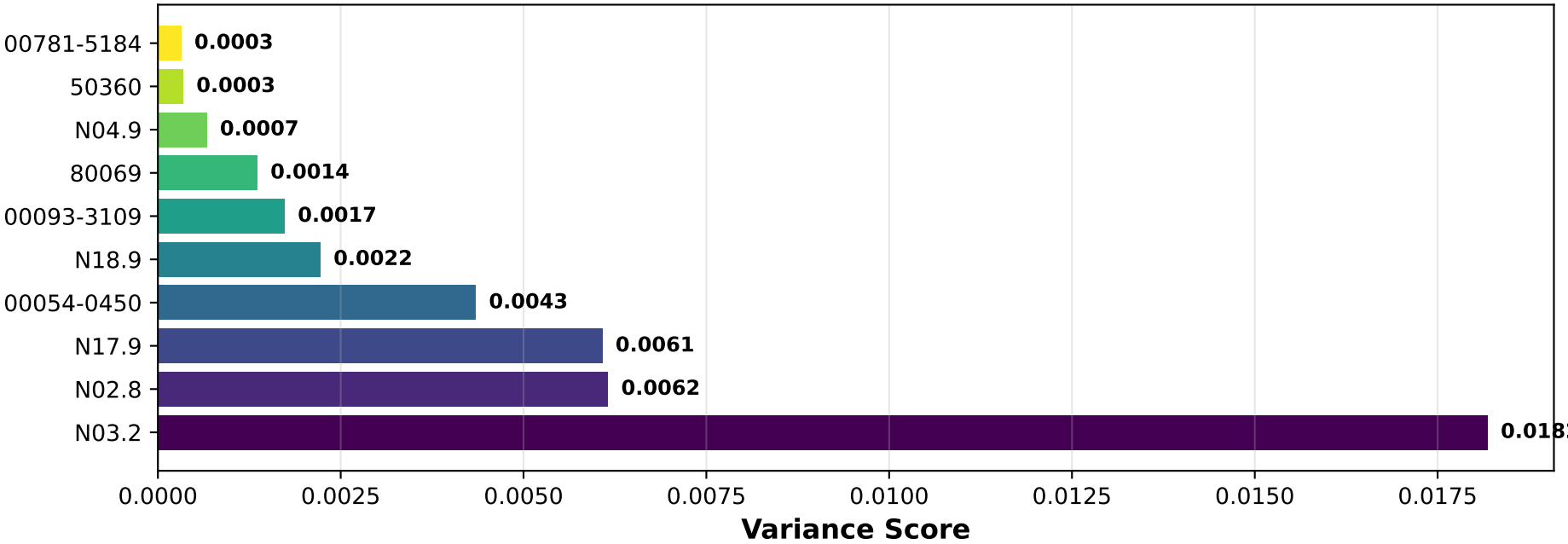
Patient Finder Model Performance Report

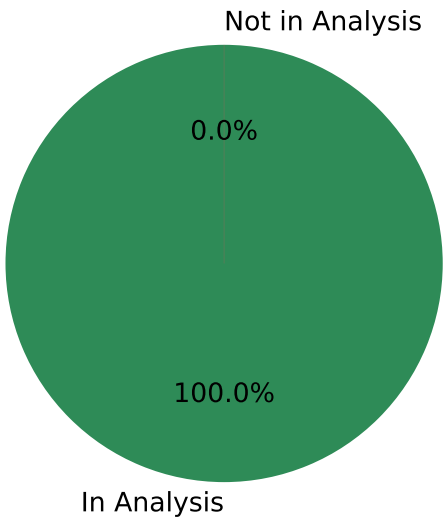
Generated on August 21, 2025

FEATURE IMPORTANCE ANALYSIS

Top Features by Variance



Business-Important Features Overlap Analysis

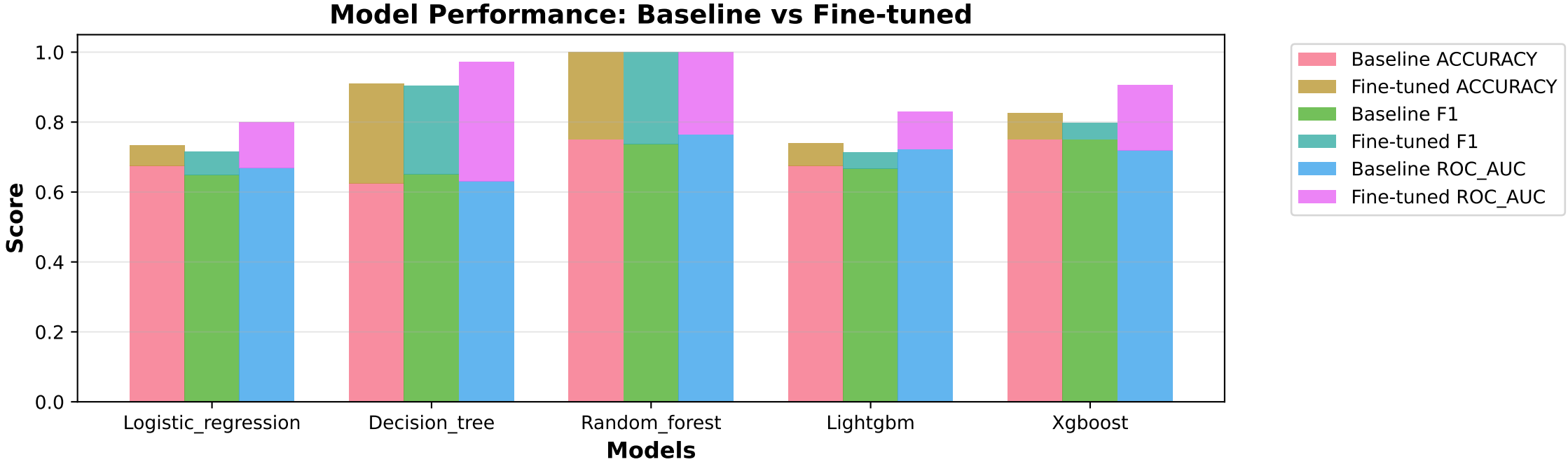


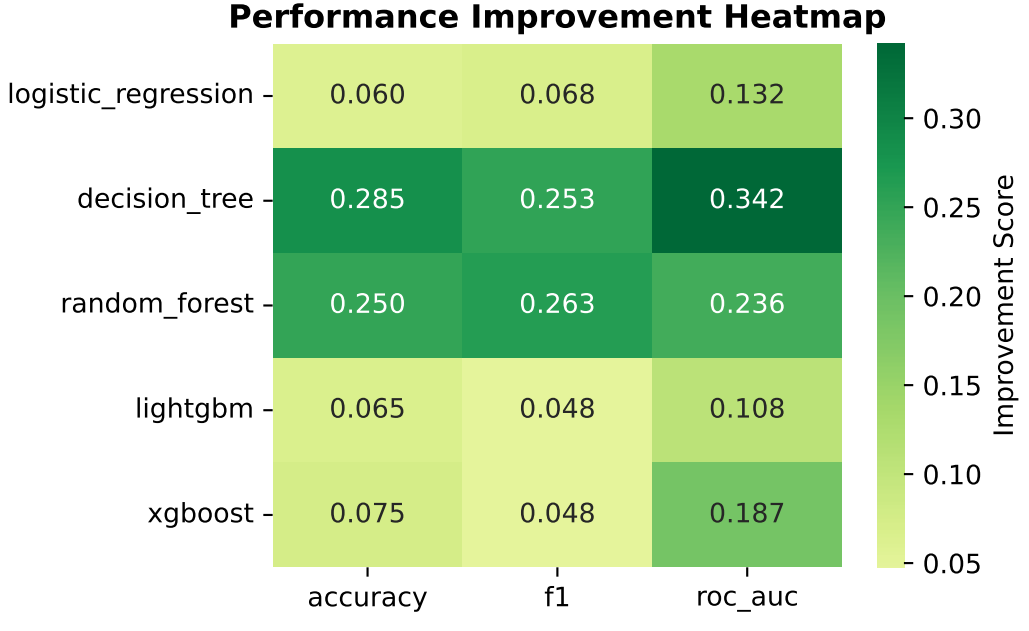
FEATURE STATISTICS

Total Features Analyzed: 14
Business-Important Features: 10
Overlap Count: 10
Overlap Percentage: 100.0%

