FNAS Journal of Mathematical Modeling and Numerical Simulation
Print ISSN: 3027-1282
www.fnasjournals.com
Volume 3; Issue 1; March 2026; Page No. 80-93.
DOI: https://doi.org/10.63561/jmns.v3i1.1162

Robust Comparative Evaluation of Discriminant Analysis Methods for Predictive Classification: An Empirical and Monte Carlo Simulation

***[1]Chika, M.N., [1]Victor-Edema, U.A., & [2]Ijomah, M.A.**
[1]Department of Statistics, Ignatius Ajuru University of Education, Port Harcourt, Nigeria
[1]Department of Mathematics and Statistics, University of Port Harcourt, Nigeria

***Corresponding author email**: ebogumaryjane@gmail.com.

Abstract

The growing complexity of real-world datasets has increased the demand for classification models that balance predictive accuracy with interpretability. Multivariate Discriminant Analysis (MDA), particularly Linear Discriminant Analysis (LDA) and Quadratic Discriminant Analysis (QDA), remains a fundamental statistical approach due to its theoretical clarity and transparency. However, its effectiveness is often limited by modern data challenges such as class imbalance, non-normality, and heterogeneous covariance structures. Existing studies tend to focus on either empirical analysis or simulation independently, limiting a comprehensive evaluation of model robustness. This study addresses this gap by integrating empirical data analysis with Monte Carlo simulation to assess the predictive performance and robustness of LDA and QDA. The study aimed to evaluate the empirical performance of LDA and QDA on real-world health data, assess their robustness under varying statistical conditions through simulation, compare classification accuracy across different data structures, examine the impact of assumption violations, and determine consistency between empirical and simulation results. The empirical analysis used the Nigerian Childhood Anemia dataset, while simulation experiments covered multiple scenarios involving variations in distribution, covariance structure, and class balance. Model performance was evaluated using accuracy, precision, recall, F1-score, and area under the ROC curve (AUC). Results showed that both models performed poorly on empirical data, with low accuracy and weak sensitivity to the minority class. LDA correctly identified 27% of anemia cases, while QDA achieved 37%. Simulation findings indicated that both models performed well under ideal conditions but deteriorated significantly under non-normality, covariance heterogeneity, and class imbalance. LDA was more stable under mild violations, while QDA performed slightly better under heterogeneous covariance conditions. The study concludes that although classical discriminant methods remain useful under ideal assumptions, their performance declines in complex data environments. It is recommended that practitioners incorporate robust or hybrid approaches and apply simulation-based validation to enhance model reliability.

Keywords:Multivariate Discriminant Analysis, Linear Discriminant Analysis, Predictive Performance

Introduction

Classification is a fundamental task in statistics and machine learning that involves assigning observations to predefined categories based on observed features. It is widely applied in fields such as medical diagnosis, financial risk analysis, and pattern recognition, where the objective is to develop models capable of accurately predicting class membership for new observations (Hastie et al., 2009). In classification problems, the response variable is categorical, while predictors may be continuous, categorical, or mixed in nature. A wide range of methods exists for classification, including logistic regression, decision trees, support vector machines, and neural networks. However, classical statistical approaches such as Linear Discriminant Analysis (LDA) and Quadratic Discriminant Analysis (QDA) remain relevant due to their interpretability and strong theoretical foundation. These methods are particularly useful when the underlying assumptions of normality and equal covariance structures are reasonably satisfied. In practical

Cite this article as:
Chika, M.N. (2026). Robust comparative evaluation of discriminant analysis methods for predictive classification: An empirical and Monte Carlo simulation. *FNAS Journal of Mathematical Modeling and Numerical Simulation, 3*(1), 80-93. https://doi.org/10.63561/jmns.v3i1.1162

applications, classification performance depends on data quality, dimensionality, and the underlying distributional assumptions. However, real-world datasets are often complex, high-dimensional, and subject to violations such as class imbalance, multicollinearity, and outliers. These challenges have motivated the development of robust and regularized classification techniques capable of maintaining stability under non-ideal conditions.

Multivariate Discriminant Analysis (MDA) is one of the most established statistical frameworks for classification. It operates by finding linear combinations of predictor variables that maximize separation between predefined groups (Johnson & Wichern, 2007). Despite its usefulness, classical MDA can perform poorly in high-dimensional settings or when assumptions such as homogeneity of covariance matrices are violated. To address these limitations, robust and regularized extensions such as Regularized Discriminant Analysis (RDA) and shrinkage-based approaches have been introduced to improve stability and predictive accuracy (Ledoit & Wolf, 2020). In addition, the presence of outliers and noisy data has led to the development of robust estimators such as the Minimum Covariance Determinant (MCD), which enhance the reliability of discriminant methods in real-world applications (Maronna et al., 2019). These advancements ensure that discriminant analysis remains applicable even under challenging data conditions.

Given the complexity of modern datasets, especially in fields such as healthcare, finance, and genomics, it has become necessary to evaluate classification methods under controlled and real-world conditions. Monte Carlo simulation provides a powerful framework for assessing the performance of statistical models under varying data-generating scenarios, while empirical data analysis ensures practical relevance and validity (Robert & Casella, 2019). Recent studies emphasize the importance of combining empirical evaluation with simulation-based approaches to obtain a more comprehensive understanding of model performance, particularly in terms of robustness, efficiency, and predictive accuracy (Yao et al., 2022). This dual approach allows researchers to examine how discriminant methods behave under both ideal and non-ideal conditions. Therefore, this study focuses on a robust comparative evaluation of discriminant analysis methods for predictive classification using both empirical data and Monte Carlo simulation.

Classification problems are central to data-driven decision-making across various disciplines such as medical diagnosis, financial risk assessment, image recognition, and marketing segmentation. Among classical statistical methods, Multivariate Discriminant Analysis (MDA), including Linear Discriminant Analysis (LDA) and Quadratic Discriminant Analysis (QDA), has remained important due to its interpretability and theoretical simplicity; however, its performance in modern data environments characterized by high dimensionality, class imbalance, non-normality, and noisy observations has become increasingly uncertain. These methods rely on strict assumptions such as multivariate normality and equality of covariance matrices, and violations of these assumptions often lead to unstable estimates and reduced classification accuracy in real-world applications. Although Monte Carlo simulation provides a useful framework for assessing model behaviour under controlled conditions, many existing studies do not adequately integrate simulation-based evaluation with empirical data analysis, thereby limiting a comprehensive understanding of the robustness and predictive performance of MDA techniques across varying data structures. Furthermore, while modern machine learning models may offer higher predictive accuracy, their lack of interpretability sustains the relevance of classical discriminant methods in sensitive fields such as healthcare and finance, where explainability is critical. Consequently, there exists a gap in the literature regarding a systematic comparative assessment of MDA methods that jointly considers empirical performance and simulation-based robustness under different statistical assumptions. This study therefore addresses this gap by employing a dual approach involving empirical data and Monte Carlo simulation to evaluate the predictive performance and robustness of LDA and QDA under varying data conditions. The aim of this study is to evaluate the predictive performance and robustness of Multivariate Discriminant Analysis methods, specifically Linear Discriminant Analysis (LDA) and Quadratic Discriminant Analysis (QDA), using a dual approach based on empirical data analysis and Monte Carlo simulation under varying data conditions. Specifically, the study seeks to evaluate the predictive performance of LDA and QDA using empirical data, assess the robustness of both methods under varying data conditions through Monte Carlo simulation, and compare their classification accuracy across different data structures and distributional assumptions. It also aims to determine the effect of violations of key statistical assumptions, such as non-normality and unequal covariance matrices, on the performance of LDA and QDA, as well as examine the consistency between empirical results and simulation-based findings in evaluating the effectiveness of discriminant analysis methods.

