Performance Evaluation of Naive Bayes Classifier on the Iris Dataset

Karol Korszun

June 2024

1 Introduction

The Naive Bayes classifier is a probabilistic machine learning model based on Bayes' theorem with strong (naive) independence assumptions between the features. In other words it assumes that features are totally independent. Despite its simplicity, it is widely used in various applications due to it s efficiency and effectiveness. This paper evaluates the performance of the Naive Bayes classifier on the Iris dataset. For additional context, we also compare its performance with other classifiers, including Support Vector Machine (SVM), Logistic Regression, and Decision Tree.

2 Methodology

The methodology involves the following steps: data preparation, model training, and performance evaluation. Each step is described in detail below.

2.1 Data Preparation

The Iris dataset consists of 150 samples from three species of Iris flowers (Iris setosa, Iris versicolor, and Iris virginica), with four features measured for each sample. The dataset was normalized using standard scaling to ensure that each feature contributes equally to the model's performance. Various normalization techniques were tested, but they did not significantly affect the results. The dataset was split into training (70%) and testing (30%) subsets.

2.2 Model Training and Evaluation

We implemented the Naive Bayes classifier and compared its performance with three other classifiers: SVM, Logistic Regression, and Decision Tree. The models were trained on the training subset and evaluated on the testing subset using accuracy, precision, recall, and F1-score as performance metrics.

3 Results

The performance of the Naive Bayes classifier and other models is summarized in Table 1. Additionally, Table 2 provides a detailed classification report for the Naive Bayes classifier.

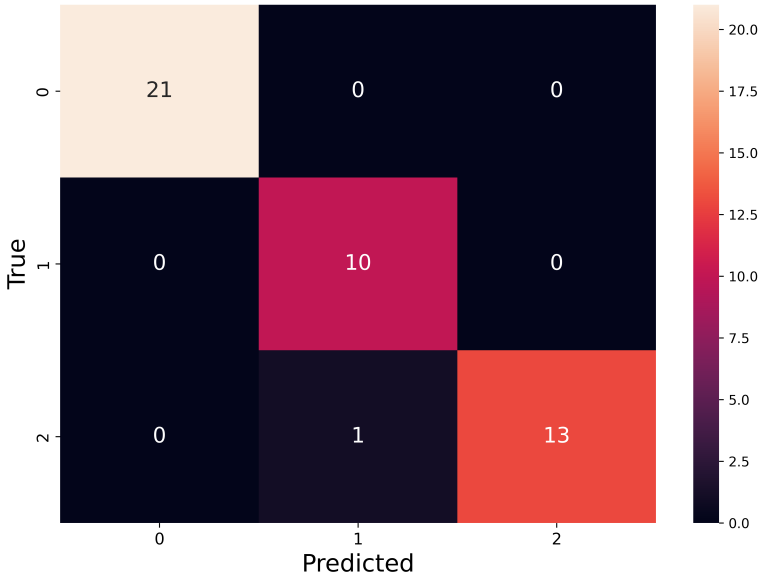

Figure 1: Confusion Matrix for the Naive Bayes Classifier

Model	Accuracy
Naive Bayes	0.9778
SVM	0.9556
Logistic Regression	0.9556
Decision Tree	0.9778

Table 1: Model Performance Comparison

The detailed classification report in Table 2 includes the precision, recall, F1-score, and support for each class.

- **Precision** is the ratio of correctly predicted positive observations to the total predicted positives. A high precision value indicates a low false positive rate. In this case, Class 0 and Class 2 have perfect precision (1.00), indicating no false positives.

	Precision	Recall	F1-Score	Support
0	1.00	1.00	1.00	21
1	0.91	1.00	0.95	10
2	1.00	0.93	0.96	14
Accuracy: 0.98, 45				
Macro Avg: 0.97, 0.98, 0.97, 45				
Weighted Avg: 0.98, 0.98, 0.98, 45				

Table 2: Detailed Classification Report for Naive Bayes

- **Recall** is the ratio of correctly predicted positive observations to all observations in the actual class. A high recall value indicates a low false negative rate. For instance, Class 1 has a perfect recall (1.00), meaning all actual positives are correctly identified.
- **F1-Score** is the weighted average of precision and recall, providing a balance between the two metrics. Class 0 and Class 2 have high F1-scores (1.00 and 0.96, respectively), demonstrating effective performance across both precision and recall.
- **Support** is the number of actual occurrences of each class in the dataset.

4 Conclusion

The Naive Bayes classifier achieved an accuracy of 97.78%, which is on par with the Decision Tree classifier and slightly higher than the SVM and Logistic Regression models. Despite its simplicity and the strong independence assumptions, the Naive Bayes classifier performed well on the Iris dataset. Various normalization techniques, including Min-Max scaling and Robust scaling, were evaluated, but they did not impact the performance of the classifier at all.