# **Sir Syed University of Engineering & Technology (SSUET)**

# **Software Engineering Department**

***Course Name: Artificial Intelligence (SE-314L)***

FRAUD DETECTION CODE

import pandas as pd

import matplotlib.pyplot as plt

import seaborn as sns

from sklearn.metrics import classification\_report, accuracy\_score

from sklearn.ensemble import IsolationForest

from sklearn.neighbors import LocalOutlierFactor

# Load the dataset using pandas

data = pd.read\_csv(r'C:\Users\yagom\OneDrive\デスクトップ\AI PROJECT\creditcard.csv')  # Updated file path

# Dataset exploration

print(data.columns)

# Sample 10% of the dataset for quick processing

data = data.sample(frac=0.1, random\_state=1)

print(data.shape)

print(data.describe())

# V1 - V28 are the results of PCA dimensionality reduction

# Plot histograms of each parameter

data.hist(figsize=(20, 20))

plt.show()

# Determine the number of fraud cases in the dataset

Fraud = data[data['Class'] == 1]

Valid = data[data['Class'] == 0]

outlier\_fraction = len(Fraud) / float(len(Valid))

print(f"Outlier Fraction: {outlier\_fraction}")

print(f"Fraud Cases: {len(Fraud)}")

print(f"Valid Transactions: {len(Valid)}")

# Correlation matrix

corrmat = data.corr()

plt.figure(figsize=(12, 9))

sns.heatmap(corrmat, vmax=.8, square=True)

plt.show()

# Prepare the data

columns = data.columns.tolist()

columns = [c for c in columns if c not in ["Class"]]  # Exclude the target column

target = "Class"

X = data[columns]

Y = data[target]

# Print shapes of the feature set and target

print(f"Feature set shape: {X.shape}")

print(f"Target shape: {Y.shape}")

# Define random states

state = 1

# Define outlier detection tools

classifiers = {

    "Isolation Forest": IsolationForest(max\_samples=len(X),

                                        contamination=outlier\_fraction,

                                        random\_state=state),

    "Local Outlier Factor": LocalOutlierFactor(

        n\_neighbors=20,

        contamination=outlier\_fraction)

}

# Apply models and evaluate

plt.figure(figsize=(9, 7))

n\_outliers = len(Fraud)

for clf\_name, clf in classifiers.items():

    if clf\_name == "Local Outlier Factor":

        y\_pred = clf.fit\_predict(X)

        scores\_pred = clf.negative\_outlier\_factor\_

    else:

        clf.fit(X)

        scores\_pred = clf.decision\_function(X)

        y\_pred = clf.predict(X)

    # Reshape the prediction values: 0 for valid, 1 for fraud

    y\_pred = [0 if pred == 1 else 1 for pred in y\_pred]