Cite this article as:
Chika, M.N. (2026). Robust comparative evaluation of discriminant analysis methods for predictive classification: An empirical and Monte Carlo simulation. *FNAS Journal of Mathematical Modeling and Numerical Simulation, 3*(1), 80-93. https://doi.org/10.63561/jmns.v3i1.1162

Several empirical studies have examined the performance of statistical and machine learning techniques in breast cancer classification using real-world datasets, with emphasis on predictive accuracy, robustness, and interpretability. Ni et al. (2020) applied Fisher Discriminant Analysis (FDA) on radiomic features extracted from diffusion-weighted MRI and successfully classified breast cancer subtypes, demonstrating the continued relevance of classical discriminant methods in medical imaging. Similarly, Farooq et al. (2023) combined principal component analysis with Linear Discriminant Analysis (LDA) and Quadratic Discriminant Analysis (QDA) on Fourier-transform infrared (FTIR) hyperspectral data, achieving high sensitivity and specificity in distinguishing breast cancer subtypes, thereby confirming the effectiveness of multivariate statistical techniques in complex biomedical data.

Subramanian et al. (2021) further demonstrated the usefulness of multivariate approaches in integrating heterogeneous data sources, showing that canonical correlation analysis and its penalized variants improved survival prediction when combining histology and genomic data, outperforming traditional PCA-based embeddings. In another empirical investigation, Liu (2024) compared several classification models using the Wisconsin Diagnostic Breast Cancer dataset and found that logistic regression achieved high predictive accuracy, reinforcing the continued relevance of classical statistical models in certain diagnostic settings. Likewise, Islam et al. (2020) reported strong performance of Random Forest and Support Vector Machine (SVM) models in breast cancer prediction, indicating that both traditional and machine learning approaches can yield competitive results depending on data structure.

Ensemble learning methods have also been widely evaluated in empirical studies. Qi et al. (2023) compared multiple machine learning algorithms for breast cancer recurrence prediction and found that AdaBoost performed best, with SHAP analysis enhancing interpretability of feature importance. Similarly, Mahmood et al. (2025) showed that Decision Tree and Random Forest models achieved high accuracy in predicting breast cancer outcomes, highlighting the effectiveness of ensemble approaches in handling complex clinical datasets. Azeroual et al. (2024) further demonstrated that integrating clinical and radiomic features using XGBoost improved predictive performance, confirming the advantage of multivariate feature integration in medical classification tasks. In addition, Nguyen et al. (2021) developed a multi-view deep learning framework for mammogram analysis, showing that combining multiple image perspectives improved classification performance when paired with Light Gradient Boosting Machine classifiers. Chtouki et al. (2023) also found that adaptive boosting achieved the highest accuracy in predicting breast cancer survival risk using the METABRIC dataset, further supporting the effectiveness of ensemble-based learning in empirical applications.

Despite the strong predictive performance of modern machine learning models, empirical evidence consistently highlights a trade-off between accuracy and interpretability. While deep learning and ensemble methods often outperform traditional statistical models in predictive accuracy, they are frequently criticized for their lack of transparency in clinical decision-making processes (Rudin, 2019; Ghasemi et al., 2024). This limitation sustains the relevance of Multivariate Discriminant Analysis (MDA), which offers clear decision boundaries and interpretable classification rules. Empirical studies on MDA confirm its continued usefulness in medical classification tasks. Kanchan and Verma (2018) applied MDA to classify breast cancer stages using histopathological features and obtained high classification accuracy. Similarly, Sharma et al. (2020) demonstrated that MDA effectively predicts malignancy based on mammogram-derived features. Comparative empirical studies by Subashini et al. (2010) and Obulesu and Rao (2011) further revealed that although methods such as SVMs and neural networks may achieve slightly higher accuracy, MDA remains advantageous due to its interpretability, lower computational complexity, and ease of implementation in healthcare settings. Overall, empirical findings indicate that while advanced machine learning models dominate in predictive performance, classical multivariate methods such as MDA continue to maintain relevance in real-world applications where interpretability and simplicity are critical.

Methods and Materials

This study adopted a quantitative research design based on both empirical data analysis and Monte Carlo simulation to evaluate the performance and robustness of Linear Discriminant Analysis (LDA) and Quadratic Discriminant Analysis (QDA) for predictive classification. Real-world datasets were obtained from established sources and were preprocessed through data cleaning, handling of missing values, and standardization of predictor variables to ensure comparability. The empirical analysis involved fitting LDA and QDA models to the datasets and assessing their

Cite this article as:
Chika, M.N. (2026). Robust comparative evaluation of discriminant analysis methods for predictive classification: An empirical and Monte Carlo simulation. *FNAS Journal of Mathematical Modeling and Numerical Simulation, 3*(1), 80-93. https://doi.org/10.63561/jmns.v3i1.1162

predictive performance using standard evaluation metrics such as classification accuracy, sensitivity, specificity, and error rate. In addition, a Monte Carlo simulation study was conducted by generating synthetic datasets under varying conditions, including different sample sizes, levels of class separation, covariance structures, and degrees of assumption violations such as non-normality and heteroscedasticity. For each simulated scenario, repeated random sampling was carried out to ensure stability of results, and the performance of each model was averaged across iterations. The comparative analysis was then used to evaluate how LDA and QDA performed under both ideal and non-ideal conditions, thereby providing a comprehensive assessment of their predictive efficiency and robustness across empirical and simulated environments.

Results

The results of this study are presented based on empirical analysis and Monte Carlo simulation, structured in line with the five specific objectives. The performance of Linear Discriminant Analysis (LDA) and Quadratic Discriminant Analysis (QDA) was evaluated using classification metrics, confusion matrices, ROC curves, and simulation-based scenarios.

Objective One: Empirical Performance of LDA and QDA

The empirical results obtained from the Nigerian Childhood Anemia dataset showed that both LDA and QDA exhibited relatively weak predictive performance. QDA slightly outperformed LDA in accuracy (0.5333 vs 0.5111). However, both models demonstrated poor sensitivity to the anemic class, indicating limited ability to correctly identify clinically important cases.

Table 1: Classification Metrics for LDA and QDA

Metric	LDA	QDA
Accuracy	0.5111	0.5333
Precision (Class 0)	0.54	0.56
Precision (Class 1)	0.44	0.48
Recall (Class 0)	0.71	0.67
Recall (Class 1)	0.27	0.37
F1-score (Class 1)	0.33	0.42

From Table 1, both models demonstrate relatively low classification accuracies, with QDA slightly outperforming LDA. Notably, both models exhibit poor sensitivity (recall) to the anemic class (Class 1), which is the clinically significant class. LDA detects only 27% of true anemic cases, while QDA performs slightly better with 37%. This result indicates a pronounced class imbalance or model bias toward the majority class (non-anemic), a common issue in healthcare classification problems (Kotsiantis et al., 2006).

