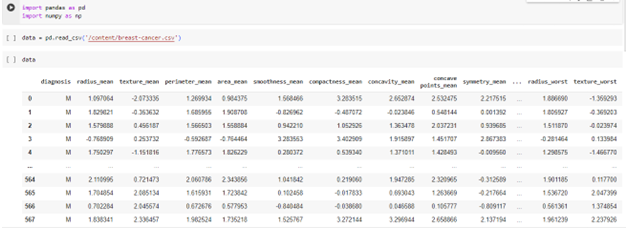
**Name :** Pathipati Brahma Teja

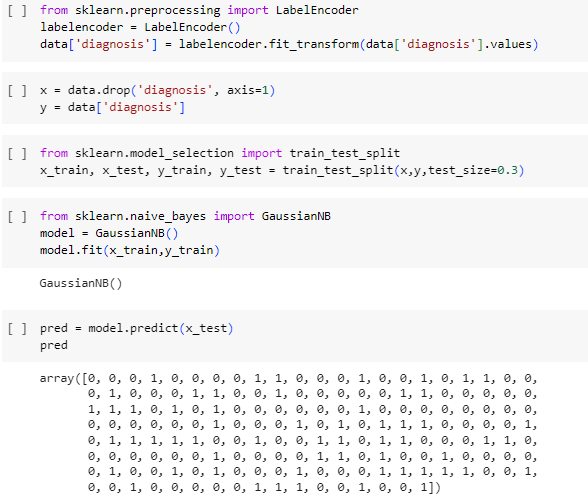
**Registration Number:** 99220041299

1. **Utilize a dataset of your preference to implement and demonstrate all the feature engineering techniques covered in the workshop. Present the results and outputs obtained from the application of these techniques.**

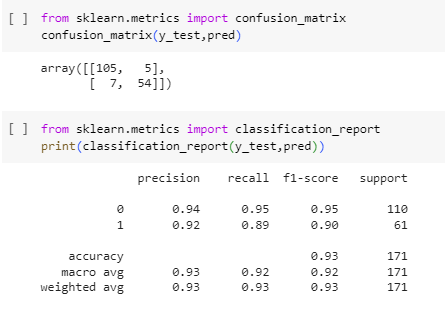
* The given data set has both Qualitative and Quantitative data features.



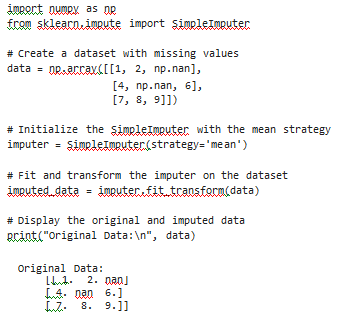
* Feature Improvement:



* Feature Selection:



* Imputing:



* Encoding:

import pandas as pd

# Create a DataFrame with a categorical column

data = {'Color': ['Red', 'Green', 'Blue', 'Red', 'Green']} df = pd.DataFrame(data)

# Perform one-hot encoding

one\_hot\_encoded = pd.get\_dummies(df, columns=['Color'], prefix='Color')

# Display the original and one-hot encoded DataFrames print("Original DataFrame:\n", df)

print("\nOne-Hot Encoded DataFrame:\n", one\_hot\_encoded)

Original DataFrame:

Color

1. Red
2. Green
3. Blue
4. Red
5. Green

One-Hot Encoded DataFrame:

Color\_Blue Color\_Green Color\_Red 0 0 0 1

1 0 1 0

2 1 0 0

3 0 0 1

4 0 1 0

* Log Transformation:

import numpy as np

import pandas as pd

import matplotlib.pyplot as plt

# Generate some example data

data = {'Feature': np.random.exponential(scale=2, size=1000)}

# Create a DataFrame

df = pd.DataFrame(data)

