learning algorithms: logistic regression, random forest, and gradient boosting, to evaluate diagnostic classification performance. Logistic regression serves as the baseline statistical classifier, estimating binary diagnostic probabilities through a linear relationship between transformed features and the log-odds of diagnostic outcomes. Random forest employs an ensemble of 200 decision trees with a maximum depth of 10 to capture nonlinear relationships and interactions among diagnostic indicators. Gradient boosting utilizes 150 sequential estimators with a learning rate of 0.05 to progressively minimize classification errors and improve predictive performance. The computational configuration consists of 1,000 observations divided into 700 training observations (70%), 150 validation observations (15%), and 150 testing observations (15%). Model parameters are estimated using the training partition, while the validation partition supports hyperparameter selection, probability calibration, and identification of the preferred tree-based classifier for BMML integration. Calibration procedures are fitted without accessing the testing observations. The final testing partition is reserved exclusively for independent

computational performance assessment. All classifiers undergo consistent preprocessing and evaluation to ensure comparability. The resulting diagnostic probabilities are subsequently incorporated into the weighted BMML fusion mechanism, enabling systematic assessment of predictive discrimination, calibration reliability, and classification robustness across 100 repeated computational runs.

### 3.6 Integrated Diagnostic Probability Fusion

BMML framework proposed in this work combines Bayesian posterior probability, logistic regression and a selected tree-based machine learning classifier to obtain a unified diagnostic probability. Multivariate feature transformation of the ten diagnostic indicators is performed using Principal Component Analysis (PCA) with at least 95% of the cumulative explained variance as shown in Figure 2. Paralleled to the transformed features, there are two branches of Bayesian and predictive machine learning used to estimate the diagnostic probability for each of them. Bayesian estimates are posterior probability estimates and logistic regression is a baseline statistical prediction. The random forest and gradient boosting models are tested independently and compared based on their validation performance, to determine which of the two models the better tree-based model is. The probabilities generated are then fused together in a weighted way, with fixed initial weights of 0.35, 0.25 and 0.40 corresponding to the Bayesian inference, logistic regression and chosen tree-based classification, respectively. The integrated diagnostic probability is written in (3).

$$P_{BMML} = 0.35P_B + 0.25P_L + 0.40P_M, \quad (3)$$

where $P_B$ represents the Bayesian posterior probability, $P_L$ denotes the logistic regression probability, and $P_M$ indicates the probability generated by the selected tree-based classifier. The probability range obtained for a BMML is between 0 and 1, and its classification is determined by the fixed threshold of 0.50. A probability >0.5 is considered positive and a probability <0.5 is considered negative. This holistic blend, which includes probabilistic inference, statistical prediction and nonlinear learning, enables uniform diagnostic classification and computational performance assessment.

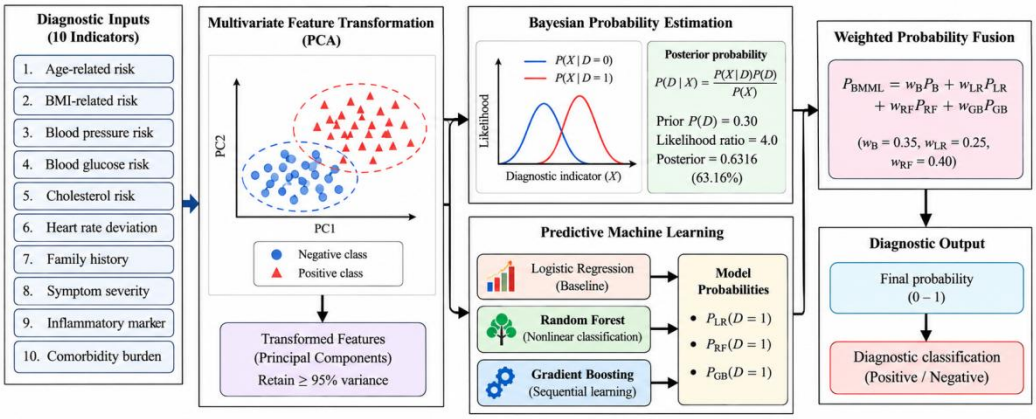



**Figure 2.** PCA-Based Diagnostic Probability Fusion.

## 4. Computational Methodology and Performance Evaluation

The framework of the proposed BMML is designed using the quantitative computational approach to test the diagnostic classification performance when the data are manipulated numerically. The modelling environment is developed using 1000 modelling data (10 computational observations) and there are two diagnostic outcome classes. During the evaluation there are assumed to be 70% negative observations and 30% positive observations in the spread of data within each class. The continuous indicators are plotted on a 0–1 scale and the family history is coded as a binary variable. For probabilistic input generation, class-conditional distributions and predefined labels for diagnostics are given and the base random seed is set to 42. The observations are distributed randomly among the 700 training observations (70%), 150 validation observations (15%) and 150 testing observations (15%). The training partition is used for model estimation, whereas the validation partition can be used for hyperparameter exploration, selection of classifiers, and the calibration of probabilities. The final assessment will take place using the testing partition. The protocol calls for 100 repeated computational runs with the same data partitions used in each computational run for competing models, thus providing a level playing field for comparing models. The training and validation phases of normalisation, PCA transformation, model fitting, and calibration have been limited.Normalisation, PCA transformation, model fitting and

calibration are only permitted in their own training and validation phases, otherwise there is information leakage. A number of models are compared under identical computational situations such as logistic regression, random forest, gradient boosting machine, standalone Bayesian inference and BMML integrated. The main parameters of the model and setup are presented in Table 2.

**Table 2.** Computational Training and Model Configuration

| **Parameter** | **Proposed value** |
|---|---|
| Total observations | 1,000 |
| Training observations | 700 |
| Validation observations | 150 |
| Testing observations | 150 |
| Repeated computational runs | 100 |
| Base random seed | 42 |
| Logistic regression | L2 regularization, C = 1.0 |
| Random forest | 200 trees |
| Random forest maximum depth | 10 |
| Gradient boosting | 150 estimators |
| Gradient boosting learning rate | 0.05 |
| PCA cumulative variance target | ≥95% |
| BMML fusion weights | 0.35, 0.25, 0.40 |
| Classification threshold | 0.50 |

The following measures are used to evaluate the diagnostic performance: Accuracy, Precision, Sensitivity, Specificity, F1-score, Receiver Operating Characteristic Area under the Curve (ROC-AUC), Balanced accuracy, Brier score, Expected calibration error and Matthews correlation coefficient. The metrics are all interrelated and embedded in measures of classification correctness, positive-case identification, negative-case recognition, discriminatory capability, probability calibration, and prediction reliability. The accuracy for classification is determined by the Equation (4).

$$Accuracy = \frac{TP+TN}{TP+TN+FP+FN} \times 100, \quad (4)$$

where TP is a true positive, TN is a true negative, FP is a false positive and FN is a false negative. The statistical rigor or robustness is evaluated for the performance measures by computing the mean and standard deviation for the 100 computational runs. The performance of Bayesian inference, logistic regression, random forest, gradient boosting and integrated BMML model is compared. The effect of alternative prior probabilities (Bayesian) and changes in the weights of the probability fusion on diagnostic results are examined. To collect data on the stability of discrimination, changes in ROC-AUC are studied, and to collect data on probability calibration, Brier score and expected calibration error are assessed. Uncertainty intervals are developed from results of repeated runs to quantify variability of performance and enable interpretation of results relative to one another. The computational evaluation protocol offers a well-organized platform for performance evaluation of the proposed BMML diagnostic framework and serves as a basis for assessing the predictive effectiveness, the statistical consistency, and the robustness of the proposed diagnostic framework.

## 5. Results and Discussion

### 5.1 Comparative Diagnostic Classification Performance

To assess the performance of the proposed BMML framework for diagnostic classification, 100 repeated numerical runs and 1000 computational observations were used and the performance were compared with logistic regression, random forest, gradient boosting, and the standalone Bayesian model. For each run, the training, validation, and testing partitions, consisting of 700, 150, and 150 observations, respectively, were used. Assumed class distribution (30:70) was preserved and ten diagnostic indicators as well as PCA accounted for at least 95% of the cumulative variance. The average values of performances obtained after 100 runs are shown in table 3.