Top 5 Business Features in Analysis:
1. N03.2
2. N02.8
3. N17.9
4. 00054-0450
5. N18.9

MODEL PERFORMANCE COMPARISON





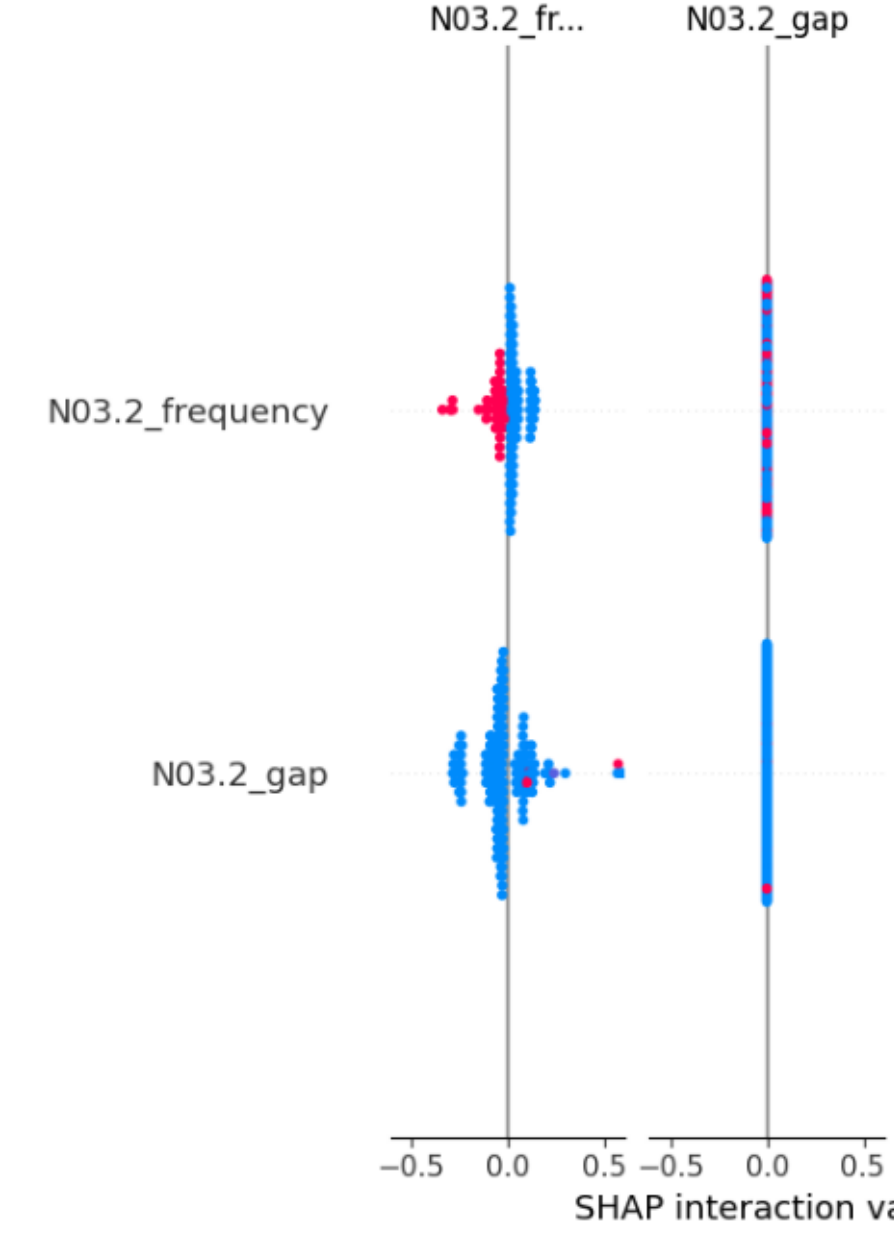
BEST PERFORMING MODEL

Model: RANDOM_FOREST

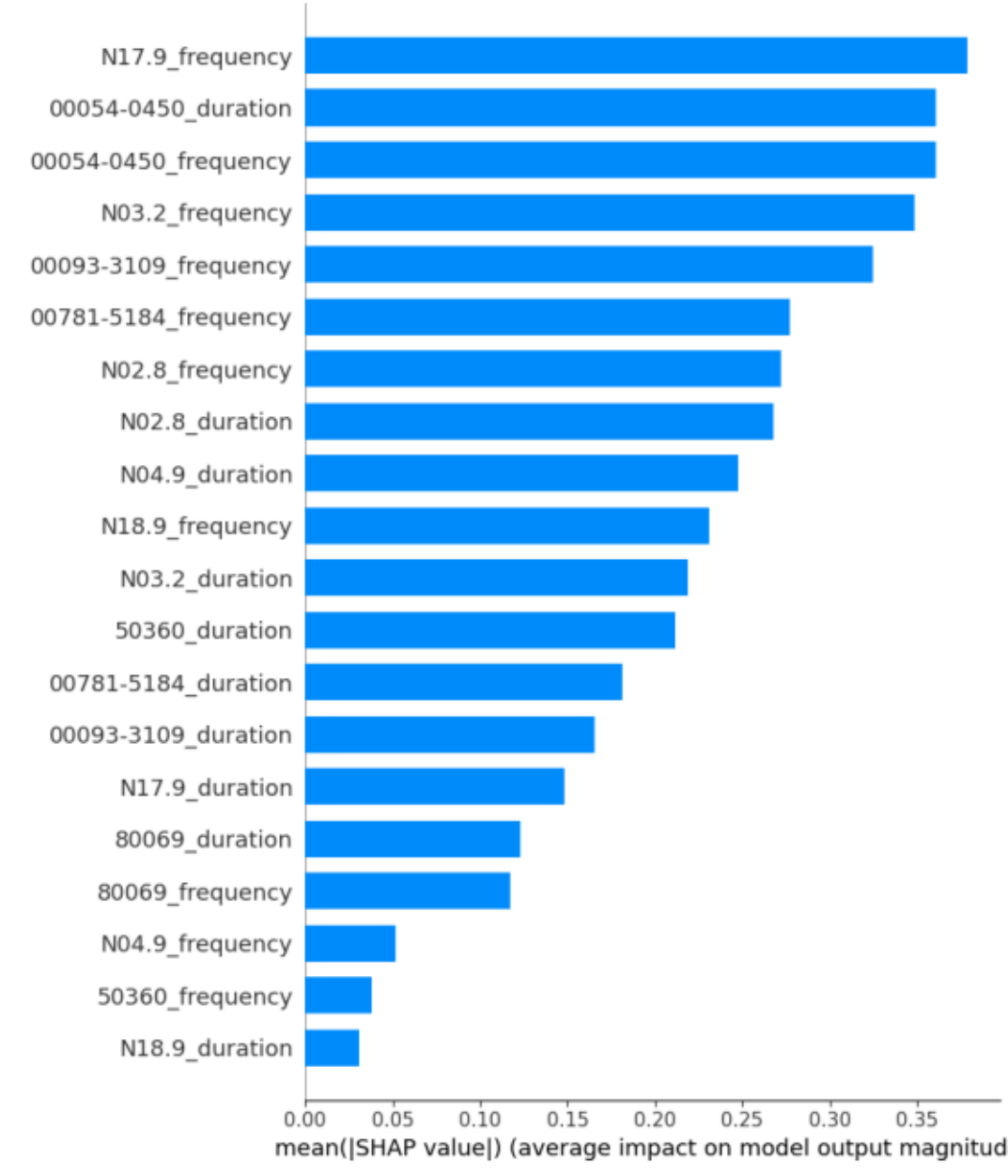
Performance Metrics:
- Accuracy: 1.000
- F1 Score: 1.000
- AUC Score: 1.000

