⌄ *Prodigy InfoTech Task-3*

```
#importing libraries
import pandas as pd
from sklearn.model_selection import train_test_split
from sklearn.tree import DecisionTreeClassifier
from sklearn.metrics import accuracy_score, classification_report
from sklearn.datasets import fetch_openml
from sklearn.preprocessing import LabelEncoder


# Load the Adult dataset
adult = fetch_openml(name='adult', as_frame=True)
```

> ⇥ /usr/local/lib/python3.10/dist-packages/sklearn/datasets/_openml.py:301: UserWarning: Multiple active versions of the dataset matchi
>   error_msg,
>   /usr/local/lib/python3.10/dist-packages/sklearn/datasets/_openml.py:968: FutureWarning: The default value of `parser` will change fr
>   Data columns (total 15 columns):

```
# Preprocess the data
X = adult.data
y = adult.target


# Identify categorical columns (correctly this time!)
categorical_cols = X.select_dtypes(include=['category', 'object']).columns


# Apply label encoding to categorical features
label_encoders = {}
for col in categorical_cols:
    le = LabelEncoder()
    X[col] = le.fit_transform(X[col])
    label_encoders[col] = le
```

> ⇥ <ipython-input-11-23f5f4d8fb8b>:5: SettingWithCopyWarning:
>   A value is trying to be set on a copy of a slice from a DataFrame.
>   Try using .loc[row_indexer,col_indexer] = value instead
>
>   See the caveats in the documentation: https://pandas.pydata.org/pandas-docs/stable/user_guide/indexing.html#returning-a-view-vers
>   X[col] = le.fit_transform(X[col])
>   <ipython-input-11-23f5f4d8fb8b>:5: SettingWithCopyWarning:
>   A value is trying to be set on a copy of a slice from a DataFrame.
>   Try using .loc[row_indexer,col_indexer] = value instead
>
>   See the caveats in the documentation: https://pandas.pydata.org/pandas-docs/stable/user_guide/indexing.html#returning-a-view-vers
>   X[col] = le.fit_transform(X[col])
>   <ipython-input-11-23f5f4d8fb8b>:5: SettingWithCopyWarning:
>   A value is trying to be set on a copy of a slice from a DataFrame.
>   Try using .loc[row_indexer,col_indexer] = value instead
>
>   See the caveats in the documentation: https://pandas.pydata.org/pandas-docs/stable/user_guide/indexing.html#returning-a-view-vers
>   X[col] = le.fit_transform(X[col])
>   <ipython-input-11-23f5f4d8fb8b>:5: SettingWithCopyWarning:
>   A value is trying to be set on a copy of a slice from a DataFrame.
>   Try using .loc[row_indexer,col_indexer] = value instead
>
>   See the caveats in the documentation: https://pandas.pydata.org/pandas-docs/stable/user_guide/indexing.html#returning-a-view-vers
>   X[col] = le.fit_transform(X[col])
>   <ipython-input-11-23f5f4d8fb8b>:5: SettingWithCopyWarning:
>   A value is trying to be set on a copy of a slice from a DataFrame.
>   Try using .loc[row_indexer,col_indexer] = value instead
>
>   See the caveats in the documentation: https://pandas.pydata.org/pandas-docs/stable/user_guide/indexing.html#returning-a-view-vers
>   X[col] = le.fit_transform(X[col])
>   <ipython-input-11-23f5f4d8fb8b>:5: SettingWithCopyWarning:
>   A value is trying to be set on a copy of a slice from a DataFrame.
>   Try using .loc[row_indexer,col_indexer] = value instead
>
>   See the caveats in the documentation: https://pandas.pydata.org/pandas-docs/stable/user_guide/indexing.html#returning-a-view-vers
>   X[col] = le.fit_transform(X[col])
>   <ipython-input-11-23f5f4d8fb8b>:5: SettingWithCopyWarning:
>   A value is trying to be set on a copy of a slice from a DataFrame.
>   Try using .loc[row_indexer,col_indexer] = value instead
>
>   See the caveats in the documentation: https://pandas.pydata.org/pandas-docs/stable/user_guide/indexing.html#returning-a-view-vers
>   X[col] = le.fit_transform(X[col])
>   <ipython-input-11-23f5f4d8fb8b>:5: SettingWithCopyWarning:
>   A value is trying to be set on a copy of a slice from a DataFrame.
>   Try using .loc[row_indexer,col_indexer] = value instead
>
>   See the caveats in the documentation: https://pandas.pydata.org/pandas-docs/stable/user_guide/indexing.html#returning-a-view-vers
```

```
      X[col] = le.fit_transform(X[col])
    <ipython-input-11-23f5f4d8fb8b>:5: SettingWithCopyWarning:
    A value is trying to be set on a copy of a slice from a DataFrame.
    Try using .loc[row_indexer,col_indexer] = value instead

    See the caveats in the documentation: https://pandas.pydata.org/pandas-docs/stable/user_guide/indexing.html#returning-a-view-vers
      X[col] = le.fit_transform(X[col])
    <ipython-input-11-23f5f4d8fb8b>:5: SettingWithCopyWarning:
    A value is trying to be set on a copy of a slice from a DataFrame.
    Try using .loc[row_indexer,col_indexer] = value instead
```

```python
# Split the 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)
```

```python
# Create a decision tree classifier
clf = DecisionTreeClassifier(random_state=42)
```

```python
# Train the model
clf.fit(X_train, y_train)
```

```
        ▾         DecisionTreeClassifier
    DecisionTreeClassifier(random_state=42)
```

```python
# Make predictions
y_pred = clf.predict(X_test)
```

```python
# Evaluate the model
accuracy = accuracy_score(y_test, y_pred)
print('Accuracy:', accuracy)
print('Classification Report:')
print(classification_report(y_test, y_pred))
```

```
    Accuracy: 0.7968062237690654
    Classification Report:
                  precision    recall  f1-score   support

           <=50K       0.87      0.87      0.87      7414
            >50K       0.58      0.58      0.58      2355

        accuracy                           0.80      9769
       macro avg       0.72      0.72      0.72      9769
    weighted avg       0.80      0.80      0.80      9769
```