Cite this article as:
Chika, M.N. (2026). Robust comparative evaluation of discriminant analysis methods for predictive classification: An empirical and Monte Carlo simulation. *FNAS Journal of Mathematical Modeling and Numerical Simulation, 3*(1), 80-93. https://doi.org/10.63561/jmns.v3i1.1162

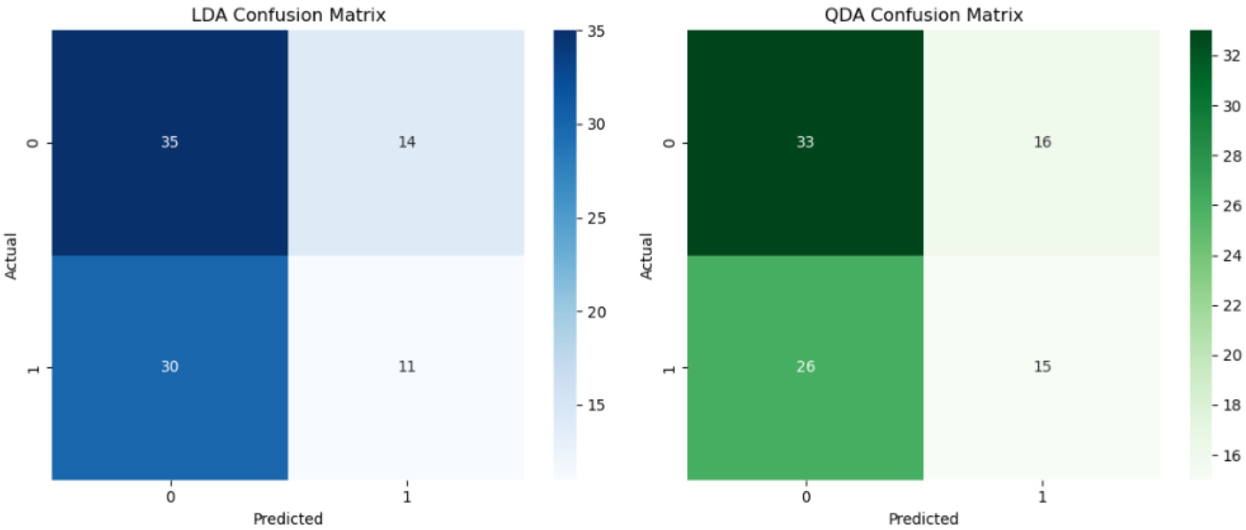

Figure 1: Confusion Matrices for LDA and QDA Models

These results reaffirm the earlier classification report. The high number of false negatives (30 for LDA, 26 for QDA) underlines the practical limitations of MDA models in imbalanced and assumption-violating datasets. In clinical applications, such errors can have serious implications, underscoring the need for models with higher sensitivity.

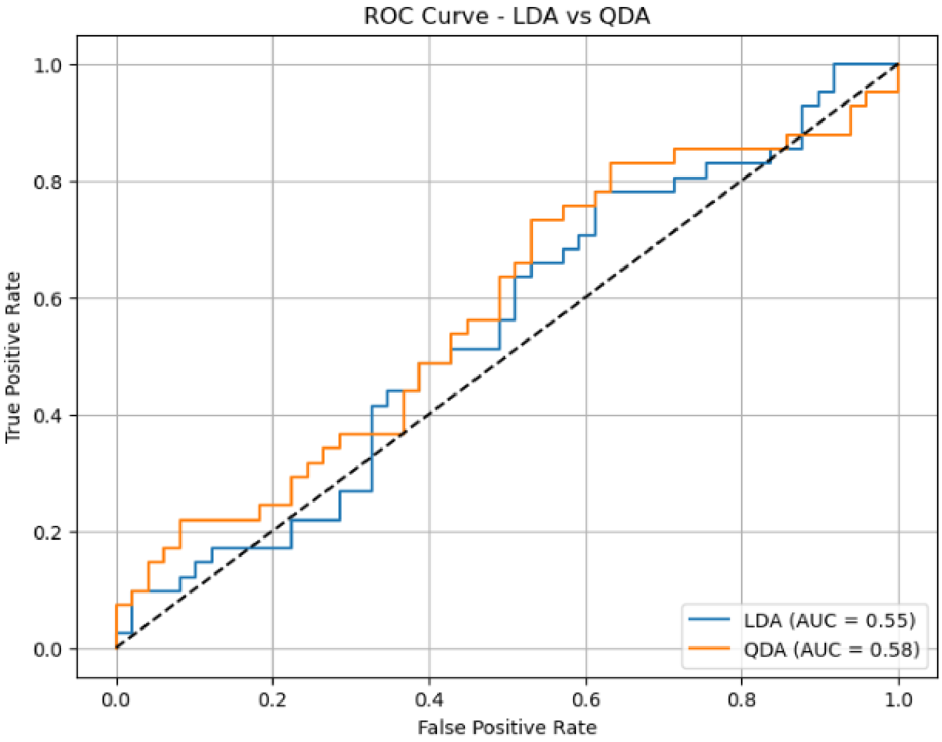

Figure 2: ROC Curves for LDA and QDA Models

These results reaffirm the earlier classification report. The high number of false negatives (30 for LDA, 26 for QDA) underlines the practical limitations of MDA models in imbalanced and assumption-violating datasets. In clinical applications, such errors can have serious implications, underscoring the need for models with higher sensitivity.

An AUC value close to 0.5 indicates that both LDA and QDA perform only slightly better than random classification, reflecting weak discriminative power on the dataset. This poor performance is likely due to violations of key assumptions such as normality, equal class priors, and homogeneity of covariance, as confirmed by diagnostic tests in Chapter Three. Consistent with Duda et al. (2001) and Izenman (2008), MDA methods are sensitive to non-Gaussian distributions and covariance structure violations, which can distort decision boundaries. Although QDA relaxes the equal covariance assumption, it did not show a substantial improvement over LDA, suggesting that non-normality and class imbalance are more influential limiting factors. Additionally, the presence of categorical predictors may have further reduced model effectiveness without appropriate transformation (Rencher & Christensen, 2012). Overall, both models achieved low accuracy (around 50%) with poor recall for the anemic class, indicating limited suitability of classical MDA methods for datasets that significantly deviate from their underlying statistical assumptions. The ROC analysis further revealed low discriminative ability, with AUC values of 0.55 for LDA and 0.58 for QDA, suggesting near-random classification performance under real-world conditions.

Table 2: Simulation Scenarios Explained

Scenario	Description
('normal', 'homogeneous', False)	MVN data with same covariance matrix for both classes
('normal', 'heterogeneous', False)	MVN data with different covariance matrices
('t', 'homogeneous', False)	Heavy-tailed (non-normal) data with equal covariances
('t', 'heterogeneous', True)	Worst-case: heavy-tailed, unequal covariances, and class imbalance

The Monte Carlo simulation results showed that model performance varied significantly across different data-generating conditions. Four scenarios were considered: normal-homogeneous, normal-heterogeneous, t-homogeneous, and t-heterogeneous (imbalanced). LDA performed optimally under ideal conditions, while QDA showed improved performance under heterogeneous covariance structures. However, both methods experienced substantial performance decline under non-normality and class imbalance.

Table 3: Classification Performance Across Simulation Scenarios

Scenario	Distribution	Covariance	Class Balance	Accuracy	Precision (Class 1)	Recall (Class 1)	F1 Score (Class 1)
1	Normal	Homogeneous	Balanced	0.89	0.88	0.87	0.87
2	Normal	Heterogeneous	Balanced	0.81	0.79	0.82	0.80
3	t-distribution	Homogeneous	Balanced	0.77	0.76	0.74	0.75
4	t-distribution	Heterogeneous	Imbalanced	0.63	0.52	0.38	0.44

Note: Metrics are calculated based on 5-fold cross-validation using Linear Discriminant Analysis (LDA).