Improvements over Baseline:
- Accuracy: +0.250
- F1 Score: +0.263
- AUC Score: +0.236

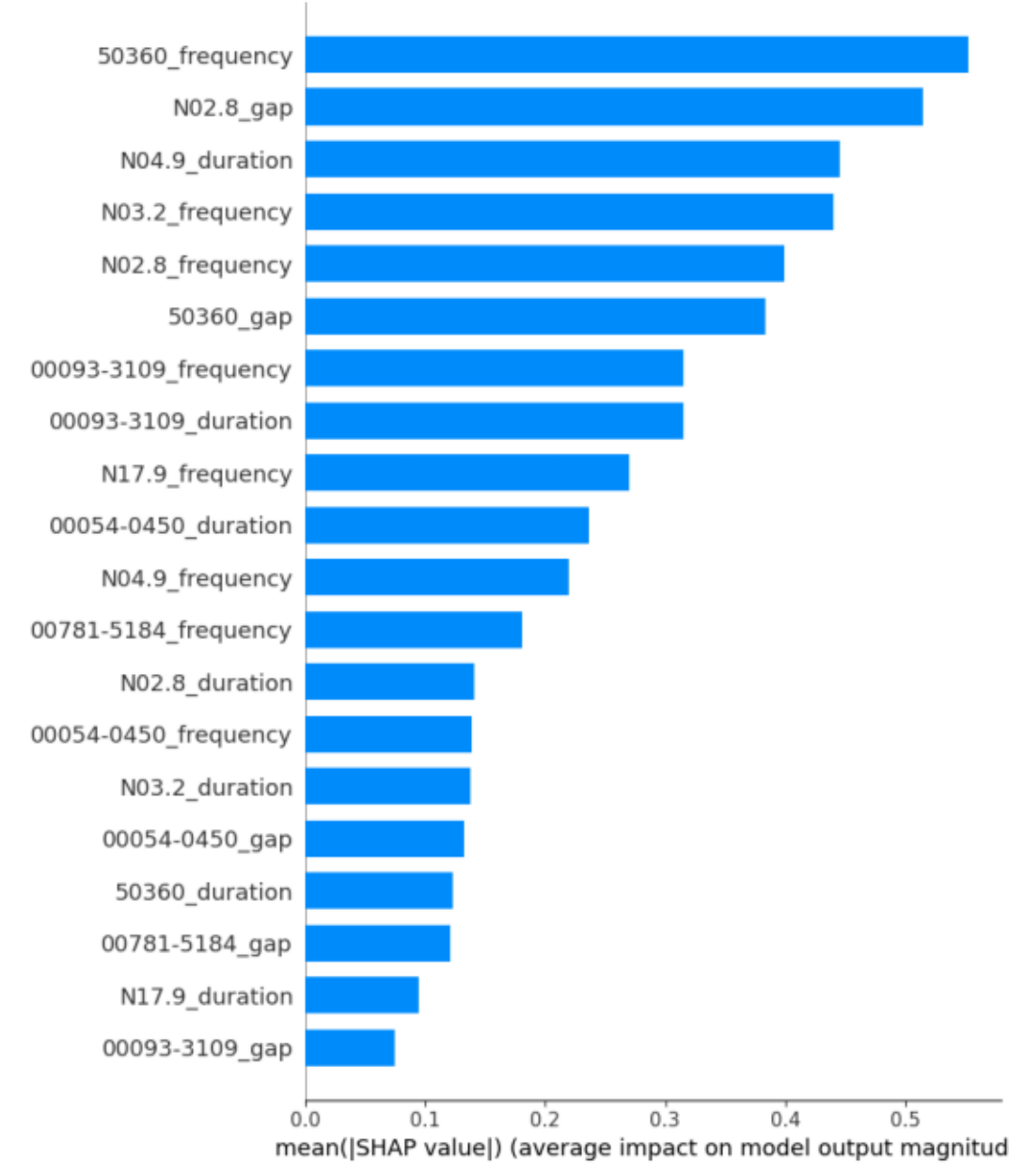
SHAP Explanation: Shap Bar Decision Tree



