#!/usr/bin/env python

# coding: utf-8

# In[6]:

import pandas as pd

import numpy as np

import matplotlib.pyplot as plt

from sklearn.datasets import load\_diabetes, load\_breast\_cancer

from sklearn.model\_selection import train\_test\_split

from sklearn.linear\_model import LinearRegression, LogisticRegression

from sklearn.metrics import mean\_squared\_error, accuracy\_score

# Load diabetes dataset

diabetes = load\_diabetes()

df\_diabetes = pd.DataFrame(diabetes.data, columns=diabetes.feature\_names)

df\_diabetes['target'] = diabetes.target

# Perform EDA on diabetes dataset

print(df\_diabetes.describe())

for col in df\_diabetes.columns:

plt.hist(df\_diabetes[col])

plt.title(col)

plt.show()

# Split diabetes dataset into train and test sets

X\_diabetes = df\_diabetes.drop('target', axis=1)

y\_diabetes = df\_diabetes['target']

X\_train, X\_test, y\_train, y\_test = train\_test\_split(X\_diabetes, y\_diabetes, test\_size=0.2, random\_state=0)

# Train linear regression model on diabetes dataset

reg = LinearRegression().fit(X\_train, y\_train)

y\_pred = reg.predict(X\_test)

mse = mean\_squared\_error(y\_test, y\_pred)

print("Mean Squared Error:", mse)

# In[7]:

# Load breast cancer dataset

cancer = load\_breast\_cancer()

df\_cancer = pd.DataFrame(cancer.data, columns=cancer.feature\_names)

df\_cancer['target'] = cancer.target

# Perform EDA on breast cancer dataset

print(df\_cancer.describe())

for col in df\_cancer.columns:

plt.hist(df\_cancer[col])

plt.title(col)

plt.show()

# Split breast cancer dataset into train and test sets

X\_cancer = df\_cancer.drop('target', axis=1)

y\_cancer = df\_cancer['target']

X\_train, X\_test, y\_train, y\_test = train\_test\_split(X\_cancer, y\_cancer, test\_size=0.2, random\_state=0)

# Train logistic regression model on breast cancer dataset

clf = LogisticRegression().fit(X\_train, y\_train)

y\_pred = clf.predict(X\_test)

acc = accuracy\_score(y\_test, y\_pred)

print("Accuracy:", acc)

# In[ ]: