**Classification Model Comparison Summary**

In evaluating eight classification algorithms on the Breast Cancer Wisconsin dataset, K-Nearest Neighbors (k=5) emerged with the highest accuracy (97.71%), indicating strong performance in correctly identifying malignant and benign cases. Close behind were SVM with RBF kernel (96.57%) and several models clustered around the 96% mark—including Logistic Regression, Naive Bayes, and XGBoost.

While Decision Tree showed the lowest accuracy (93.71%), this is unsurprising given its known tendency to overfit on small datasets. Ensemble methods like Random Forest (95.43%) and XGBoost (96.00%) demonstrated robustness, balancing complexity and predictive power.

Confusion matrices reveal that all models maintained relatively low false-positive and false-negative rates, though subtle differences in misclassifications can be critical in a medical context. For instance, even a few false negatives (misclassified malignant tumors) may be clinically unacceptable, suggesting that top-performing models like KNN or SVM-RBF should be favored for deployment.

These results suggest that while many models perform comparably, models like KNN and SVM (RBF) offer the best trade-off between accuracy and interpretability for this classification task. See the table below:

|  |  |  |
| --- | --- | --- |
| **Classification Model Comparison Summary** | | |
| **Model** | **Accuracy** | **Confusion Matrix** |
| Logistic Regression | 0.9600 | [[117, 1], [6, 51]] |
| K-Nearest Neighbors (k=5) | 0.9771 | [[116, 2], [2, 55]] |
| SVM (Linear Kernel) | 0.9600 | [[116, 2], [5, 52]] |
| SVM (RBF Kernel) | 0.9657 | [[115, 3], [3, 54]] |
| Naive Bayes | 0.9600 | [[113, 5], [2, 55]] |
| Decision Tree | 0.9371 | [[115, 3], [8, 49]] |
| Random Forest | 0.9543 | [[115, 3], [5, 52]] |
| XGBoost | 0.9600 | [[115, 3], [4, 53]] |
|  |  |  |