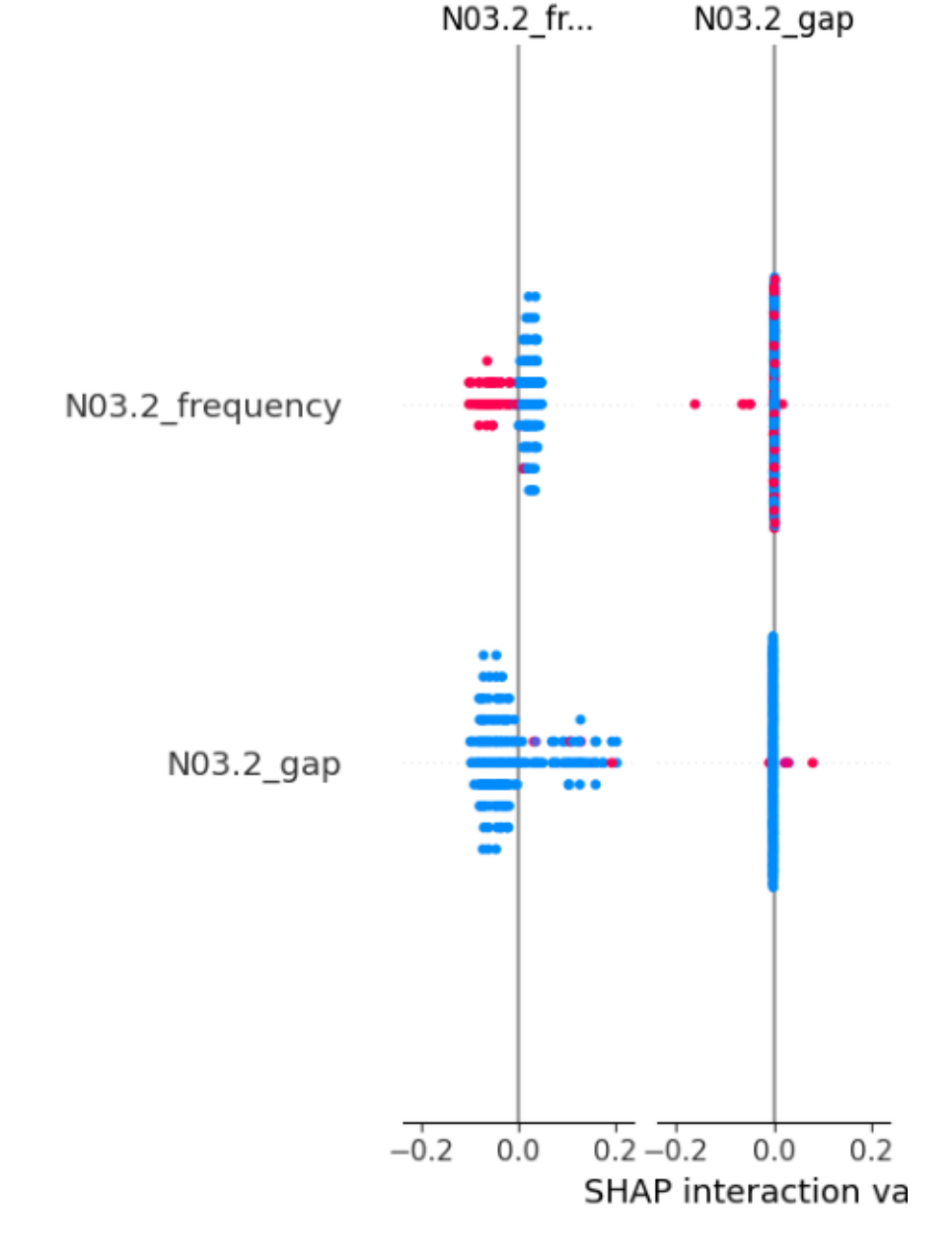
SHAP Explanation: Shap Bar Lightgbm



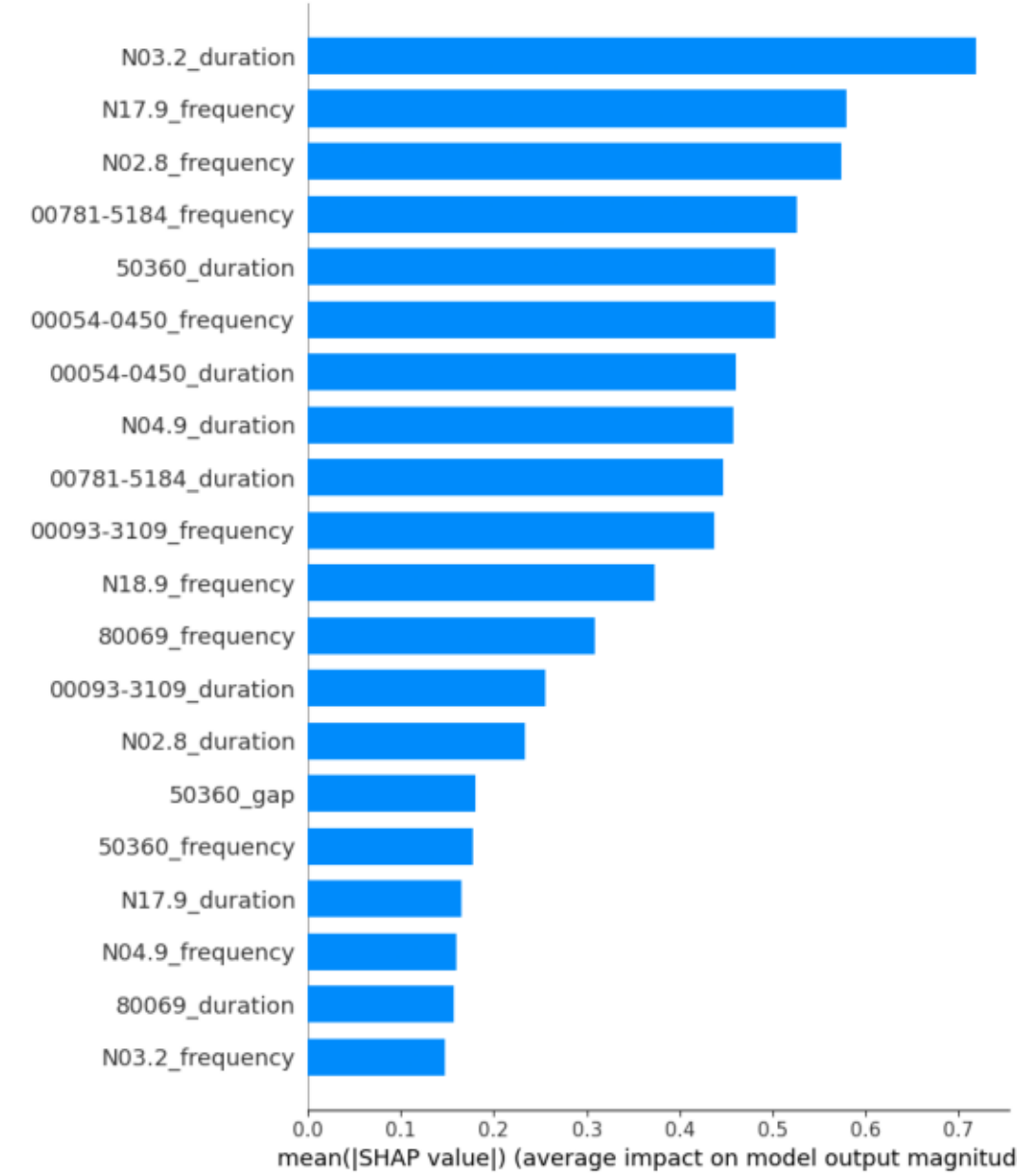
SHAP Explanation: Shap Bar Logistic Regression



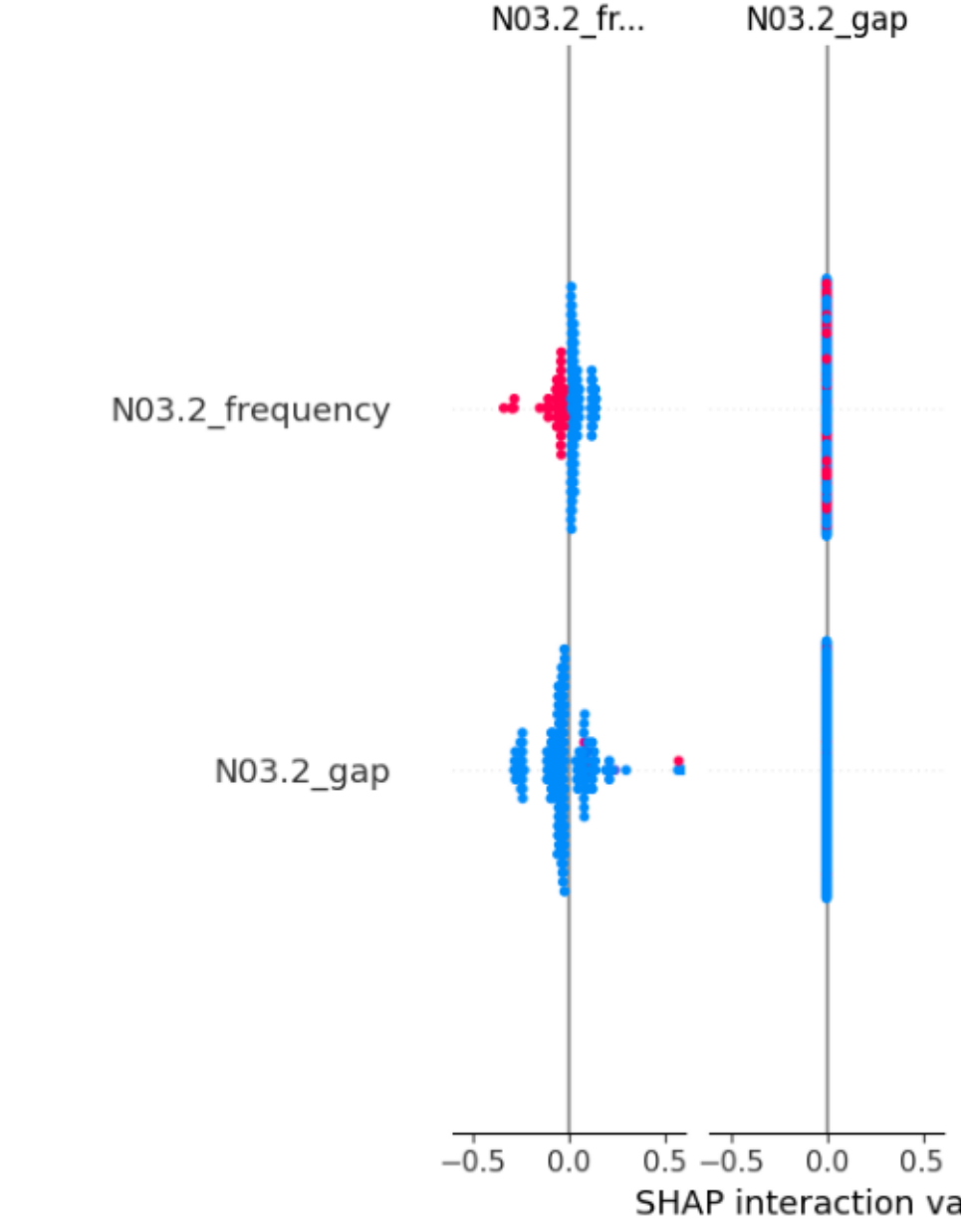
SHAP Explanation: Shap Bar Random Forest