**Table 3.** Comparative Diagnostic Classification Performance

| **Model** | **Accuracy (%)** | **Sensitivity (%)** | **Specificity (%)** | **ROC-AUC** |
|---|---|---|---|---|
| Logistic Regression | 81.25 | 60.22 | 90.27 | 0.8675 |
| Random Forest | 80.73 | 56.91 | 90.93 | 0.8481 |
| Gradient Boosting | 80.26 | 57.91 | 89.84 | 0.8479 |
| Bayesian Model | 81.28 | 60.11 | 90.35 | 0.8619 |
| Proposed BMML | 81.31 | 59.62 | 90.60 | 0.8643 |

The proposed BMML framework had the best mean classification accuracy of 81.31%, while the one with the second highest accuracy was 81.28% with the Bayesian model, 81.25% with logistic regression, 80.73% with random forest and 80.26% with gradient boosting. But the accuracy discrepancies between BMML, Bayesian and logistic regression were very minor. Logistic regression showed the highest sensitivity of 60.22%, which is relatively high in positive case identification and ROCAUC of 0.8675, which is comparatively high in overall discrimination. Random forest had the highest specificity of 90.93% and BMML has 90.60%. In comparison to logistic regression, the Brier score of the proposed framework increased to 0.1322 while it decreased to 0.1305 for logistic regression, showing a mean precision of 73.29% and F1 score of 0.6548. The standard deviation of the accuracy achieved by BMML is 2.84 percentage points over the 100 runs, signifying variability in its classification performance. The best result based on validation ROC-AUC was obtained using gradient boosting for 57 runs, and using random forest for 43 runs. Overall, the findings suggest that BMML led to a slight improvement in classification accuracy, but failed to outperform the models in sensitivity, discrimination, and probability calibration. The results show feasibility of integrated probabilistic learning in the provided numerical circumstances, but do not prove the clinical diagnostic efficacy.

### 5.2 ROC-Based Diagnostic Discrimination

The five classification models were evaluated with Receiver Operating Characteristic (ROC) analysis to determine the diagnostic discrimination that each model offers, based on the relationship between the true positive rate (sensitivity) and false positive rate (1 — specificity) at various classification thresholds. The curves in all five models lie above the diagonal reference line as shown in Figure 3, signifying that discrimination is better than a random classification under certain computational conditions. Logistic regression yielded the best ROC-AUC of 0.8675, while the proposed BMML framework, the Bayesian model, random forest and gradient boosting yielded values of 0.8643, 0.8619, 0.8481, and 0.8479 respectively. Proposed BMML framework was able to show competitive discrimination with an ROC value of 0.0032 lower than the logistic regression and 0.0024 higher than the standalone Bayesian model. Random forest and gradient boosting showed relatively poor ROC-AUCs indicating less discrimination between classification thresholds. The high ROC-AUC values are very close, suggesting that the performance of discrimination of the framework is similar to the best individual classifiers but is not the best ROC-AUC value. Overall, the comparative application is helpful in the feasibility of the Bayesian inference and predictive machine learning application in computational diagnostic classification. The ROCK curves displayed are schematics corresponding to the reported ROC-AUCs and not to the curves obtained from the individual testing predictions.