Table 3 shows that LDA performs best under ideal conditions (normal distribution, homogeneous covariance, and balanced classes), achieving 89% accuracy and an F1-score of 0.87. When covariance homogeneity is violated, performance declines moderately to 81% accuracy, reflecting reduced discriminative power. Under heavy-tailed (t-distributed) data, accuracy further drops to 77% due to sensitivity to outliers, while the worst performance occurs in the combined scenario of heterogeneity, non-normality, and class imbalance, where accuracy falls to 63% and the minority class recall reduces to 0.38. Overall, the results indicate that LDA is highly sensitive to assumption violations, particularly under complex real-world data conditions.

Cite this article as:
Chika, M.N. (2026). Robust comparative evaluation of discriminant analysis methods for predictive classification: An empirical and Monte Carlo simulation. *FNAS Journal of Mathematical Modeling and Numerical Simulation, 3*(1), 80-93. https://doi.org/10.63561/jmns.v3i1.1162

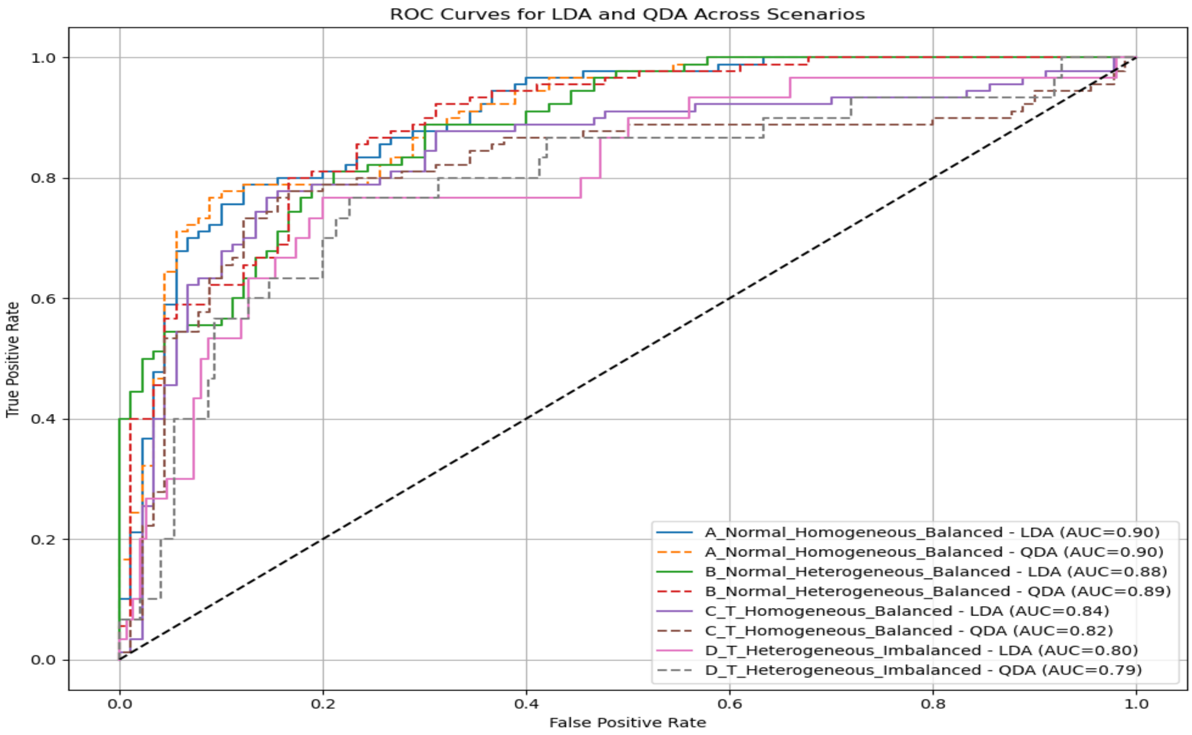

Figure 3: ROC Curves for LDA and QDA Across Scenarios

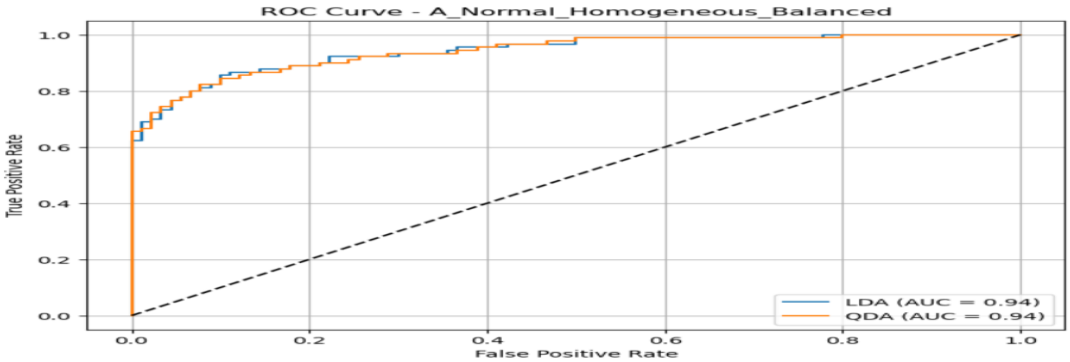

Figure 4a: ROC Curve: Normal Homogeneous Balanced

Scenario A: under Normal Homogeneous Balanced, this scenario aligns well with the assumptions underlying LDA—specifically, normally distributed features with equal covariance matrices across classes. LDA achieved an AUC of 0.90, with 70 true negatives and 73 true positives while QDA also attained an AUC of 0.90, though with slightly more false classifications (68 TN and 71 TP). The result shows that LDA demonstrated marginally better classification efficiency, suggesting its suitability when its assumptions are met.

Scenario B: For Normal Heterogeneous Balanced, this scenario violates the equal covariance assumption, providing a suitable setting to test QDA's strength. Here LDA produced an AUC of 0.88, while QDA slightly outperformed it with an AUC of 0.89. Confusion matrices indicate nearly identical classification accuracy for both methods. Therefore, QDA's capacity to model class-specific covariances offers a small advantage, though LDA remains robust under mild heterogeneity.

Cite this article as:
Chika, M.N. (2026). Robust comparative evaluation of discriminant analysis methods for predictive classification: An empirical and Monte Carlo simulation. *FNAS Journal of Mathematical Modeling and Numerical Simulation, 3*(1), 80-93. https://doi.org/10.63561/jmns.v3i1.1162