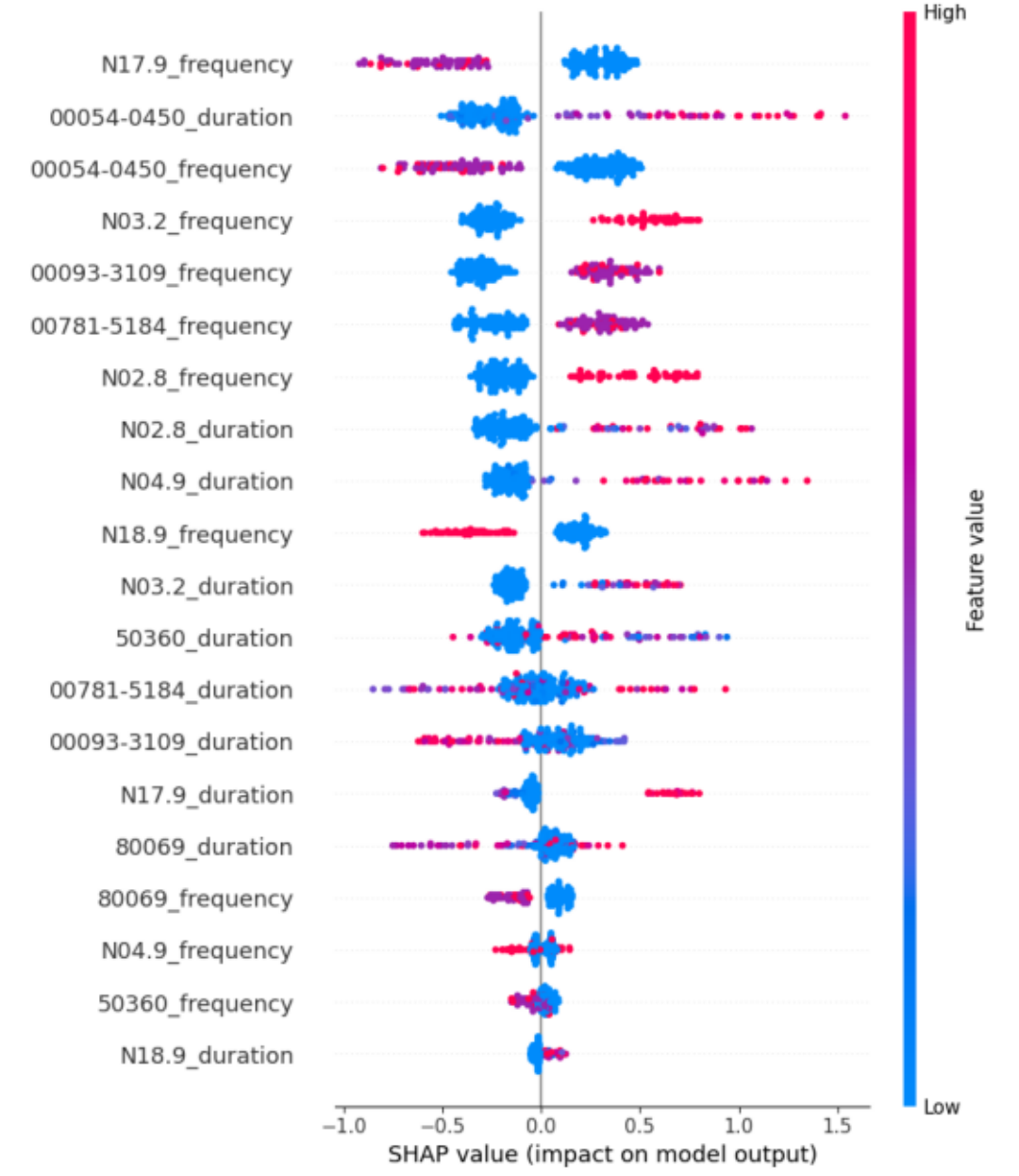
SHAP Explanation: Shap Bar Xgboost



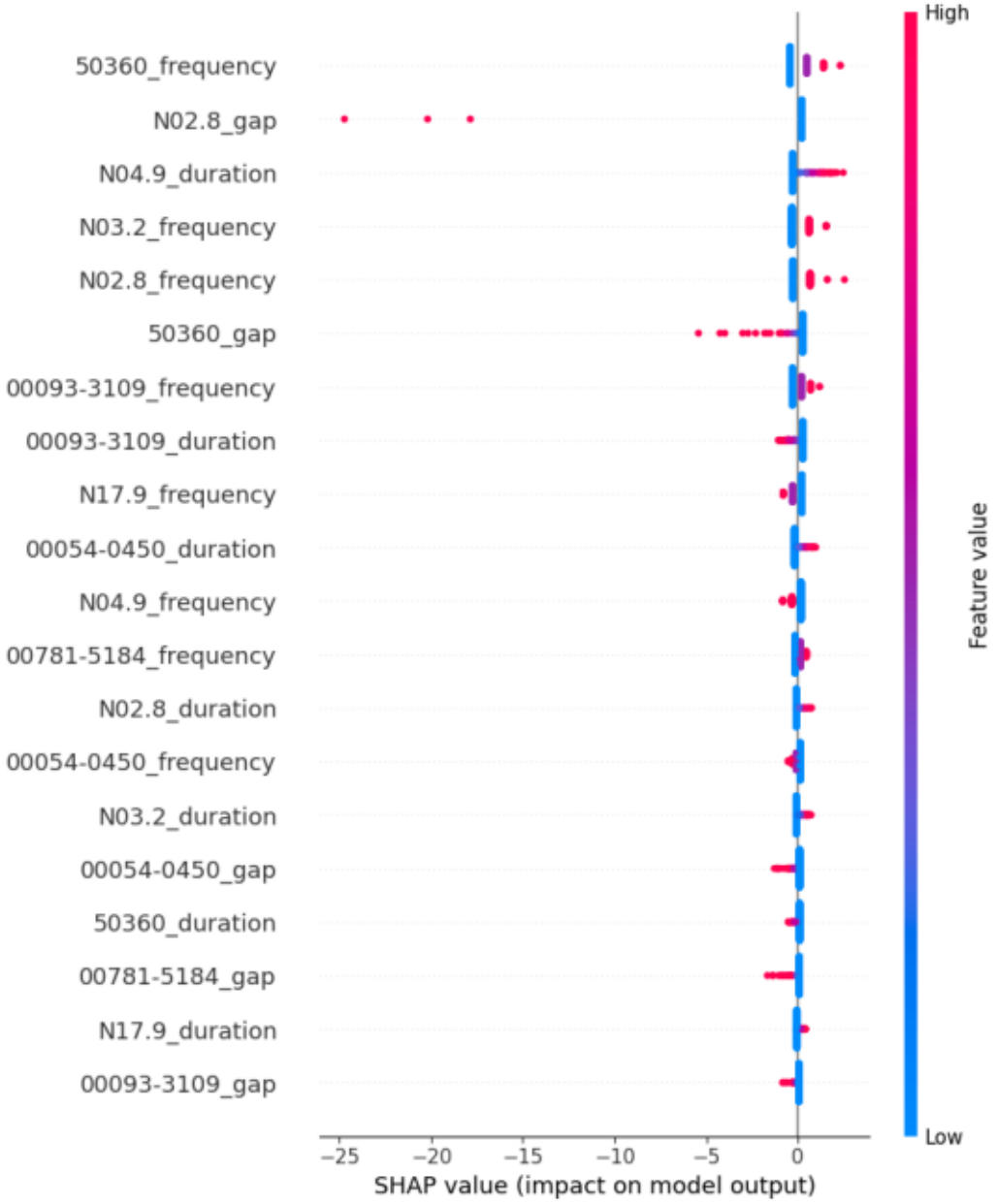
SHAP Explanation: Shap Summary Decision Tree