# Plot the original distribution plt.figure(figsize=(12, 5))

plt.subplot(1, 2, 1)

plt.hist(df['Feature'], bins=30, color='blue', edgecolor='black', alpha=0.7) plt.title('Original Distribution')

plt.xlabel('Feature') plt.ylabel('Frequency')

# Apply a logarithmic transformation to the feature

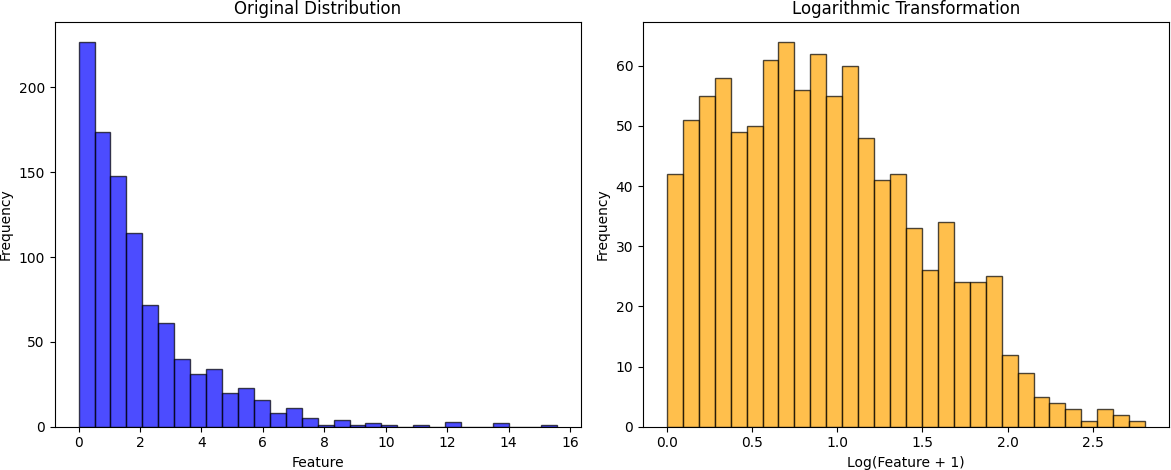
df['LogFeature'] = np.log1p(df['Feature']) # Adding 1 to handle zero values

# Plot the transformed distribution plt.subplot(1, 2, 2)

plt.hist(df['LogFeature'], bins=30, color='orange', edgecolor='black', alpha=0.7) plt.title('Logarithmic Transformation')

plt.xlabel('Log(Feature + 1)') plt.ylabel('Frequency')

plt.tight\_layout() plt.show()



**2.Describe and elaborate on the practical applications of the feature engineering techniques covered in the workshop/seminar within real-life scenarios and industries. Provide insights into how these techniques are utilized and their significance in various contexts**

**Practical Applications of Feature Engineering:**

1. Finance: In the finance industry, feature engineering is used to predict stock prices, detect fraudulent transactions, and assess credit risk. For example, feature engineering can be used to extract features such as moving averages, relative strength index, and momentum from stock prices to predict future prices.

2. Healthcare: In the healthcare industry, feature engineering is used to predict disease outcomes, identify high-risk patients, and improve patient care. For example, feature engineering can be used to extract features such as age, gender, and medical history from patient data to predict the likelihood of developing a particular disease.

3. Marketing: In the marketing industry, feature engineering is used to predict customer behavior, segment customers, and personalize marketing campaigns. For example, feature engineering can be used to extract features such as purchase history, demographics, and web browsing behavior from customer data to predict which products a customer is likely to buy.

4. Manufacturing: In the manufacturing industry, feature engineering is used to optimize production processes, detect equipment failures, and improve product quality. For example, feature engineering can be used to extract features such as temperature, pressure, and humidity from sensor data to predict when a machine is likely to fail.

5. Transportation: In the transportation industry, feature engineering is used to optimize routes, predict travel times, and improve safety. For example, feature engineering can be used to extract features such as traffic volume, weather conditions, and road conditions from transportation data to predict travel times.