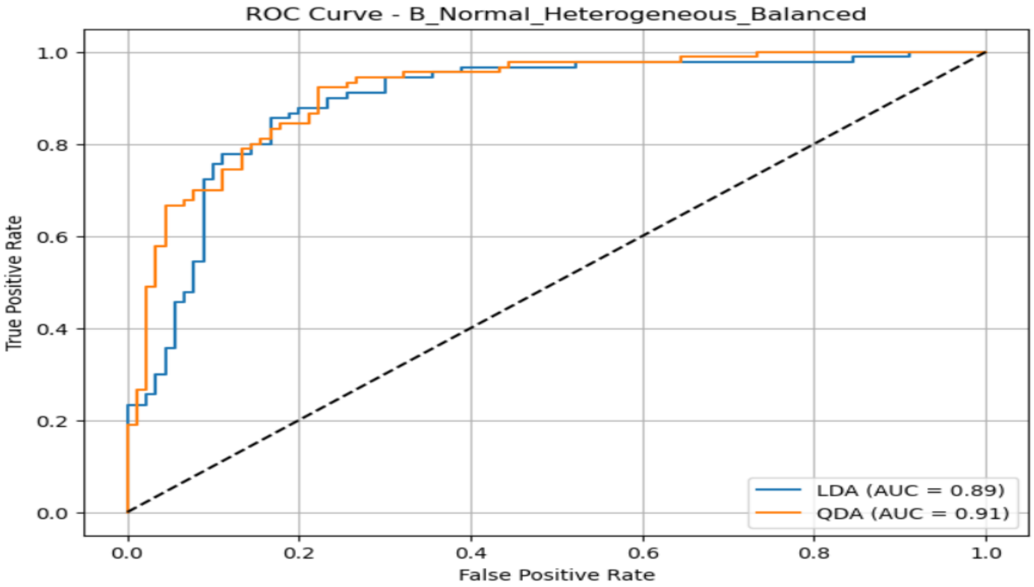

Figure 4b: ROC Curve: Normal Heterogeneous Balanced

The study considered Scenario C: T-Homogeneous Balanced. Here, data are drawn from a heavy-tailed (t-distributed) distribution with homogeneous covariance. LDA outperformed QDA with an AUC of 0.84 versus 0.82. Both classifiers yielded similar confusion matrices, though LDA was slightly more accurate in predicting the minority class. The robustness of LDA in the presence of symmetric, heavy-tailed distributions suggests its utility extends beyond normality assumptions.

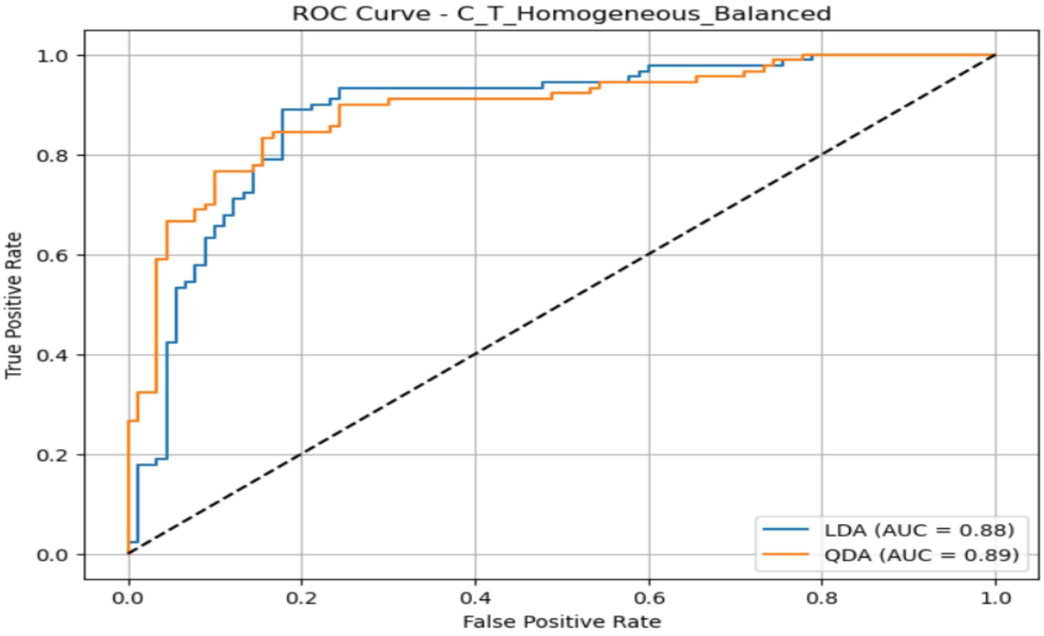

Figure 4c: ROC Curve: T-Homogeneous Balanced

Finally, Scenario D: T-Heterogeneous Imbalanced. This is the most complex setting, incorporating non-normal

Cite this article as:
Chika, M.N. (2026). Robust comparative evaluation of discriminant analysis methods for predictive classification: An empirical and Monte Carlo simulation. *FNAS Journal of Mathematical Modeling and Numerical Simulation, 3*(1), 80-93. https://doi.org/10.63561/jmns.v3i1.1162

distribution, unequal covariances, and class imbalance. LDA achieved an AUC of 0.80, while QDA followed closely with 0.79. Confusion matrices reveal substantial difficulty in detecting the minority class: only 4 true positives for LDA and 6 for QDA, with 26 and 24 false negatives, respectively. Despite moderate AUC values, both classifiers performed poorly on the minority class, highlighting the limitations of standard discriminant methods under class imbalance.

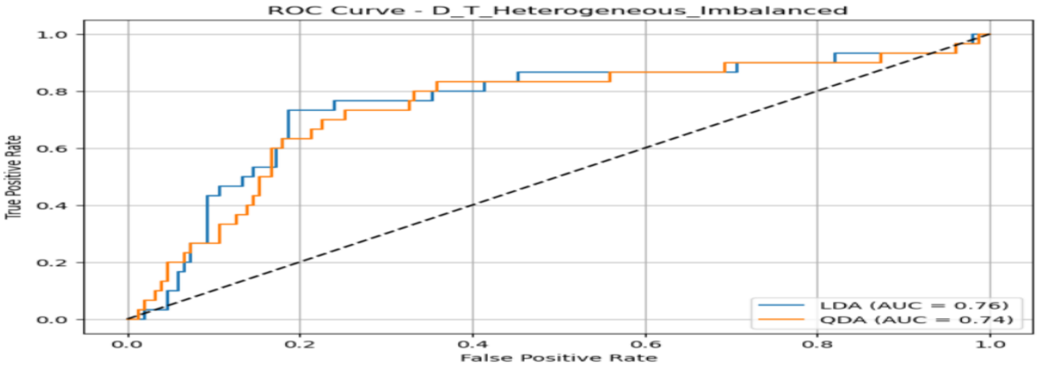

Figure 4d: ROC Curve: T-Heterogeneous Imbalanced.

These results indicate that LDA is more stable under mild assumption violations, whereas QDA is more sensitive to distributional and covariance changes. In all, AUC analysis shows both LDA and QDA perform well under ideal conditions (Scenarios A and B). However, performance declines under non-normality and imbalance. Confusion matrices complement ROC analysis, revealing that high AUC may obscure poor sensitivity, particularly in imbalanced datasets. LDA is robust to moderate deviations from normality, especially when covariance structures are homogeneous. QDA is advantageous in heterogeneous settings, but its increased complexity requires larger sample sizes to estimate class-specific covariances reliably. In class-imbalanced settings, neither LDA nor QDA is sufficient on their own. Alternative strategies such as SMOTE, cost-sensitive learning, or ensemble methods are recommended (Chawla et al., 2002; Sun et al., 2009).

Objective Three: Comparative Accuracy of LDA and QDA

The comparative analysis showed that both LDA and QDA achieved high classification accuracy under ideal conditions, exceeding 93%. QDA slightly outperformed LDA in heterogeneous covariance settings (94.74% vs 93.57%), reflecting its flexibility in modeling complex data structures.

