#### Classification Performance

The Naive Bayes algorithm achieved a classification accuracy of \*\*54.7%\*\* on the test dataset, demonstrating its ability to differentiate between skill levels of surgeons. The detailed classification metrics are presented below:

- \*\*Expert:\*\*

- Precision: 0.58

- Recall (Sensitivity): 0.26

- F1-Score: 0.35

- Support: 43

- \*\*Intermediate:\*\*

- Precision: 0.53

- Recall (Sensitivity): 0.67

- F1-Score: 0.59

- Support: 108

- \*\*Novice:\*\*

- Precision: 0.56

- Recall (Sensitivity): 0.54

- F1-Score: 0.55

- Support: 94

- \*\*Overall Accuracy:\*\* 0.55

#### Aggregated Metrics

A summary of the key evaluation metrics is as follows:

| Metric | Score |

|----------------|-----------|

| Accuracy | 0.546939 |

| Sensitivity | 0.488345 |

| Specificity | 0.745214 |

| F1-Score | 0.499594 |

| MCC | 0.251210 |

| AUC | 0.650108 |

These results indicate that the classifier shows moderate performance, with higher specificity (74.5%) compared to sensitivity (48.8%). The area under the ROC curve (AUC) value of 0.65 suggests some ability to distinguish between classes but highlights room for improvement in model optimization. Further exploration of features, tasks, and muscle groups may enhance classification performance.

### Methods

#### Hyperparameter Optimization

Hyperparameter tuning was performed to ensure optimal performance of the Naive Bayes classifier. While Naive Bayes inherently has fewer hyperparameters compared to other algorithms, the following configurations were explored:

- \*\*Smoothing Parameter (alpha):\*\* The Laplace smoothing parameter was kept at its default value of 1 to handle potential zero probabilities in categorical data.

- \*\*Feature Scaling:\*\* Features were normalized using the StandardScaler to standardize the input feature space, ensuring compatibility with the algorithm's assumptions.

No additional parameter tuning was required as Gaussian Naive Bayes assumes a normal distribution of continuous features.

#### Naive Bayes Algorithm

The Gaussian Naive Bayes algorithm was utilized for classification. This algorithm operates on the principle of Bayes' theorem, assuming feature independence given the target class. Key steps include:

1. \*\*Data Preprocessing:\*\*

- Features were scaled to have zero mean and unit variance.

- Missing values were imputed using column-wise mean values.

- The dataset was split into training (70%) and testing (30%) subsets, stratified by class distribution.

2. \*\*Model Training:\*\*

- The Gaussian Naive Bayes classifier computes class-conditional probabilities using the Gaussian probability density function.

- The model calculates the posterior probability for each class given the input features, assigning the class with the highest probability to each instance.

3. \*\*Model Evaluation:\*\*

- Predictions were generated on the test set, and evaluation metrics were computed, including accuracy, sensitivity, specificity, F1-score, Matthews correlation coefficient (MCC), and area under the ROC curve (AUC).

- Class probabilities for each instance were extracted to generate ROC curves and assess the classifier's discriminative performance across skill levels.

4. \*\*Feature Importance and Task Analysis:\*\*

- Although Naive Bayes does not directly provide feature importance, the class-wise means (μ) and variances (σ²) were analyzed to identify features contributing to skill-level classification.

- Tasks and muscle groups were aggregated to determine their relevance in distinguishing skill levels, supporting interpretability and actionable insights for surgeon training.

This methodology ensured that the results were robust, interpretable, and aligned with the study's objectives of assessing skill-level classification using nonlinear variability measures.