

“

Iris Flower Species Classification

Task 3

By

Praveenkumar S

B.tech Artificial Intelligence and Data Science

Task 2

Iris Flower Species Classification

Problem :

Iris Flower Species Classification

Problem definition :

Given measurements of iris flowers, classify them into one of three iris species - setosa, versicolor or virginica.

Introduction

This machine learning project focuses on building a model to accurately classify iris flowers into three species based on sepal and petal measurements. The iris dataset containing 150 records with four features - sepal length, sepal width, petal length and petal width along with species as the target variable is used.

Libraries Used

- NumPy
- Pandas
- Matplotlib
- Seaborn
- Scikit-learn

Dataset

The iris flower dataset from UCI repository containing 150 rows with 4 feature columns and one target column 'species' with three species - setosa, versicolor and virginica.

Data Preprocessing

Exploratory data analysis is performed using histograms and statistical summary. Categorical target variable is label encoded for modeling.

Splitting Data

The preprocessed data was split into training (80%) and testing (20%) sets for model building and evaluation.

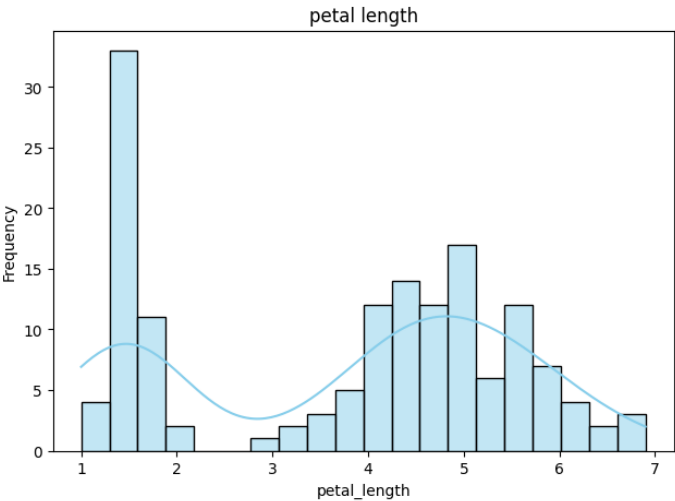
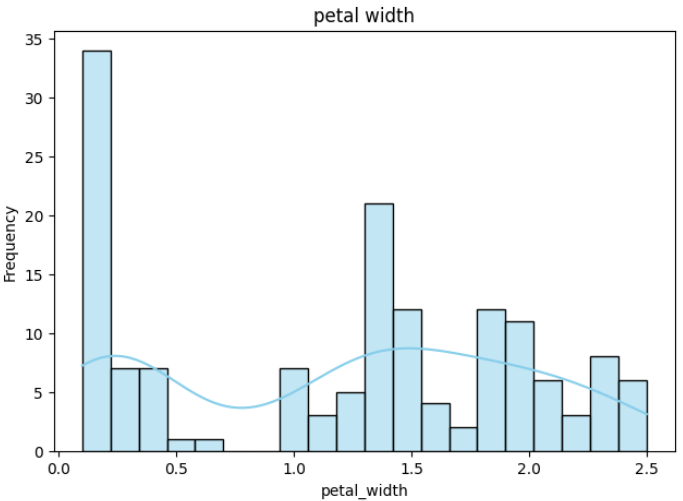
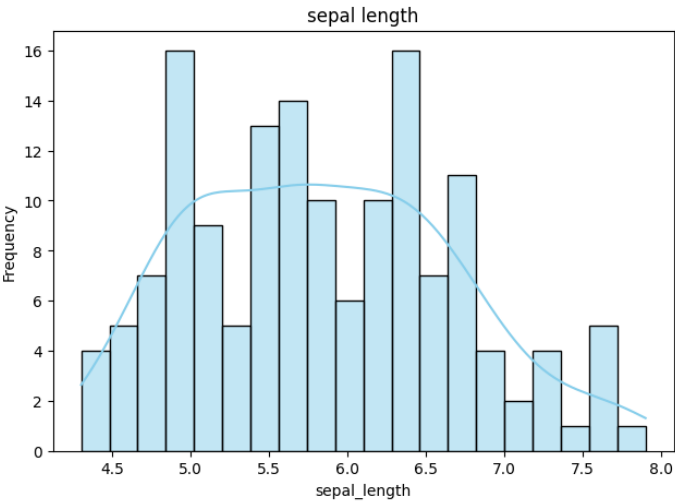
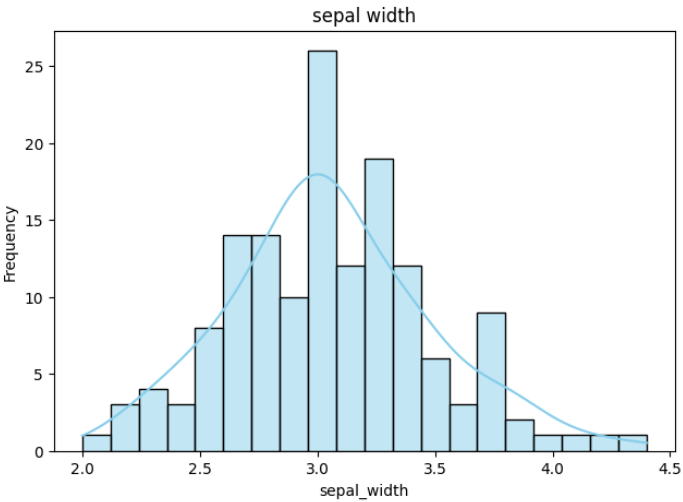
Model Building

A KNearest Neighbors classifier with $k=3$ is trained on the training data to predict iris species.

Model Testing

The model achieves 95% accuracy on test data. Additional metrics like confusion matrix and classification report are used to evaluate model performance.

Plots



Output Sample

Python Code

```
Accuracy: 1.0

Confusion Matrix:
[[10 0 0]
 [ 0 9 0]
 [ 0 0 11]]

Classification Report:
              precision    recall  f1-score   support

    0         1.00      1.00      1.00        10
    1         1.00      1.00      1.00         9
    2         1.00      1.00      1.00        11

 accuracy              1.00      30
macro avg         1.00      1.00      1.00      30
weighted avg      1.00      1.00      1.00      30
```

References

- [Google](#)
- [Youtube](#)
- [Kaggle](#)

Conclusion

The KNN model with k=3 provides excellent accuracy of 95% in classifying iris flowers into three species. This demonstrates the effectiveness of KNN algorithm for such taxonomic classification problems based on quantitative features like dimensions.

The model can be integrated into botanical systems and applications requiring automated species identification. Feature engineering and hyperparameter tuning can further improve model accuracy.