ml-lab4

February 28, 2024

[22]:
```
#importing libraries
from sklearn.datasets import load_diabetes
#warning.filterwarning('ignore')
from sklearn.model_selection import KFold
import matplotlib.pyplot as plt
from sklearn.tree import DecisionTreeClassifier, plot_tree
from sklearn.metrics import mean_absolute_error, mean_squared_error, r2_score
```

[23]:
```
diabetes = load_diabetes()
X = diabetes.data
y = diabetes.target
```

[24]:
```
kf = KFold(n_splits=6, shuffle=True, random_state=42)
```

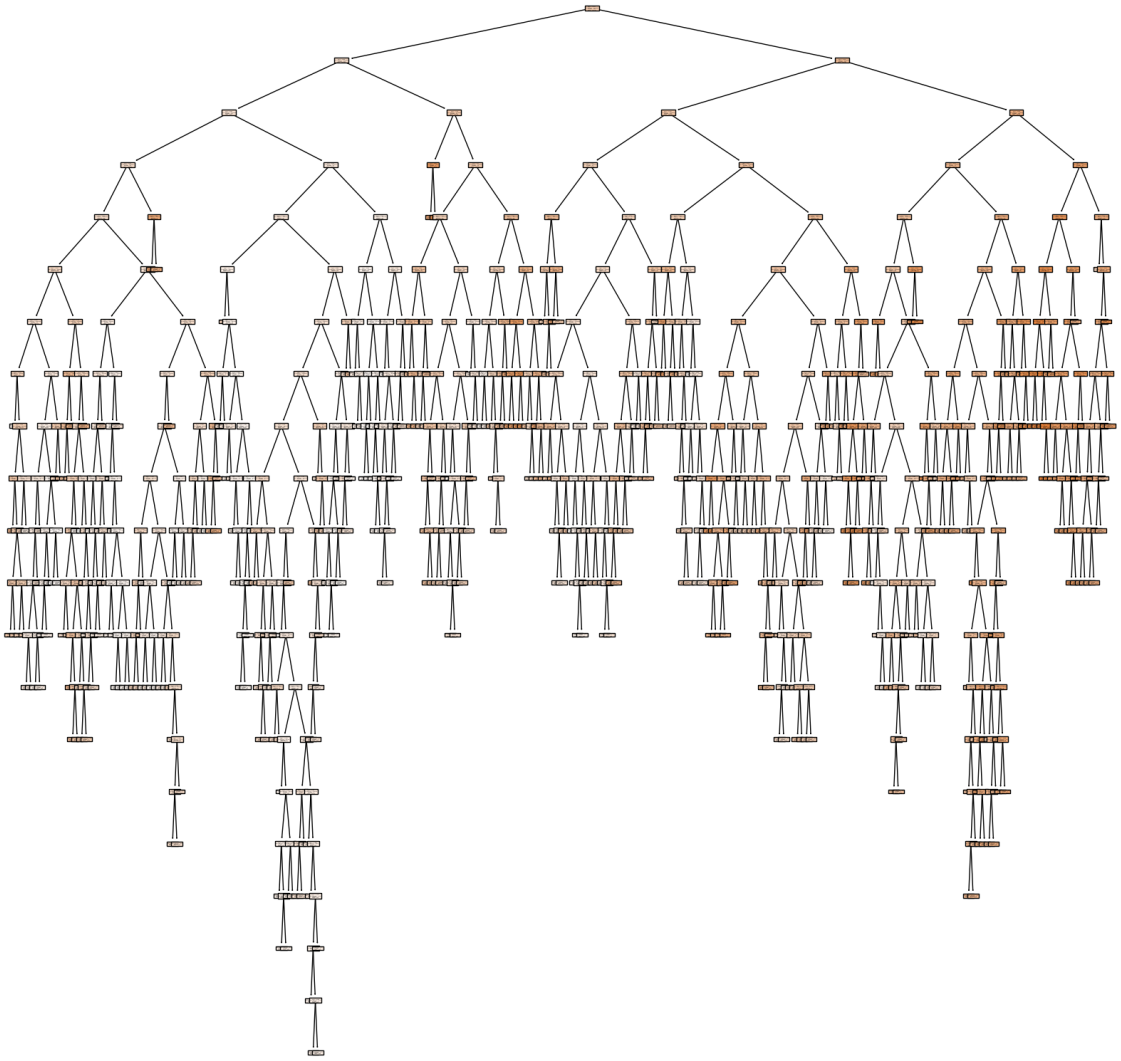
[30]:
```
for fold_idx, (train_index, test_index) in enumerate(kf.split(X)):
    X_train, X_test = X[train_index], X[test_index]
    y_train, y_test = y[train_index], y[test_index]

print(f"Fold {fold_idx + 1}:")
print(f"  Training samples: {len(X_train)}")
print(f"  Test samples: {len(X_test)}")
```

```
Fold 6:
  Training samples: 369
  Test samples: 73
```

[35]:
```
regressor = DecisionTreeRegressor()
regressor.fit(X_train, y_train)
```

[35]:
```
DecisionTreeRegressor()
```

[29]:
```
# Plot the decision tree
plt.figure(figsize=(20,20))
plot_tree(regressor, filled=True)
plt.show()
```



[36]:
```
y_pred = regressor.predict(X_test)
```

[40]:
```
from sklearn.metrics import mean_absolute_error, mean_squared_error, r2_score

mae = mean_absolute_error(y_test, y_pred)
mse = mean_squared_error(y_test, y_pred)
r2 = r2_score(y_test, y_pred)
```

[44]:
```
print("Mean Absolute Error:", mae)
print("Mean Squared Error:", mse)
print("R-squared:", r2)
```

```
Mean Absolute Error: 71.04109589041096
Mean Squared Error: 7882.328767123287
```

```
R-squared: -0.7273596458834859
```