Building Interpretable Models with scikit-learn

Hands-on Workflow & Code Examples

Implementation Steps

1 **Import & Load Data**
Load dataset and dependencies

↓

2 **Preprocessing**
Feature scaling, categorical encoding

↓

3 **Train Models**
Linear, Tree-based classifiers

↓

Code Examples

📊 Linear Model

```
from sklearn.linear_model import LogisticRegression

model = LogisticRegression()
model.fit(X_train, y_train)

# Access coefficients
coef = model.coef_
print(f"Coefficients: {coef}")
```

4 **Extract Insights**
Coefficients, feature importance

5 **Visualize & Compare**
Accuracy vs interpretability

Decision Tree

```
from sklearn.tree import DecisionTreeClassifier

tree = DecisionTreeClassifier(max_depth=3)
tree.fit(X_train, y_train)

# Feature importance
importance = tree.feature_importances_
```

Inspection Tools

```
from sklearn.inspection import PartialDependenceDisplay

# Visualize partial dependence
PartialDependenceDisplay.from_estimator(
    model, X, features=[0, 1]
)
```

Key sklearn Modules

`linear_model`	`tree`
`inspection`	`preprocessing`