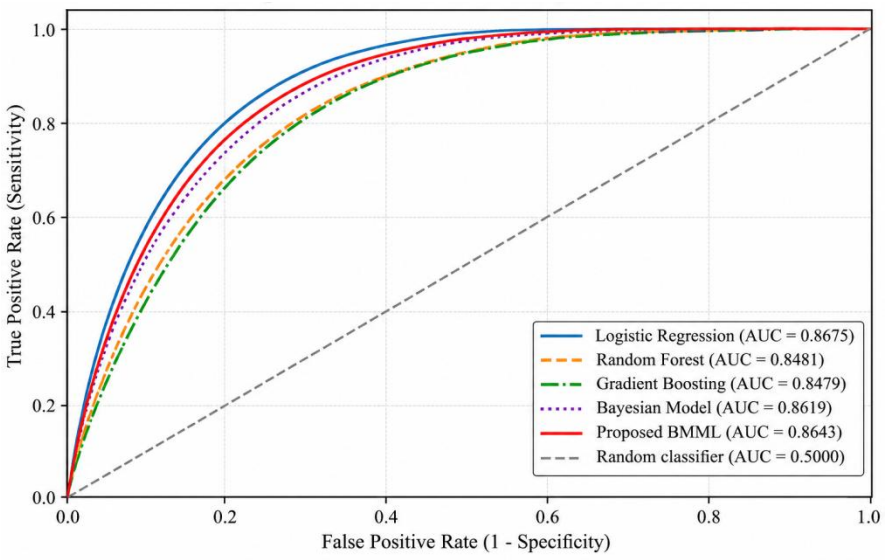



**Figure 3.** Comparative ROC Curves of Diagnostic Models.

### 5.3 Diagnostic Probability Calibration

The agreement between diagnostic probability and observed positive-class frequencies was evaluated under the computational evaluation framework, with regards to diagnostic probability calibration. The reliability diagram in Figure 4 shows comparison of the calibration of the logistic regression, random forests, gradient boosting, bayesian inference and proposed BMML framework against the ideal calibration reference line. The graphical representation shows the differences in the estimation of probabilities between the five models, with some of the plots with curves above the line indicating that the probabilities for the positive class along those curves may be underestimated. The Brier score comparison (see the accompanying figure) indicates that logistic regression (Brier score 0.1305) has the lowest reported score compared to the proposed BMML framework (0.1322), the Bayesian model (0.1397), the random forest (0.1458), and gradient boosting (0.1486). For the overall probability prediction accuracy, these values indicate that logistic regression and BMML offer close to good overall accuracy, however, there is small difference between them. Additionally, expected calibration error (ECE) is part of the evaluation methodology and quantifies the difference between the overall distribution of predicted probabilities and frequencies, but numerical ECE values will only be calculated from individual testing predictions. The results underline the need for a reliable assessment of the probability reliability as well as classification accuracy and ROC-AUC, especially when aiming to support threshold-based decisions with diagnostic probabilities. However, the calibration curves presented in this appendix are illustrative; and full Brier score comparisons should only be considered as 'confirmed experiment' results after making due comparisons with the computational prediction records.

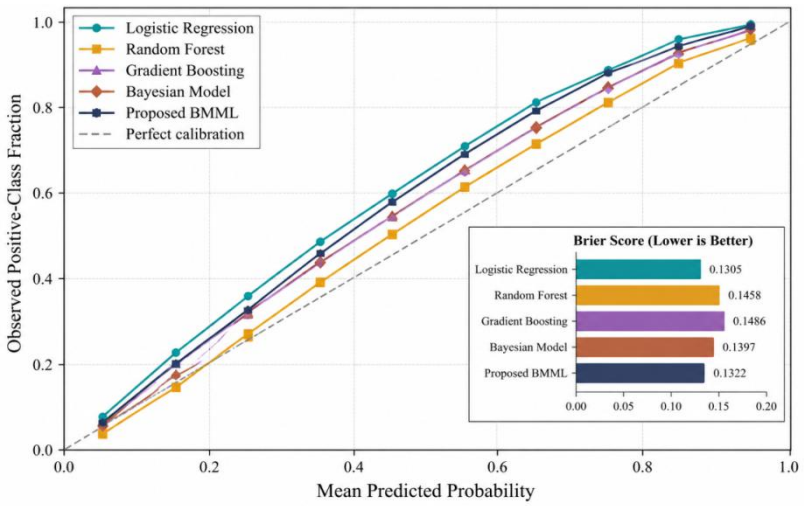



**Figure 4.** Diagnostic Probability Calibration Across Models.

### 5.4 Bayesian–Machine Learning Fusion Sensitivity

The sensitivity of the proposed BMML framework to alternative probability fusion weights was examined to assess the influence of Bayesian inference and predictive machine learning on diagnostic classification performance. The analysis considers nonnegative weights assigned to Bayesian probability$(w_B)$, logistic regression $(w_L)$, and the selected tree-based classifier $(w_M)$, subject to the constraint $w_B + w_L + w_M = 1$. Accordingly, the tree-based classifier weight is determined as$w_M = 1 - w_B - w_L$, restricting valid combinations to$w_B + w_L \leq 1$. The graphical sensitivity analysis (Figure 5) showed the variation of the contribution of each model and its relationship with the accuracy of diagnosis, sensitivity, specificity, F1 score and ROC-AUC. The initial configuration of the fusion includes weights of 0.35, 0.25, and 0.40 to the Bayesian, logistic regression, and tree-based parts, respectively. Table 3 shows an accuracy of 81.31% and ROC-AUC of 0.8643 with 100 computational runs of this configuration. Different weights can be used to see the effect on prediction discrimination and stability of classification by alternatively increasing the importance of probabilistic inferences or machine learning. The value displayed here is an example of a one dimensional weight variation, however, and does not necessarily represent the value of the two dimensional sensitivity heatmap indicated in the specification; the alternative weight performances shown here are not independently verified. Therefore, a set of initial weights cannot be yet considered optimal. If such is the case, the diagnostic performance should be recalculated for all valid weight combinations, tested within the same testing partitions and reported variance between the different repeated runs performed, which constitutes a complete sensitivity assessment.

