Classification Report: Diabetic Prediction Models

Overview

This report presents a comparative analysis of several machine learning models used to predict diabetes status. The models include both baseline versions and versions that incorporate class weighting to address class imbalance in the dataset. Each model's performance is evaluated using precision, recall, and F1-score for both non-diabetic (`class 0`) and diabetic (`class 1`) categories.

Key Findings

1. Class Imbalance Effects

Across all models, we observe better performance on the non-diabetic class (0) than on the diabetic class (1). This suggests that the dataset is imbalanced—likely with more non-diabetic examples—which skews model predictions toward the majority class.

2. Impact of Class Weighting

Introducing class weights generally improved recall for the diabetic class, which is especially important in healthcare contexts where failing to identify diabetic patients can have serious consequences. However, this often came at the cost of slightly reduced precision or F1-score for the non-diabetic class.

Model-by-Model Analysis

Gaussian Naive Bayes (GNB)

- **Baseline GNB** performed very well on non-diabetics (F1: 85.82%) but had a lower recall (59.54%) for diabetics.

- **Weighted GNB** significantly improved recall for diabetics (83.64%) and also produced the **highest F1-score for class 1** at 73.02%.

Conclusion: Weighted GNB offers the best balance for diabetic detection, making it the top performer among all models for this task.

Logistic Regression (LR)

- **Baseline LR** had decent performance across both classes but fell short on diabetic recall (61.82%).

- **Weighted LR** improved diabetic recall to 76.36% and raised the F1-score for diabetics to 70.00%.

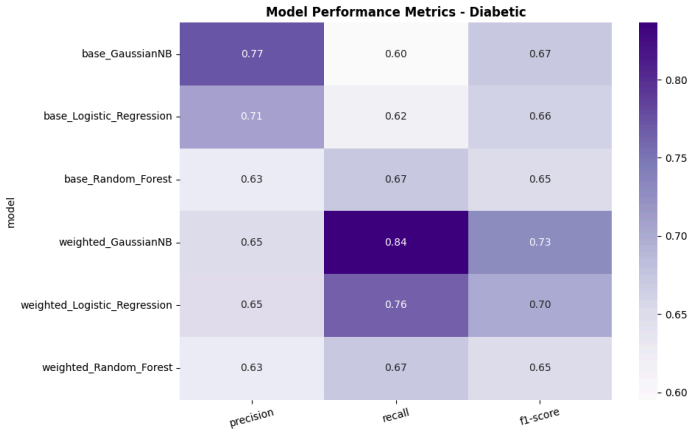
Conclusion: Class weighting enhances LR's ability to detect diabetics, with a modest trade-off in non-diabetic performance.

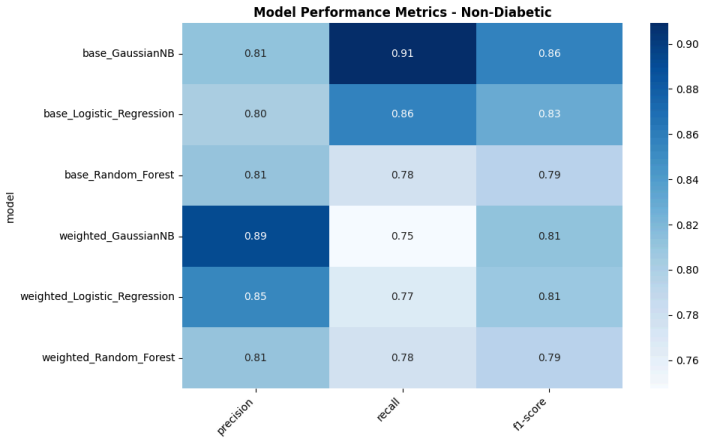
Random Forest (RF)

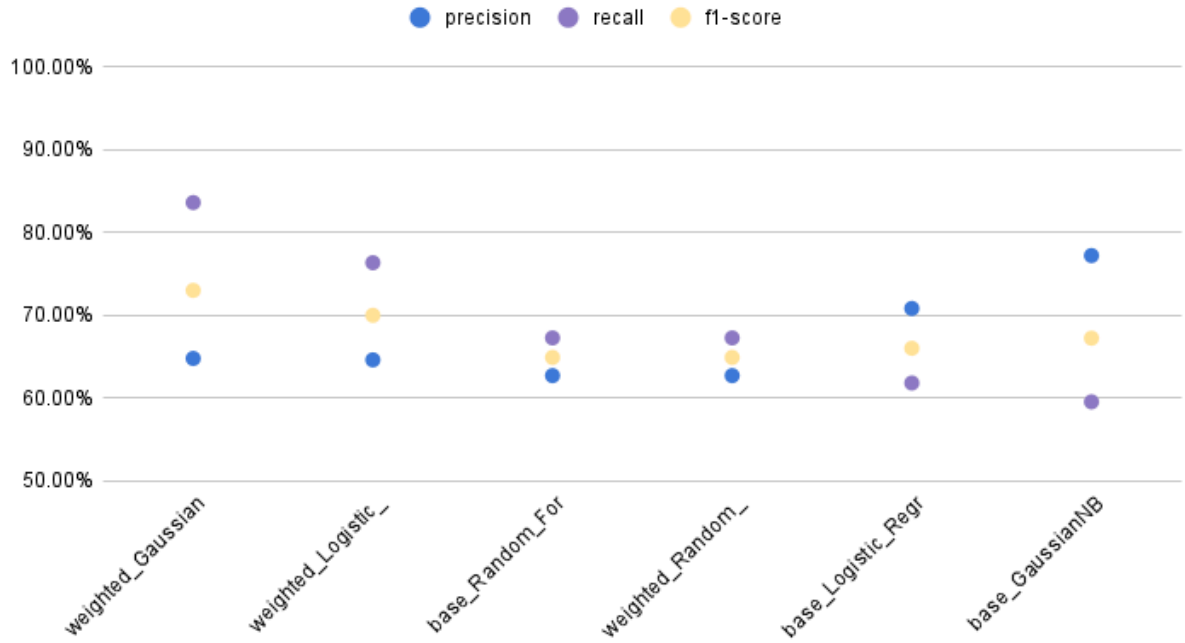
- Both the **baseline and weighted RF** models delivered identical results, showing no improvement with class weighting.