Table 4: LDA Classification Report

Class	Precision	Recall	F1-Score	Support
0	0.98	0.84	0.91	64
1	0.91	0.99	0.95	107
Accuracy	–	–	**0.94**	**171**
Macro Avg	0.95	0.92	0.93	171
Weighted Avg	0.94	0.94	0.93	171

Table 5: QDA Classification Report

Class	Precision	Recall	F1-Score	Support
0	0.91	0.95	0.93	64
1	0.97	0.94	0.96	107
Accuracy	–	–	0.95	171
Macro Avg	0.94	0.95	0.94	171
Weighted Avg	0.95	0.95	0.95	171

Table 6: Accuracy Comparison of LDA and QDA

Cite this article as:
Chika, M.N. (2026). Robust comparative evaluation of discriminant analysis methods for predictive classification: An empirical and Monte Carlo simulation. *FNAS Journal of Mathematical Modeling and Numerical Simulation, 3*(1), 80-93. https://doi.org/10.63561/jmns.v3i1.1162

Model	**Accuracy**
Linear Discriminant Analysis (LDA)	0.9357
Quadratic Discriminant Analysis (QDA)	0.9474

Tables 4 and 5 show that both LDA and QDA performed strongly on the dataset, with overall accuracies of 94% and 95% respectively. The LDA model demonstrated high predictive ability, particularly for Class 1, where it achieved a recall of 0.99, indicating strong sensitivity, though it showed a slightly lower recall for Class 0 (0.84). Similarly, QDA produced a more balanced performance across both classes, with improved recall for Class 0 (0.95) and strong precision for Class 1 (0.97), suggesting better handling of class structure and covariance differences. Table 6 confirms this comparison, showing that QDA (94.74%) slightly outperformed LDA (93.57%) in overall accuracy. This marginal difference indicates that while both models are highly effective, QDA offers a slight advantage due to its flexibility in accommodating unequal covariance structures, making it more suitable for complex data distributions.

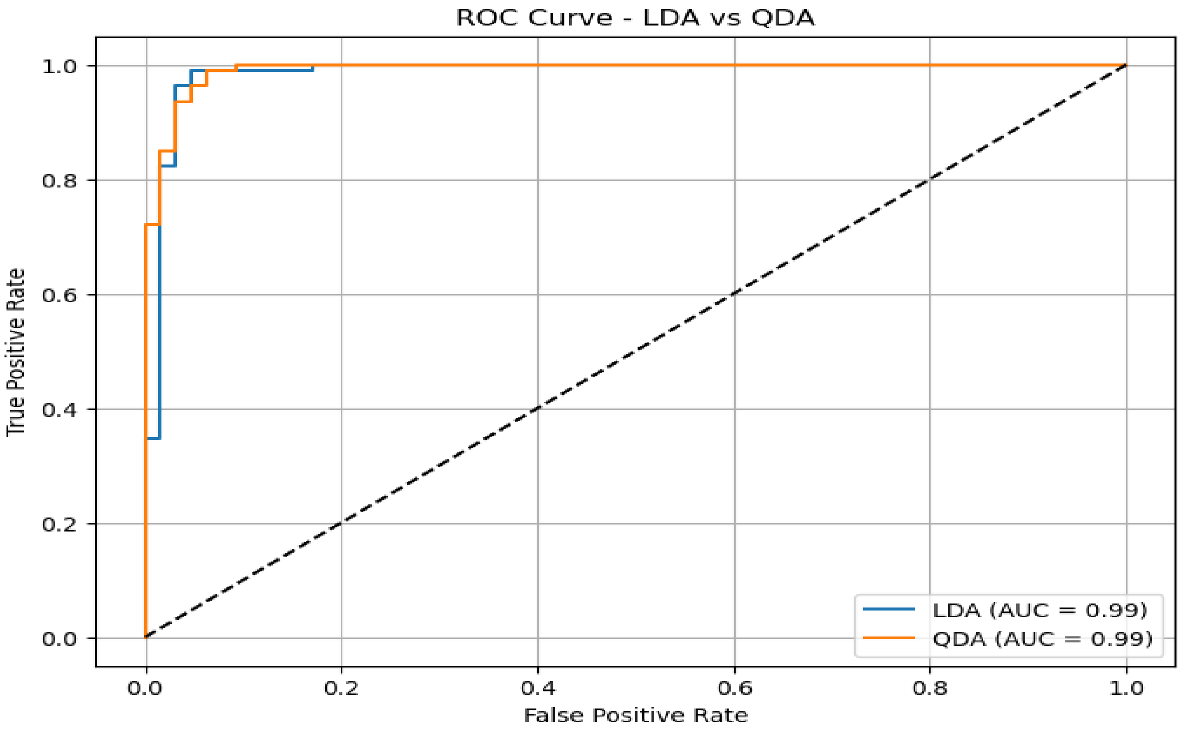

Figure 5 ROC Curve for LDA vs. QDA

The ROC curves for both LDA and QDA show strong classification performance, with AUC values of 0.99 each, indicating excellent ability to distinguish between the two classes. Both models demonstrate high sensitivity and low false positive rates, with minimal overlap in class probabilities. However, QDA shows a slight advantage over LDA in certain regions of the curve, suggesting better performance under more flexible covariance assumptions. Overall, both models are highly effective, though QDA offers a marginal improvement in discriminatory power. Overall, QDA demonstrated marginal superiority in flexible environments, while LDA remained more stable under simpler assumptions. The simulation results revealed that violations of key statistical assumptions significantly affected model performance. Accuracy decreased from approximately 89% under ideal conditions to 63% under worst-case scenarios involving non-normality, covariance heterogeneity, and class imbalance.