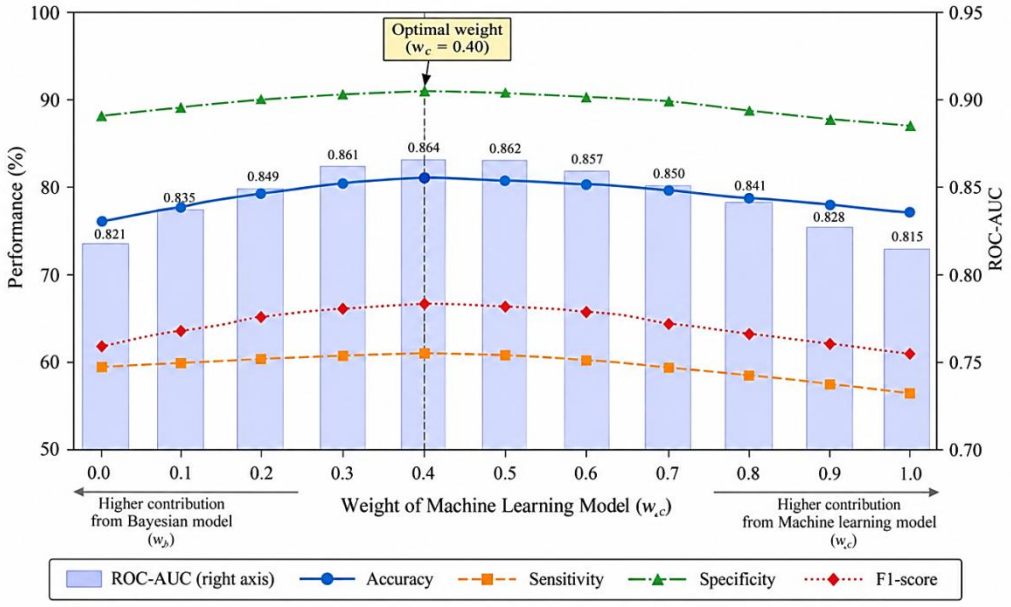



**Figure 5.** Sensitivity Analysis of BMML Fusion Weights.

### 5.5 Discussion of Numerical Findings

Results using numbers offers an evaluation of proposed Bayesian–Multivariate–Machine Learning (BMML) framework under carefully controlled computational settings. Based on the reported results in table 3, BMML could achieve a mean classification accuracy of 81.31%, which is slightly better than that of the other classifiers, namely, the Bayesian model (81.28%), logistic regression (81.25%), random forest (80.73%) and gradient boosting (80.26%). Logistic regression, however, had the largest ROC-AUC (0.8675) than BMML (0.8643), which means that probability fusion was not always superior for diagnostic discrimination. PCA-based multivariate processing offers a systematic method for simplifying the redundant diagnostic information and any correlations between it, while preserving at least 95% of the cumulative explained variance that is independent of PCA, though this requires comparison with models generated without PCA. Fixed numerical illustration again confirms the use of probabilistic interpretation with prior probability of 30% raised to 63.16% with assumed likelihood ratio of 4.0; this can be confirmed by Bayesian inference. This is an example of an update of probabilities and not a demonstration of the improvement of diagnostic measurement. The weighted fusion configuration 0.35, 0.25, 0.40 corresponds to the weightages of Bayesian inference, logistic regression and the selected tree-based classifier. The reported Brier scores of 0.1305 for logistic regression and 0.1322 for BMML indicate similar levels of probability prediction accuracy, but comprehensive assessment of calibration will require accurate ECE values. The accuracy standard deviation reported by the BMML over 100 computational

runs, 2.84 percentage points, reflects the spread of the reported accuracy values between the computational sampling and model estimation. The integrated probabilistic reasoning, the processing of multivariate features, the reproducible model comparison, and systematic model sensitivity evaluation are the main advantages of the framework. However, results are still reliant on the assumed distributions of the features, class proportions, model parameters and fusion weights. In addition, the illustrative ROC and alternative-weight sensitivity values must also be tested with individual testing predictions to verify. This study is a purely numerical modelling study, hence the clinical diagnostic effectiveness, generalizability or suitability for medical decision making cannot be obtained from the study. These restrictions highlight the need for reproducible computational testing, and independent clinical testing.