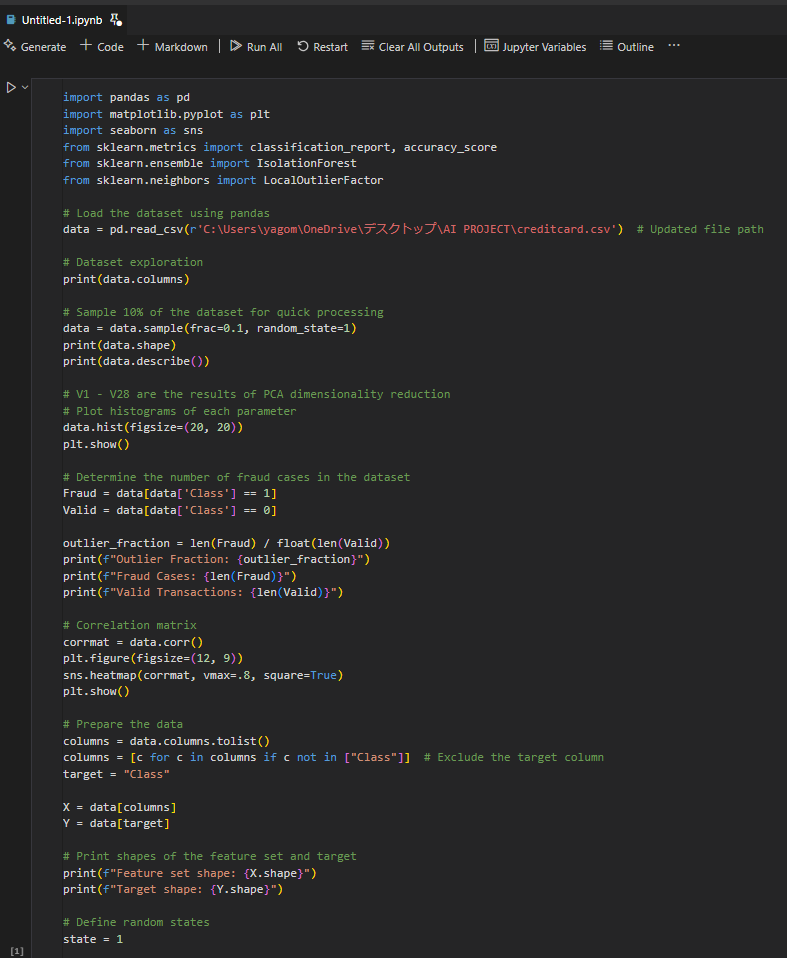
    n\_errors = (y\_pred != Y).sum()

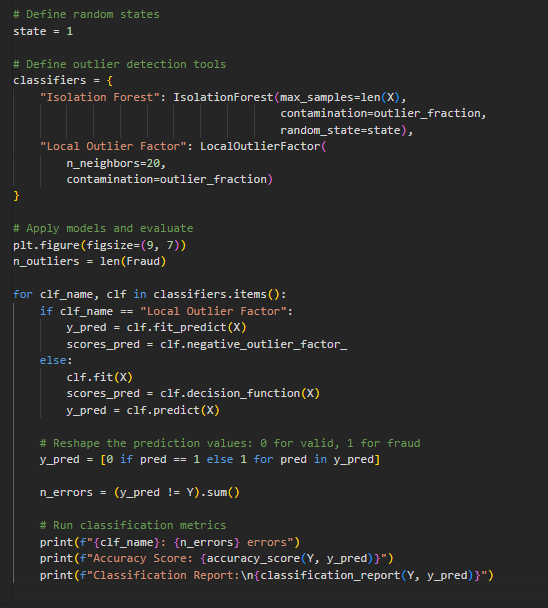
    # Run classification metrics

    print(f"{clf\_name}: {n\_errors} errors")

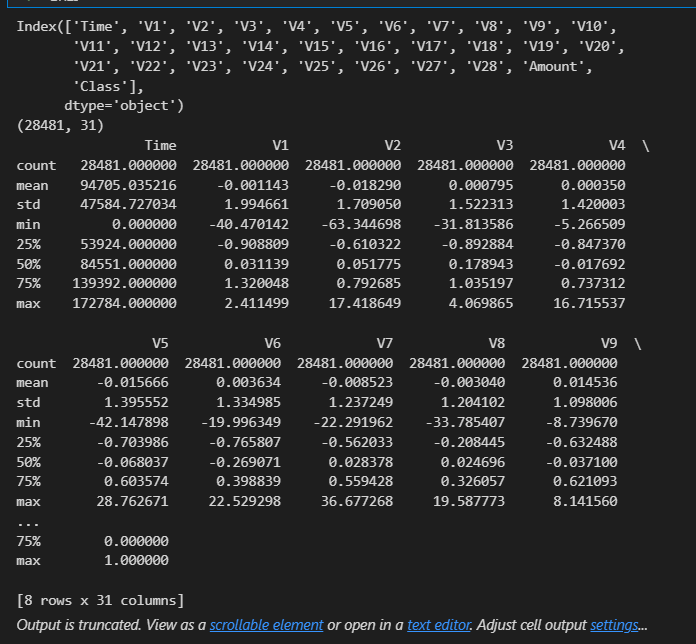
    print(f"Accuracy Score: {accuracy\_score(Y, y\_pred)}")

    print(f"Classification Report:\n{classification\_report(Y, y\_pred)}")