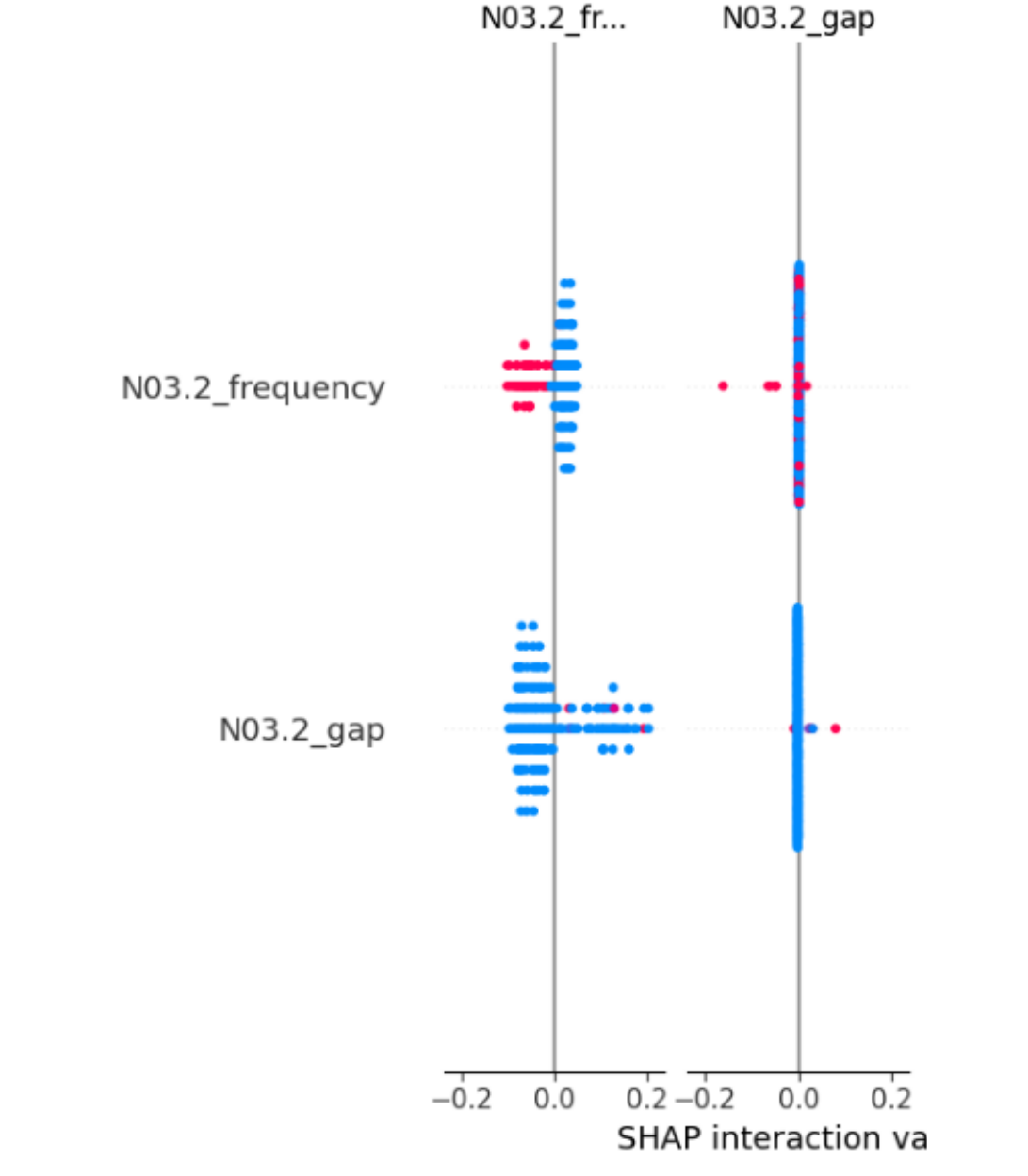
SHAP Explanation: Shap Summary Lightgbm



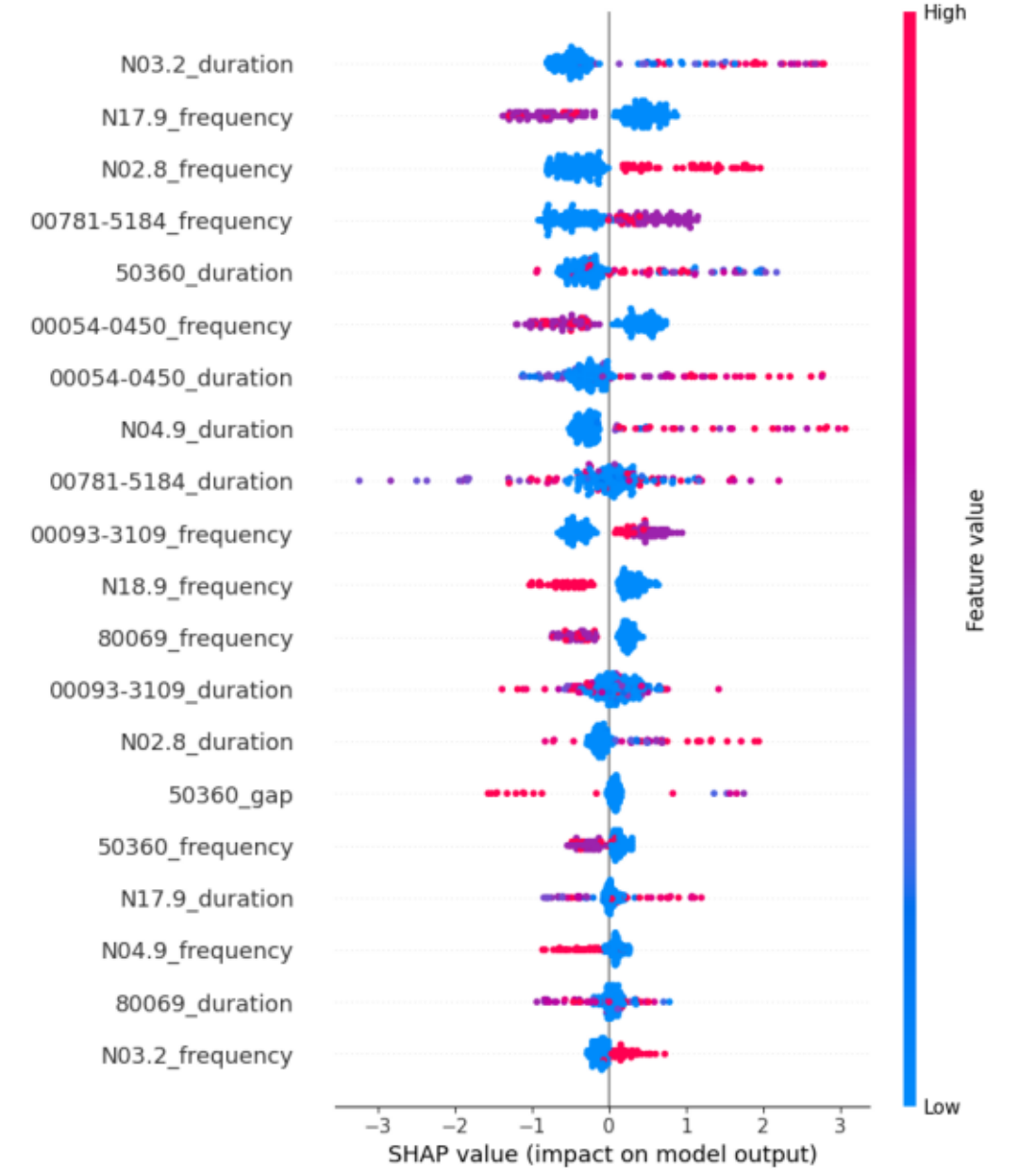
SHAP Explanation: Shap Summary Logistic Regression

