# Comparison of Machine Learning Models for Gender Classification

|  |  |  |  |  |  |  |
| --- | --- | --- | --- | --- | --- | --- |
| Model | Test Accuracy | Cross-Validation Accuracy (Mean) | Precision (F/M) | Recall (F/M) | F1-Score (Macro Avg) | Key Insights |
| Logistic Regression | 98.91% | 98.91% | High | High | High | No overfitting, robust generalization |
| Decision Tree | 93.28% | 93.13% | 91.35% / 94.33% | 89.69% / 95.28% | 92.66% | Better at classifying males, potential recall improvement for females |
| Linear SVM | 99.35% | 99.22% | Very High | Very High | Very High | Strong performance, but slightly outperformed by RBF SVM |
| RBF SVM | 99.94% | 99.94% | Extremely High | Extremely High | Extremely High | Best model overall, captures non-linear patterns |
| LSTM | 98.61% | 98.18% | High | High | High | Performs well, benefits from sequential dependencies, slightly lower than SVMs |

## Decision Matrices

|  |  |  |  |  |
| --- | --- | --- | --- | --- |
| Model | TN | FP | FN | TP |
| Logistic Regression | 1100 | 54 | 34 | 2042 |
| Decision Tree | 1035 | 119 | 213 | 1863 |
| Linear SVM | 1130 | 24 | 15 | 2061 |
| RBF SVM | 1149 | 5 | 2 | 2074 |
| LSTM | 1125 | 28 | 19 | 2057 |

## Feature Importance

MFCCs are most relevant features (mfcc\_1\_mean, mfcc\_2\_mean, etc.).

## Key Takeaways

✅ Best Model: RBF SVM (99.94% accuracy)

✅ Most Interpretable Model: Logistic Regression (easier to understand feature contributions)

⚠️ Potential Bias: Decision Tree favors male classification slightly more than female

📌 Next Steps: Consider ensemble learning, fairness adjustments, and further hyperparameter tuning.