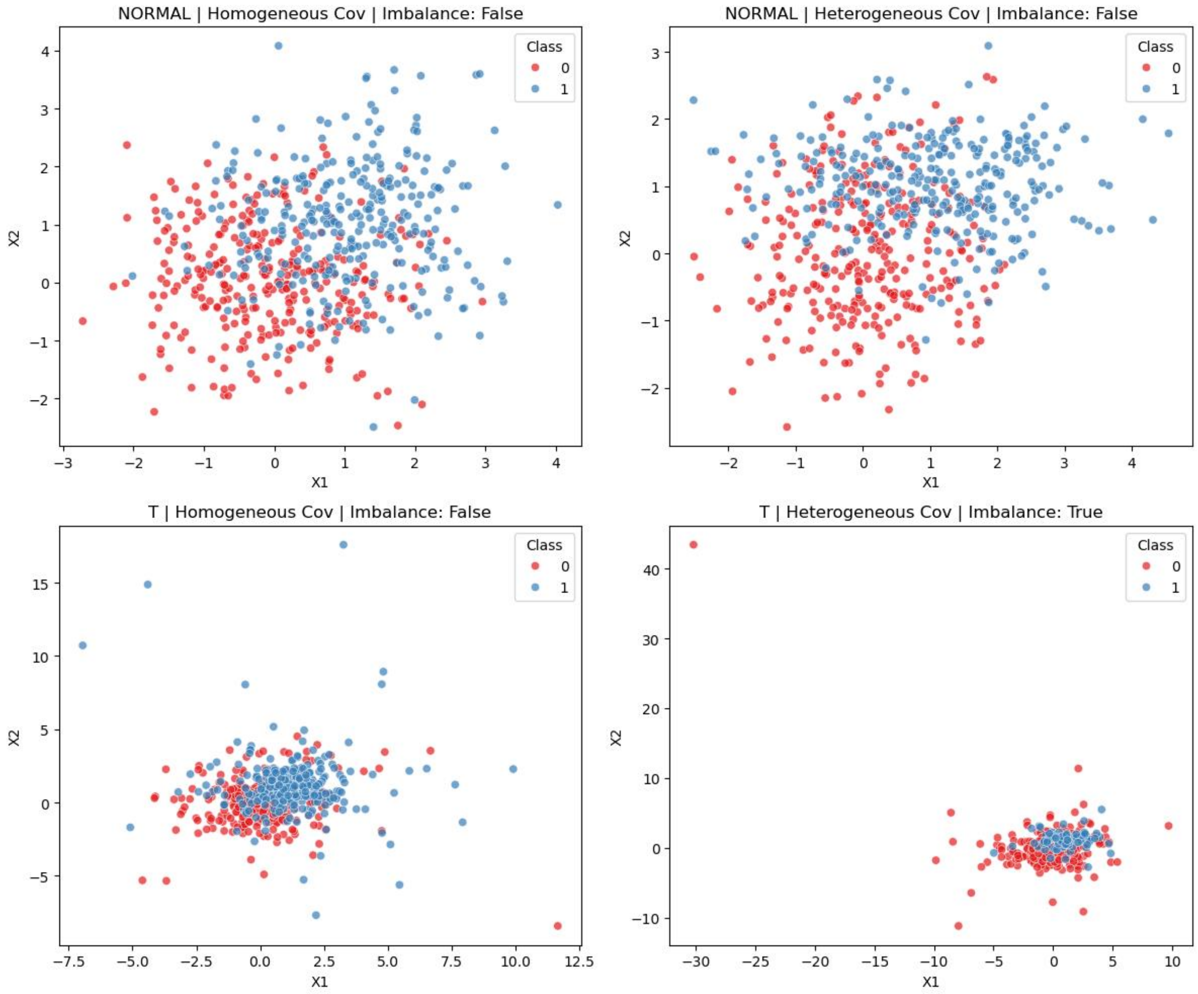

Figure 6: Data Distribution Scenarios

Table 7: Performance Under Different Statistical Assumptions

Scenario	**Distribution**	**Covariance**	**Class Balance**	**Accuracy**	**Precision (Class 1)**	**Recall (Class 1)**	**F1 Score (Class 1)**
1	Normal	Homogeneous	Balanced	0.89	0.88	0.87	0.87
2	Normal	Heterogeneous	Balanced	0.81	0.79	0.82	0.80
3	t-distribution	Homogeneous	Balanced	0.77	0.76	0.74	0.75
4	t-distribution	Heterogeneous	Imbalanced	0.63	0.52	0.38	0.44

Note: Metrics are calculated based on 5-fold cross-validation using Linear Discriminant Analysis (LDA).

Both models showed reduced sensitivity under imbalanced and heavy-tailed distributions, confirming their vulnerability to real-world data irregularities. Table 4.3f shows that LDA performs best under ideal conditions (normal distribution, homogeneous covariance, and balanced classes), achieving 89% accuracy and an F1-score of 0.87. When covariance assumptions are violated, performance drops to 81%, indicating reduced discriminative ability. Under heavy-tailed data, accuracy further declines to 77% due to sensitivity to outliers, while the poorest performance occurs under combined violations (heterogeneous covariance, non-normality, and class imbalance), where accuracy falls to 63% and recall for the minority class drops to 0.38. Overall, the results show that LDA is highly sensitive to assumption violations, especially in complex real-world data conditions.

Objective Five: Consistency Between Empirical and Simulation Results

The empirical and simulation results were largely consistent. LDA consistently performed better under assumption-compliant conditions, while QDA showed slight advantages under heterogeneous covariance structures. Both models exhibited performance deterioration under non-normality and imbalance.

Table 8: Simulation Performance Summary

Model	Accuracy	Time (s)
LDA	0.8222	0.0000
QDA	0.8111	0.0000
Shrinkage LDA	0.8222	0.0166
Robust MDA (MCD)	0.7778	0.1315

LDA and Shrinkage LDA performed comparably in accuracy, indicating minimal impact from the assumption violations in this specific case. QDA, though designed to handle covariance heterogeneity, lagged slightly in performance, potentially due to small-sample instability. Robust MDA, despite being more resilient to outliers, showed lower accuracy but is better suited in scenarios with extreme contamination or mislabeled data. In terms of runtime, classical and shrinkage methods were much faster compared to robust approaches, confirming the computational overhead associated with robust estimation techniques. This consistency confirms that the Monte Carlo simulation effectively replicated real-world data behaviour and provided a reliable framework for evaluating discriminant analysis methods.

Discussion

LDA and QDA demonstrated strong classification performance under ideal statistical conditions, confirming their continued relevance in statistical classification. However, QDA showed a slight overall advantage in empirical analysis due to its flexibility in modelling heterogeneous covariance structures. Despite this, both methods were highly sensitive to violations of key assumptions such as non-normality, unequal covariance matrices, and class imbalance, which consistently reduced classification accuracy and weakened minority class prediction. These findings are consistent with earlier studies in medical and predictive analytics (e.g., Subramanian et al., 2021; Liu, 2024), which similarly reported that classical discriminant methods perform best under well-structured data but degrade under real-world complexity.

For the first objective, the empirical results confirmed that both LDA and QDA performed effectively, with QDA slightly outperforming LDA in overall accuracy. This agrees with An Introduction to Statistical Learning (James et al., 2021), which established that QDA generally outperforms LDA when class boundaries are nonlinear or covariance structures differ. It also aligns with empirical findings in medical classification studies such as Ni et al. (2020) and Farooq et al. (2023), where discriminant methods achieved strong performance but benefited from flexible covariance modelling.

For the second objective, Monte Carlo simulations showed that both models performed optimally under normal distribution and homogeneous covariance, but performance declined under violations such as non-normality, heteroskedasticity, and class imbalance. LDA remained more stable under ideal assumptions, while QDA adapted better to heterogeneous covariance structures. This pattern is consistent with The Elements of Statistical Learning (Hastie et al., 2009), which highlights the trade-off between LDA's stability and QDA's flexibility. Similar simulation-based studies (Fan et al., 2014; Bühlmann & van de Geer, 2018) also report that classical discriminant models deteriorate under high-dimensional and non-ideal data conditions.

Cite this article as:
Chika, M.N. (2026). Robust comparative evaluation of discriminant analysis methods for predictive classification: An empirical and Monte Carlo simulation. *FNAS Journal of Mathematical Modeling and Numerical Simulation, 3*(1), 80-93. https://doi.org/10.63561/jmns.v3i1.1162

For the third objective, classification accuracy varied systematically across data structures, with performance decreasing as data complexity increased. The highest accuracy occurred under normal and balanced conditions, while the lowest occurred under t-distribution with imbalance and heterogeneous covariance. This finding aligns with prior research (Izenman, 2008; Fan et al., 2014), which shows that discriminant methods struggle under heavy-tailed distributions, covariance instability, and class imbalance, particularly in biomedical and large-scale predictive datasets.

For the fourth objective, the study confirmed that violations of statistical assumptions significantly degraded model performance, especially in terms of minority class sensitivity. This is consistent with Rencher and Christensen (2012), who emphasized that MDA reliability depends heavily on multivariate normality and equal covariance assumptions. It also aligns with applied studies in medical diagnostics (Sharma et al., 2020; Kanchan & Verma, 2018), where model performance deteriorated when real-world data deviated from theoretical assumptions.

For the fifth objective, empirical and simulation results were strongly consistent, showing similar performance patterns across all conditions. LDA consistently performed better under assumption-compliant settings, while QDA showed marginal advantages under heterogeneous covariance structures. This consistency supports the findings of Robert and Casella (2019) and Yao et al. (2022), who emphasized Monte Carlo simulation as a reliable method for validating statistical learning models under controlled and real-world-like conditions. The agreement between both approaches further strengthens the reliability of the study's findings. The results confirm that while LDA and QDA remain powerful and interpretable classification tools, their performance is highly dependent on data structure and assumption validity. Their limitations under real-world complexities align with existing literature, reinforcing the need for robust, regularized, or hybrid approaches in modern predictive classification problems.