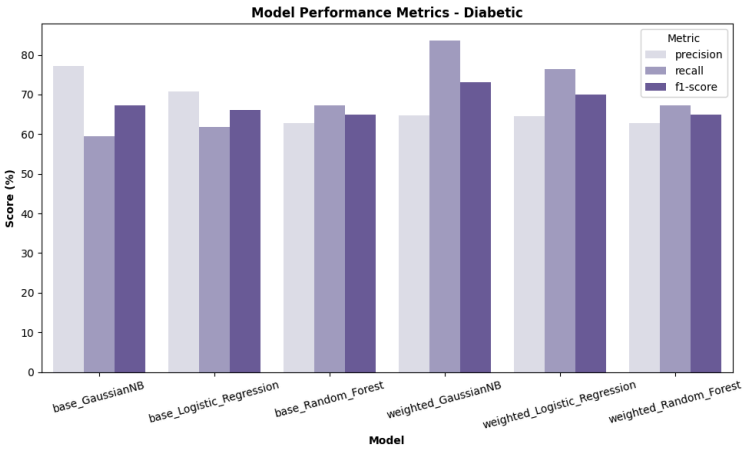
- Diabetic class F1-score remained low (64.91%) with relatively moderate recall and precision.

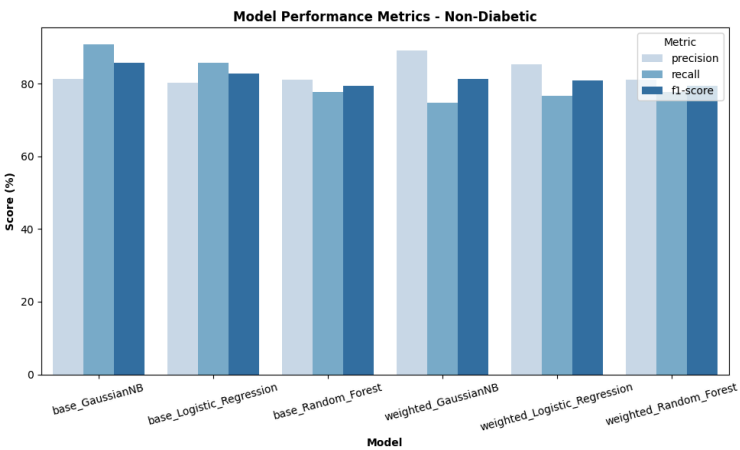
Conclusion: In this setup, Random Forest is the least effective at detecting diabetics, and unresponsive to class weighting—potentially due to default parameters or already internal handling of imbalance.


Model Performance Metrics - Diabetic
model
base_GaussianNB
base_Logistic_Regression
base_Random_Forest
weighted_GaussianNB
weighted_Logistic_Regression
weighted_Random_Forest
0.77
0.60
0.67
0.71
0.62
0.66
0.63
0.67
0.65
0.65
0.84
0.73
0.65
0.76
0.70
0.63
0.67
0.65
precision
recall
f1-score
0.80
0.75
0.70
0.65
0.60


Model Performance Metrics - Non-Diabetic
model
base_GaussianNB
base_Logistic_Regression
base_Random_Forest
weighted_GaussianNB
weighted_Logistic_Regression
weighted_Random_Forest
0.81
0.91
0.86
0.80
0.86
0.83
0.81
0.78
0.79
0.89
0.75
0.81
0.85
0.77
0.81
0.81
0.78
0.79
precision
recall
f1-score
0.90
0.88
0.86
0.84
0.82
0.80
0.78
0.76


precision
recall
f1-score
100.00%
90.00%
80.00%
70.00%
60.00%
50.00%
weighted_Gaussian
weighted_Logistic_
base_Random_For
weighted_Random_
base_Logistic_Regr
base_GaussianNB


Model Performance Metrics - Diabetic
Metric
precision
recall
f1-score
Score (%)
0
10
20
30
40
50
60
70
80
base_GaussianNB
base_Logistic_Regression
base_Random_Forest
weighted_GaussianNB
weighted_Logistic_Regression
weighted_Random_Forest
Model


Model Performance Metrics - Non-Diabetic
Metric
precision
recall
f1-score
Score (%)
0
20
40
60
80
base_GaussianNB
base_Logistic_Regression
base_Random_Forest
weighted_GaussianNB
weighted_Logistic_Regression
weighted_Random_Forest
Model

Summary of F1-Scores for Diabetic Class (1)

Model	F1-Score (Class1)
Base GaussianNB	67.24%
Weighted GaussianNB	**73.02%**
Base Logistic Regression	66.02%
Weighted Logistic Regression	70.00%
Random Forest (Base/Weighted)	64.91%

Final Recommendation

For tasks where correctly identifying diabetic patients is critical, the **weighted Gaussian Naive Bayes** model is the most effective option, offering the best balance of recall and F1-score for the diabetic class. If slightly more interpretability or regularization is desired, **weighted logistic regression** is a strong alternative.