Online Fraud Detection Code

```
import pandas as pd
import numpy as np
import plotly.express as px
from sklearn.model_selection import train_test_split
from sklearn.tree import DecisionTreeClassifier

data = pd.read_csv("../input/online-payment-fraud-detection/onlinefraud.csv")

print(data.head())

data.shape
# Now, let’s have a look at whether this dataset has any null values or not

print(data.isnull().sum())
# Exploring transaction type

print(data.type.value_counts())

type = data["type"].value_counts()
transactions = type.index
quantity = type.values
figure = px.pie(data, values=quantity, names=transactions, hole = 0.5,
title="Distribution of Transaction Type")
figure.show()
# Now let’s have a look at the correlation between the features of the data
with the isFraud column
# Checking correlation

correlation = data.corr()
print(correlation["isFraud"].sort_values(ascending=False))
# Now let’s transform the categorical features into numerical. Here we will
also transform the values of the isFraud column into
# No Fraud and Fraud labels to have a better understanding of the output
# Changing CASH_OUT to 1, PAYMENT to 2, CASH_IN to 3, TRANSFER to 4 and DEBIT
to 5

data["type"] = data["type"].map({"CASH_OUT": 1, "PAYMENT": 2, "CASH_IN": 3,
"TRANSFER": 4, "DEBIT": 5})
data["isFraud"] = data["isFraud"].map({0: "No Fraud", 1: "Fraud"})
print(data.head())
# splitting the data

x = np.array(data[["type", "amount", "oldbalanceOrg", "newbalanceOrig"]])
y = np.array(data[["isFraud"]])
```

```
# training a machine learning model

xtrain, xtest, ytrain, ytest = train_test_split(x, y, test_size=0.10,
random_state=42)
model = DecisionTreeClassifier()
model.fit(xtrain, ytrain)
print(model.score(xtest, ytest))
# prediction
#features = [type, amount, oldbalanceOrg, newbalanceOrig]

features = np.array([[1, 8900.2, 8990.2, 0.0]])
print(model.predict(features))
```