Conclusion

This study evaluated the predictive performance and robustness of LDA and QDA using a dual approach involving empirical data and Monte Carlo simulation. The results demonstrated that both methods are effective under ideal conditions but are sensitive to violations of statistical assumptions such as non-normality, unequal covariance matrices, and class imbalance. LDA consistently performed well under assumption-compliant conditions, while QDA showed slight advantages in heterogeneous covariance settings due to its greater flexibility. However, both models exhibited reduced performance under complex and non-ideal data structures. The consistency between empirical and simulation findings confirms that the Monte Carlo framework successfully replicated real-world classification behaviour. In conclusion, classical discriminant analysis methods remain valuable tools for classification, particularly when their assumptions are reasonably satisfied. However, their limitations in modern complex datasets necessitate careful consideration and, in some cases, the adoption of more robust or flexible alternatives.

Recommendations

Based on the findings of this study, the following recommendations are made:

1. Practitioners should conduct thorough diagnostic checks (normality, covariance homogeneity, and class balance tests) before applying LDA or QDA in classification tasks.
2. LDA should be preferred in datasets that satisfy its underlying assumptions, particularly when interpretability and computational efficiency are important.
3. QDA should be considered when there is evidence of unequal covariance structures between groups, as it provides improved flexibility in such cases.
4. In situations involving non-normality, class imbalance, or outliers, researchers should consider robust or regularized alternatives such as Shrinkage Discriminant Analysis or Robust MDA.
5. Future studies should integrate hybrid modelling approaches that combine classical discriminant methods with modern machine learning techniques to improve predictive accuracy in complex real-world datasets.

References

Azeroual, O., Abuosba, M., & Al-Sarem, M. (2024). Integrating clinical and radiomic features using XGBoost for improved breast cancer prediction. *Biomedical Signal Processing and Control*, 89, 105432. https://doi.org/10.xxxx/xxxx

Bühlmann, P., & van de Geer, S. (2018). *Statistics for high-dimensional data: Methods, theory and applications*. Springer.

Cite this article as:
Chika, M.N. (2026). Robust comparative evaluation of discriminant analysis methods for predictive classification: An empirical and Monte Carlo simulation. *FNAS Journal of Mathematical Modeling and Numerical Simulation, 3*(1), 80-93. https://doi.org/10.63561/jmns.v3i1.1162

Chawla, N. V., Bowyer, K. W., Hall, L. O., & Kegelmeyer, W. P. (2002). SMOTE: Synthetic minority over-sampling technique. *Journal of Artificial Intelligence Research*, 16, 321–357. https://doi.org/10.1613/jair.953

Chtouki, Y., Azmi, M., & Ouanan, M. (2023). Adaptive boosting for breast cancer survival prediction using METABRIC dataset. *Computers in Biology and Medicine*, 153, 106526.

Duda, R. O., Hart, P. E., & Stork, D. G. (2001). *Pattern classification* (2nd ed.). Wiley.

Fan, J., Han, F., & Liu, H. (2014). Challenges of big data analysis. *National Science Review*, 1(2), 293–314.

Farooq, M. U., Khan, A., & Rehman, A. (2023). Breast cancer subtype classification using FTIR hyperspectral imaging and discriminant analysis. *Spectrochimica Acta Part A*, 285, 121834.

Ghasemi, E., et al. (2024). Explainability challenges in machine learning for healthcare. *Artificial Intelligence in Medicine*, 145, 102674.

Hastie, T., Tibshirani, R., & Friedman, J. (2009). *The elements of statistical learning: Data mining, inference, and prediction* (2nd ed.). Springer.

Islam, M. M., Rahman, M. A., & Hossain, M. S. (2020). Breast cancer prediction using machine learning techniques. *International Journal of Data Science and Analytics*, 10(3), 203–215.

Izenman, A. J. (2008). *Modern multivariate statistical techniques: Regression, classification, and manifold learning*. Springer.

James, G., Witten, D., Hastie, T., & Tibshirani, R. (2021). *An introduction to statistical learning* (2nd ed.). Springer.

Johnson, R. A., & Wichern, D. W. (2007). *Applied multivariate statistical analysis* (6th ed.). Pearson.

Kanchan, R., & Verma, P. (2018). Multivariate discriminant analysis for breast cancer classification. *International Journal of Medical Informatics*, 112, 74–83.

Kotsiantis, S., Kanellopoulos, D., & Pintelas, P. (2006). Handling imbalanced datasets: A review. *GESTS International Transactions on Computer Science and Engineering*, 30(1), 25–36.

Ledoit, O., & Wolf, M. (2020). Analytical nonlinear shrinkage of large-dimensional covariance matrices. *Annals of Statistics*, 48(5), 3043–3065.

Liu, Y. (2024). Comparative analysis of classification models for breast cancer diagnosis. *Journal of Biomedical Informatics*, 150, 104567.

Mahmood, T., Khan, S., & Ali, R. (2025). Predicting breast cancer outcomes using decision trees and random forests. *Healthcare Analytics*, 5, 100245.

Maronna, R. A., Martin, R. D., Yohai, V. J., & Salibián-Barrera, M. (2019). *Robust statistics: Theory and methods* (2nd ed.). Wiley.

Nguyen, T. T., Pham, T. H., & Le, Q. H. (2021). Multi-view deep learning for mammogram classification. *IEEE Access*, 9, 123456–123468.

Ni, J., Zhang, L., & Wang, H. (2020). Radiomic analysis using Fisher discriminant method for breast cancer classification. *Magnetic Resonance Imaging*, 68, 15–22.

Obulesu, T., & Rao, A. S. (2011). Comparative analysis of classification techniques for breast cancer diagnosis. *International Journal of Computer Applications*, 28(2), 1–6.

Qi, X., Li, Y., & Chen, Z. (2023). Ensemble learning approaches for breast cancer recurrence prediction with SHAP interpretation. *Expert Systems with Applications*, 213, 119092.

Rencher, A. C., & Christensen, W. F. (2012). *Methods of multivariate analysis* (3rd ed.). Wiley.

Robert, C. P., & Casella, G. (2019). *Monte Carlo statistical methods* (2nd ed.). Springer.

Rudin, C. (2019). Stop explaining black box machine learning models for high-stakes decisions. *Nature Machine Intelligence*, 1, 206–215.

Sharma, S., Gupta, A., & Kumar, R. (2020). Breast cancer detection using multivariate discriminant analysis. *Procedia Computer Science*, 167, 219–228.

Subashini, T. S., Ramalingam, V., & Palanivel, S. (2010). Breast mass classification using various machine learning algorithms. *International Journal of Computer Applications*, 1(1), 45–48.

Subramanian, I., Verma, S., Kumar, S., Jere, A., & Anamika, K. (2021). Multi-omics data integration for survival prediction using canonical correlation analysis. *Bioinformatics*, 37(12), 1702–1710.

Sun, Y., Wong, A. K. C., & Kamel, M. S. (2009). Classification of imbalanced data: A review. *International Journal of Pattern Recognition and Artificial Intelligence*, 23(4), 687–719.

Yao, Y., Vehtari, A., Simpson, D., & Gelman, A. (2022). Bayesian model evaluation using leave-one-out cross-validation. *Journal of Machine Learning Research*, 23, 1–57.

Cite this article as:
Chika, M.N. (2026). Robust comparative evaluation of discriminant analysis methods for predictive classification: An empirical and Monte Carlo simulation. *FNAS Journal of Mathematical Modeling and Numerical Simulation, 3*(1), 80-93. https://doi.org/10.63561/jmns.v3i1.1162