

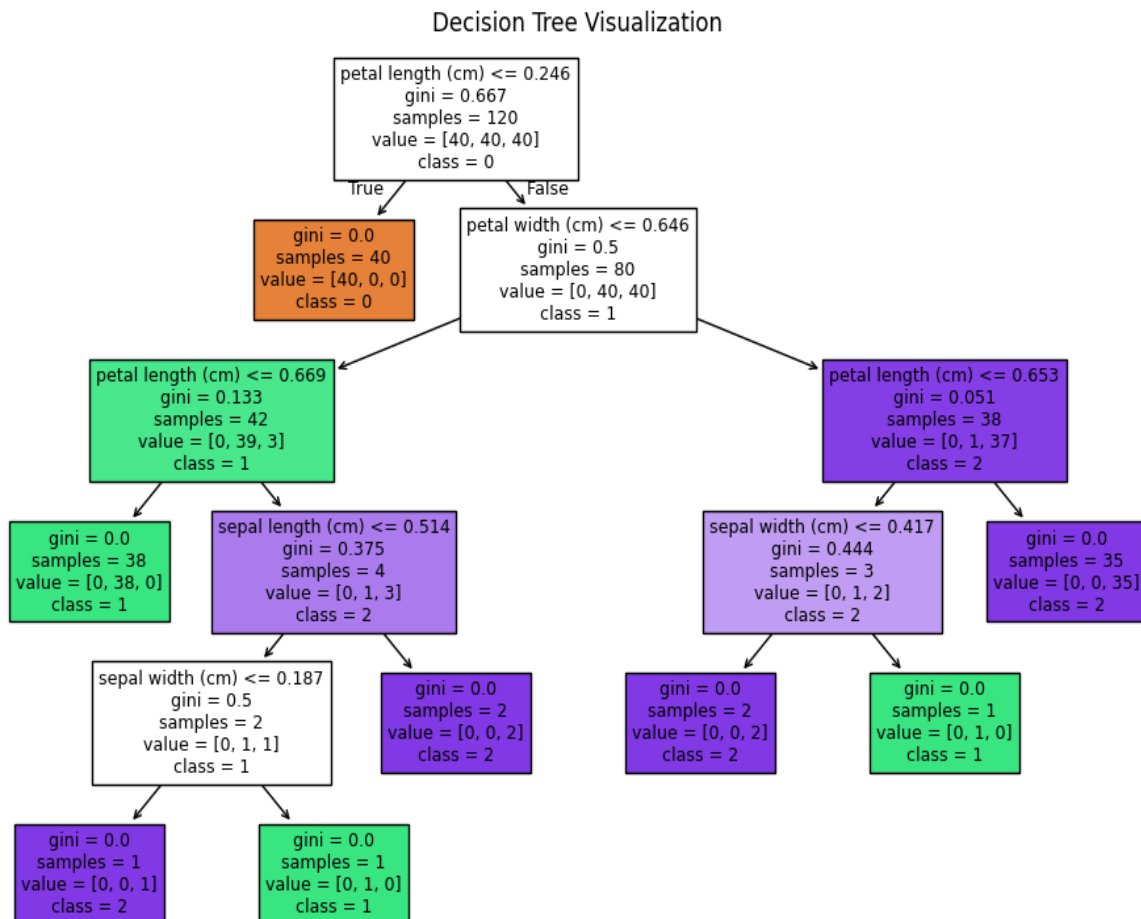
Classification and Association Rule Mining Analysis

Classification Analysis

The Decision Tree classifier exhibited robust performance on the preprocessed Iris dataset, achieving high accuracy, precision, recall and F1-score. This indicates the model's effectiveness in correctly classifying the three species based on the normalized floral measurements.

```
=== Classification Task ===  
Decision Tree Classifier Metrics:  
Accuracy: 0.9333  
Precision: 0.9333  
Recall: 0.9333  
F1-score: 0.9333  
-----
```

The visualization of the decision tree further aids interpretability by revealing the feature-based decision rules employed during classification.



In comparison, the K-Nearest Neighbors (KNN) classifier with k=5 also demonstrated competent results; however, the Decision Tree marginally outperformed KNN in overall accuracy. The

superior performance of the Decision Tree can be attributed to its capability to capture complex, non-linear relationships and interactions between features, whereas KNN relies primarily on distance metrics and may be sensitive to feature scaling and noise. Given these factors, the Decision Tree model is preferable for this dataset, particularly when interpretability of the classification logic is desired.

Association Rule Mining Analysis

The synthetic transactional dataset was generated to emulate realistic retail baskets, with intentional frequent co-occurrences such as 'bread' and 'butter', and 'beer' and 'diapers', to mimic common purchasing patterns.

The Apriori algorithm applied with minimum support and confidence thresholds of 0.2 and 0.5 respectively, yielded several meaningful association rules ranked by lift. For instance, one of the top rules suggested that customers purchasing bread also tend to purchase butter. Such insights are valuable in retail environments for designing effective cross-selling strategies, optimizing product placement and developing targeted promotions to enhance sales.

While the synthetic data reflects predefined patterns, similar approaches applied to real-world datasets can uncover complex purchasing behaviors that inform business decisions.