INDEX



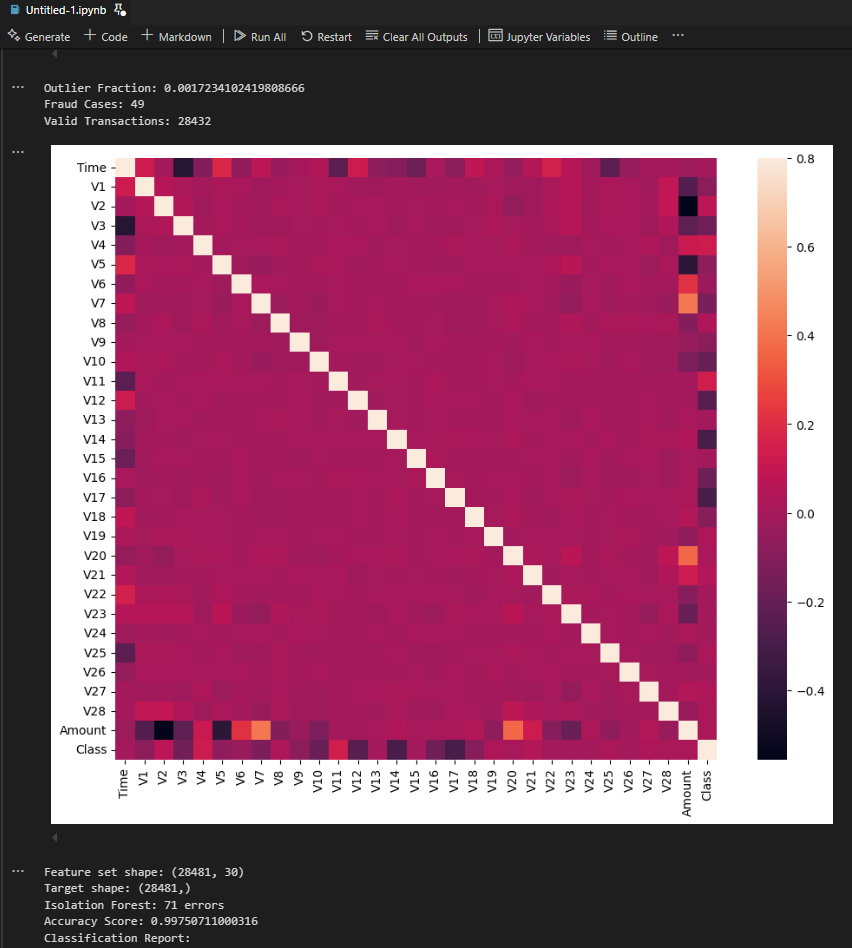
GRAPH

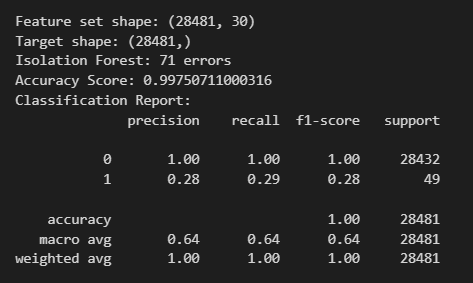
A screenshot of a graph

Description automatically generated

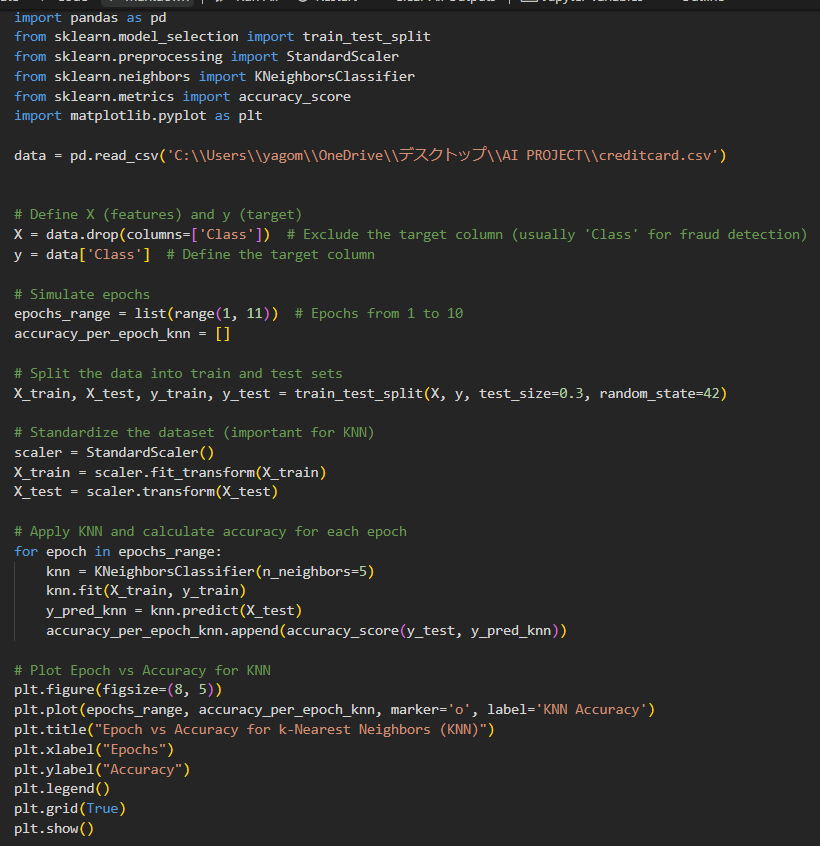
A screenshot of a computer code

Description automatically generated





KNN ALGO



OUTPUT

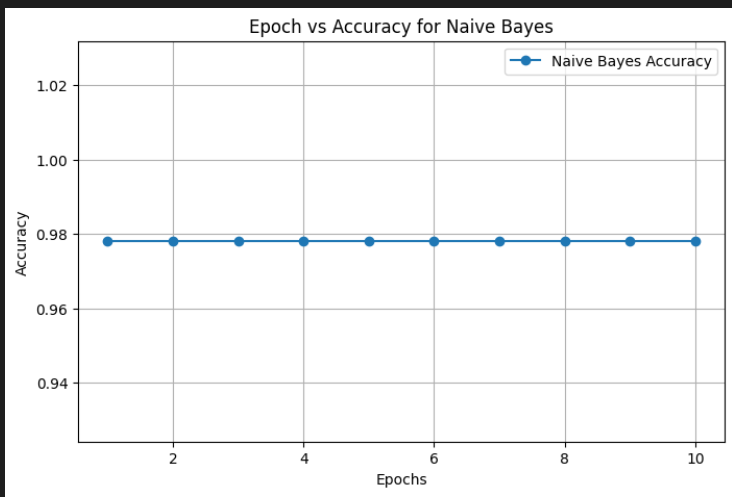
A graph with a line

Description automatically generated

NAÏVE BAYES

A computer screen shot of a program

Description automatically generatedOUTPUT



DECISION TREE

A computer screen shot of a program code

Description automatically generated

OUTPUT

A graph with a line

Description automatically generated

LINEAR REGRESSION

A computer screen shot of a program

Description automatically generated

