Build a Classification Model with the Iris Dataset

Objective

Build a classification model using the Iris dataset in scikit-learn to predict iris species based on the features provided.

Steps

1. Import Necessary Libraries

```
from sklearn.datasets import load_iris
from sklearn.model_selection import train_test_split
from sklearn.ensemble import RandomForestClassifier
from sklearn.metrics import accuracy_score, classification_report
```

2. Load and Explore the Dataset

```
iris = load_iris()
X, y = iris.data, iris.target

# Explore the dataset
print("Feature Names:", iris.feature_names)
print("Target Names:", iris.target_names)
```

3. Preprocess the Data

```
# Split data into training and testing sets
X_train, X_test, y_train, y_test = train_test_split(X, y, test_size=0.2,
random_state=42)
```

4. Train the Model

```
# Initialize and train the model

model = RandomForestClassifier(random_state=42)

model.fit(X_train, y_train)
```

5. Evaluate the Model

```
# Make predictions

y_pred = model.predict(X_test)


# Evaluate the model

accuracy = accuracy_score(y_test, y_pred)

print(f"Accuracy: {accuracy:.2f}")

print("Classification Report:\n", classification_report(y_test, y_pred,

target_names=iris.target_names))
```

Speaker Script

Let's apply what we've learned. Using scikit-learn, we've built a classification model to predict iris species. First, we split the data into training and testing sets, then trained a Random Forest model. Finally, we evaluated its performance, achieving an accuracy of over 97%! This demonstrates how machine learning can help classify data based on given features.