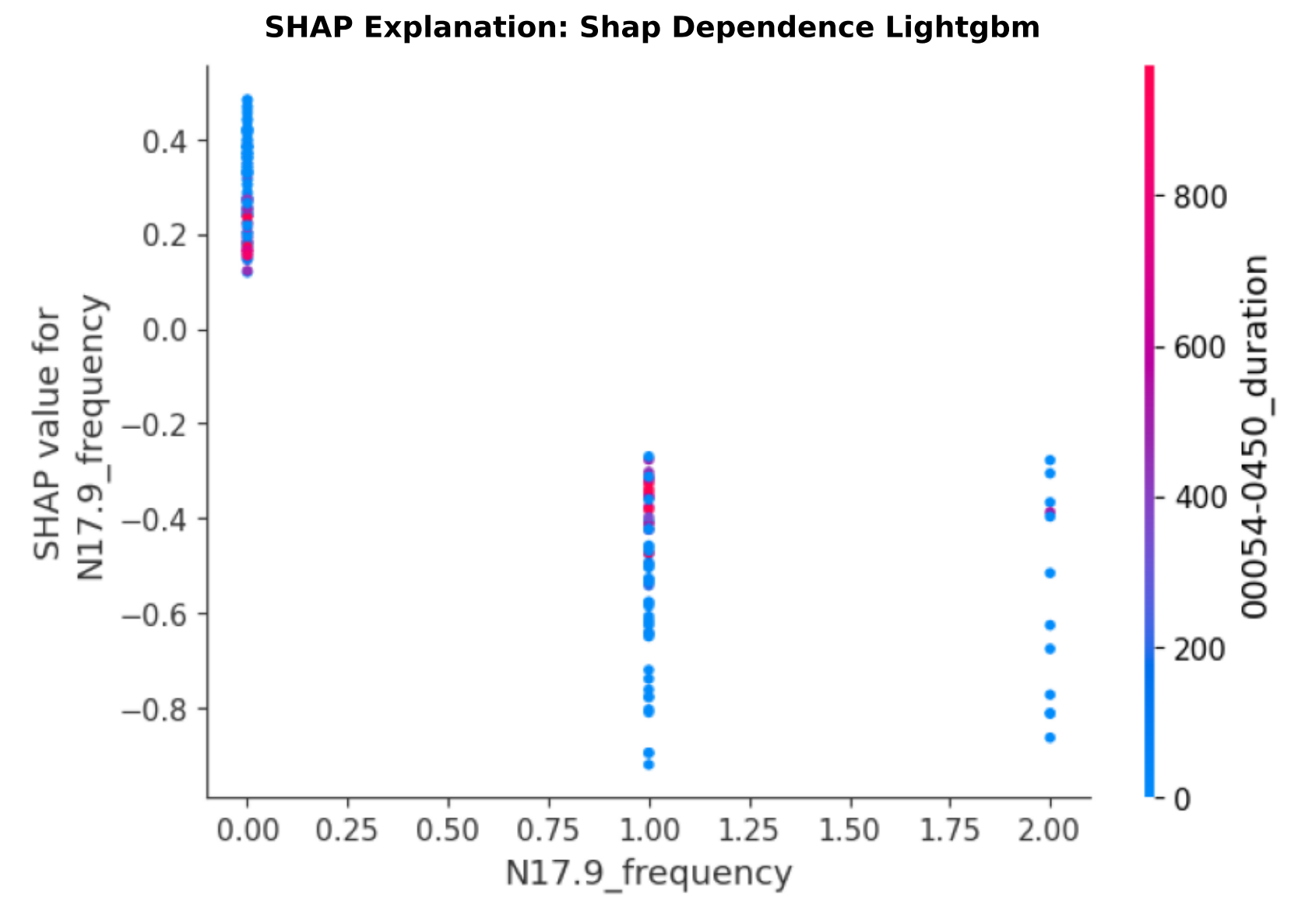


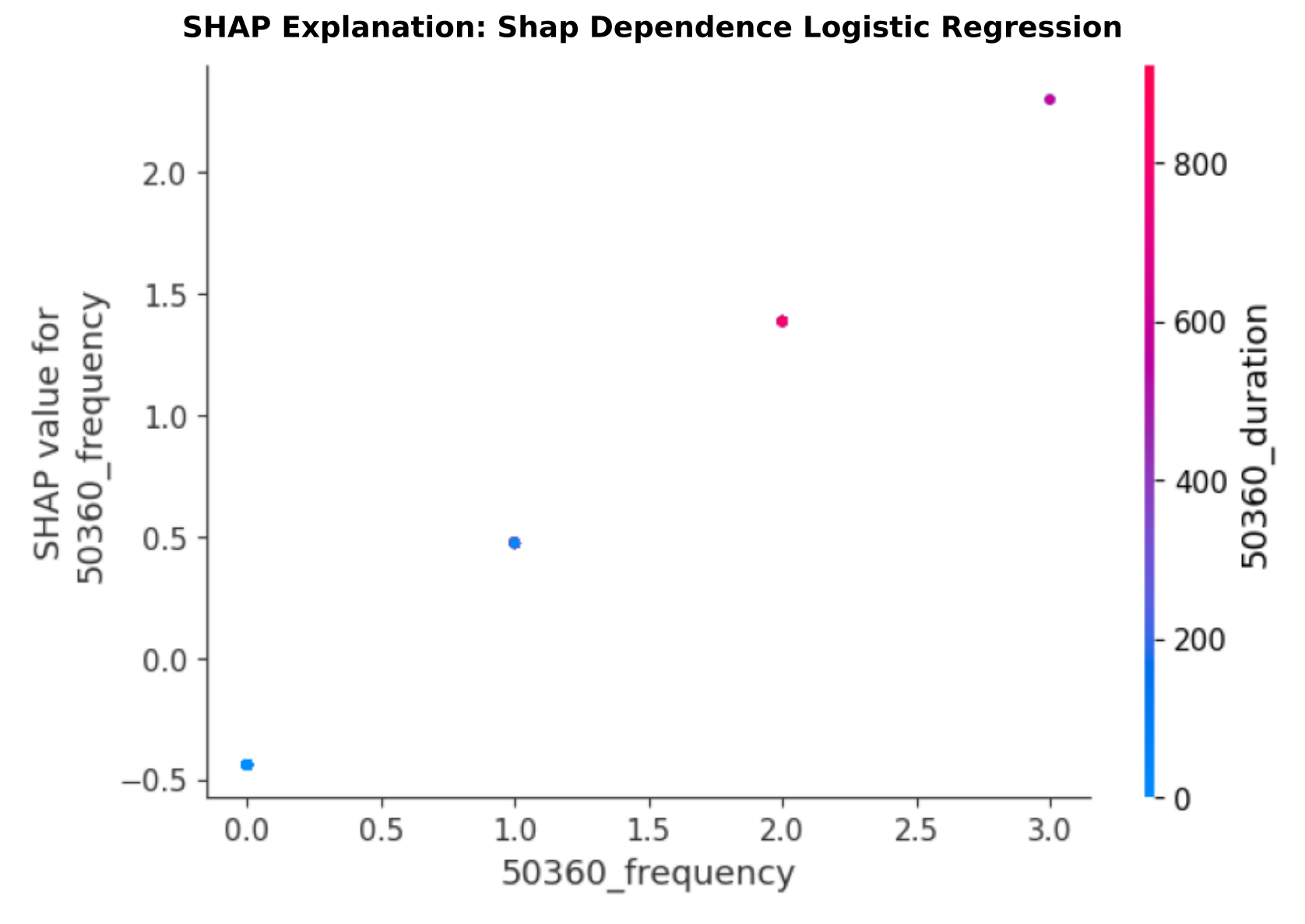
SHAP Explanation: Shap Summary Random Forest

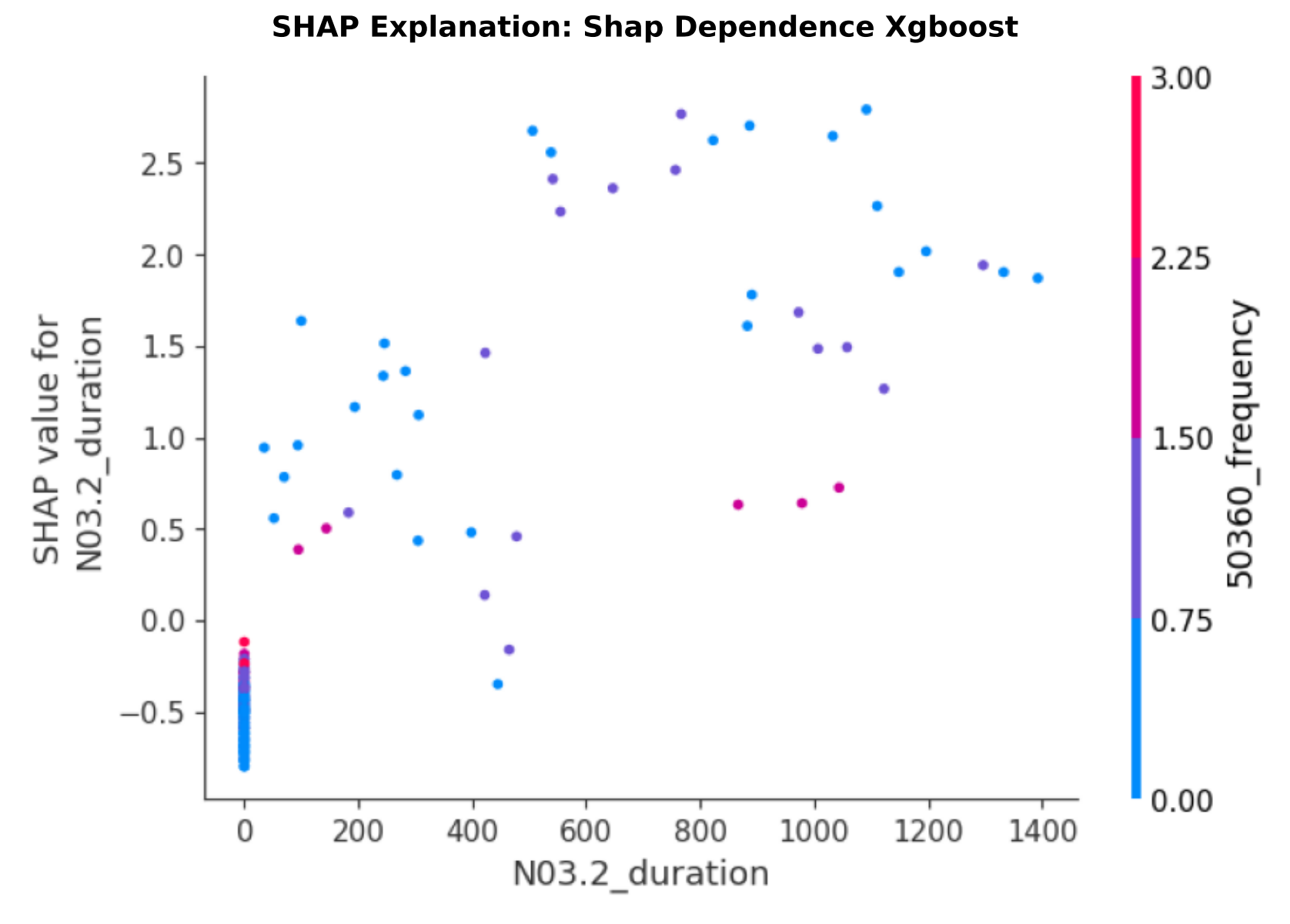


SHAP Explanation: Shap Summary Xgboost




SHAP Explanation: Shap Dependence Lightgbm
SHAP value for N17.9_frequency
0.4
0.2
0.0
−0.2
−0.4
−0.6
−0.8
0.00
0.25
0.50
0.75
1.00
1.25
1.50
1.75
2.00
N17.9_frequency
00054-0450_duration
800
600
400
200
0


SHAP Explanation: Shap Dependence Logistic Regression
SHAP value for
50360_frequency
2.0
1.5
1.0
0.5
0.0
−0.5
0.0
0.5
1.0
1.5
2.0
2.5
3.0
50360_frequency
800
600
400
200
0
50360_duration


SHAP Explanation: Shap Dependence Xgboost
SHAP value for N03.2_duration
N03.2_duration
50360_frequency
2.5
2.0
1.5
1.0
0.5
0.0
−0.5
0
200
400
600
800
1000
1200
1400
3.00
2.25
1.50
0.75
0.00

RECOMMENDATIONS & CONCLUSIONS

- On average, fine-tuning improved accuracy by 0.147.
- Feature 'N03.2' showed the highest variance importance.
- SHAP analysis highlights key drivers of model predictions, supporting interpretability and business decision-making.
- Recommended next steps: validate top features with domain experts and assess fairness across patient subgroups.