OUTPUT

A graph with a line

Description automatically generated

MULTI REGRESSION

A computer screen shot of a program

Description automatically generated

OUTPUT

A graph with a line

Description automatically generated

COMPARISON OF ALL ALGOS

A screen shot of a computer program

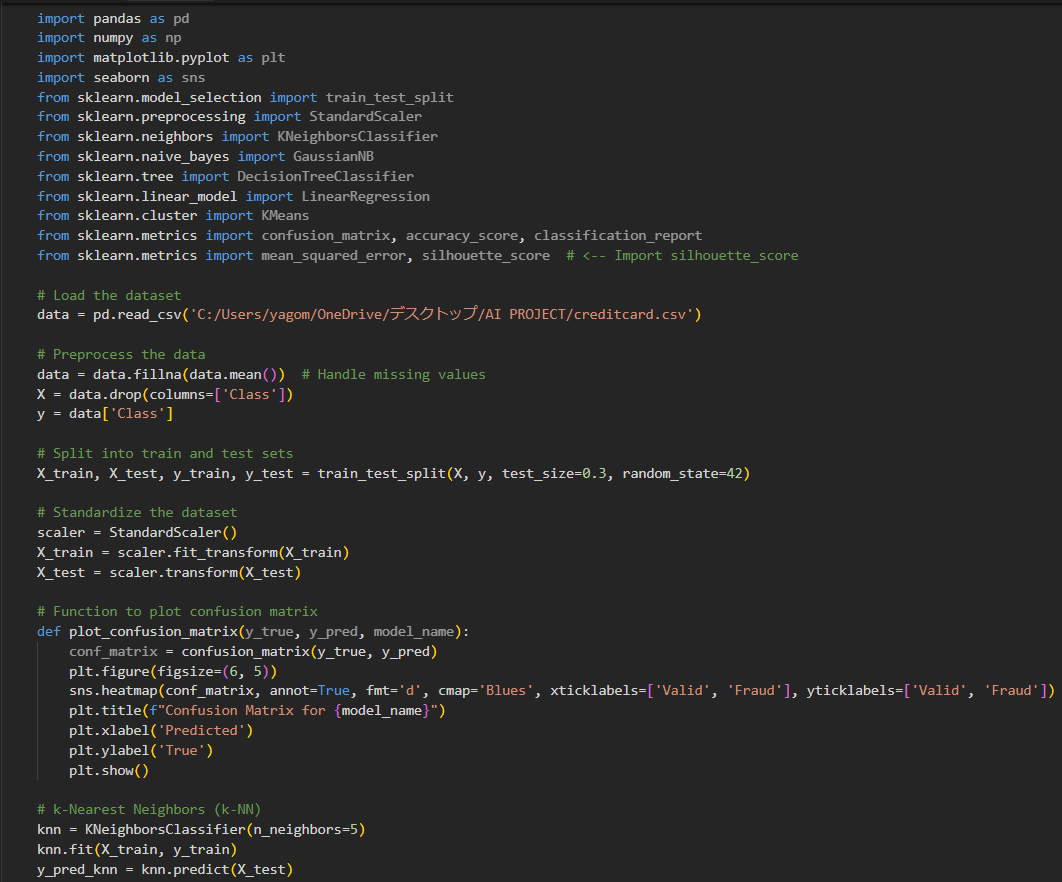
Description automatically generated

OUTPUT

A graph with colored lines and numbers

Description automatically generated

CONFUSION MATRIX AND CLASSIFICATION REPORT



A computer screen shot of a program code

Description automatically generated

A screen shot of a computer program

Description automatically generated

A computer screen shot of text

Description automatically generated

KNN ALGO

A screenshot of a computer screen

Description automatically generated

A graph with numbers and a blue square

Description automatically generated

NAÏVE

A screenshot of a computer screen

Description automatically generated

A graph showing a blue square with white text

Description automatically generated

DECISION TREE

A screenshot of a computer

Description automatically generated

A graph with numbers and a blue square

Description automatically generated

LINEAR REGRESSION

A screenshot of a computer

Description automatically generated

A graph with blue squares and numbers

Description automatically generated

MULTI LINEAR REGRESSION

K MEANS

A graph with a blue square

Description automatically generated