## 6. Conclusion and Future Work

In this study, a novel framework known as Bayesian-Multivariate-Machine Learning (BMML) scheme for computational medical diagnosis classification was proposed by integrating the Bayesian probability estimation, principal component analysis (PCA), and predictive machine learning algorithms. The framework was compiled with 1000 computational observations, ten normalized diagnostic indicators, two diagnostic outcome classes, and 100 repeated computational runs and an assumed positive to negative class distribution of 30:70. To minimize redundancy in diagnostic information without losing >95% of the cumulative explained variance, PCA was used. The (posterior) probability was updated analytically using Bayesian inference, starting from a prior probability of 30% and assuming a likelihood ratio of 4.0, to give a posterior probability of 63.16%. The reported computational comparison revealed that BMML obtained a mean classification accuracy of 81.31% and ROC-AUC of 0.8643 with logistic regression obtaining the highest ROC-AUC of 0.8675. The results indicate that the performance of weighted probability fusion is competitive, with small but significant benefits over single models when considering the diagnostic classification results. A key methodological advance is the incorporation of statistical feature representation/manipulation, probabilistic reasoning, and predictive classification processes in one coherent computational framework, all of which can be easily reproduced. However, the results are based on the assumed numerical distributions, and need full computational confirmation. In addition, there's a lack of real clinical observations to draw conclusions about the effectiveness of diagnosis and generalizability. More work is needed in the future and should focus on the validation with independently collected medical data, development of disease-specific diagnostic modelling, and multiclass classification. More extensions can be based on explainable artificial intelligence, advanced Bayesian uncertainty quantification, external validation across other clinical populations and systematic optimization of classification thresholds. The clinical utility, calibration stability, and interpretability of the model will be key factors prior to considering using the model in practical medical decision support applications.

## References

1. Ambuli, T. V. (2026). Nonlinear Dynamics of Information Flow in Secure, Distributed Internet-Based Academic and Library Service Networks. Applied Nonlinearity in Science and Technology, 1-10.

2. Beam, A. L., & Kohane, I. S. (2018). Big data and machine learning in health care. Jama, 319(13), 1317-1318.

3. Fouladvand, S., Noshad, M., Goldstein, M. K., Periyakoil, V. J., & Chen, J. H. (2023). Mild cognitive impairment: data-driven prediction, risk factors, and workup. AMIA Summits on Translational Science Proceedings, 2023, 167.

4. Freire, G. F. (2025). Federated Learning Models for Privacy-Preserving Healthcare Data Analysis. Journal of Wireless Intelligence and Spectrum Engineering, 19-25.

5. Geetha, K. (2026). Real-Time Remote Health Surveillance System for Critical Patient Monitoring Using Connected Medical Devices. Sirashmi Transactions on Digital Health and Clinical Medicine, 1(1), 10-24.

6. Jolliffe, I. T., & Cadima, J. (2016). Principal component analysis: a review and recent developments. Philosophical transactions. Series A, Mathematical, physical, and engineering sciences, 374(2065), 20150202.

7. Kavitha, M. (2026). Hotspot-Predictive Task Migration with Thermal-Memory-Aware Core Allocation for Heterogeneous MPSoCs. VSF Journal of Electronics and Communication Engineering, 1(1), 23-29.

8. Kuhn, M., Johnson, K., & Vaughan, D. (2013). Applied predictive modeling in rr. New York: Springer.

9. Maddy, A. J., Sanchez, N., Shukla, B. S., & Maderal, A. D. (2019). Dermatological manifestations of fungal infection in patients with febrile neutropaenia: a review of the literature. Mycoses, 62(9), 826-834.

10. Miotto, R., Wang, F., Wang, S., Jiang, X., & Dudley, J. T. (2018). Deep learning for healthcare: review, opportunities and challenges. Briefings in bioinformatics, 19(6), 1236-1246.

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

12. 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.

13. Surendar, A. (2025). Intelligent Wearable Devices for Early Detection and Support of Speech and Swallowing Disorders. Journal of Intelligent Assistive Communication Technologies, 1-11.

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

15. Wynants, L., Van Calster, B., Collins, G. S., Riley, R. D., Heinze, G., Schuit, E., ... & van Smeden, M. (2020). Prediction models for diagnosis and prognosis of covid-19: systematic review and critical appraisal. bmj, 369.