```python
# =========================
# Install PySpark and Import Libraries
# =========================

# Install PySpark
!pip install pyspark --quiet

# Import necessary libraries
import pyspark
from pyspark.sql import SparkSession
from pyspark.sql.functions import col
from pyspark.ml.feature import StringIndexer, OneHotEncoder,
VectorAssembler, StandardScaler
from pyspark.ml import Pipeline, PipelineModel
from pyspark.ml.classification import (
    LogisticRegression, RandomForestClassifier,
MultilayerPerceptronClassifier,
    LogisticRegressionModel, RandomForestClassificationModel,
MultilayerPerceptronClassificationModel
)
from pyspark.ml.evaluation import MulticlassClassificationEvaluator
import shutil
import os


import urllib.request

# =========================
# Start Spark Session
# =========================

# Initialize a Spark session
spark =
SparkSession.builder.appName("HeartAttackPrediction").getOrCreate()

# =========================
# Download the dataset
# =========================

url = "https://sp8138-heart-attack-dataset.s3.us-east-
2.amazonaws.com/heart_attack_data.csv"
local_file = "heart_attack_data.csv"
urllib.request.urlretrieve(url, local_file)

('heart_attack_data.csv', <http.client.HTTPMessage at 0x7c8fdf7beed0>)

# =========================
# Load and Preprocess Data
# =========================

# Load the dataset
```

```python
df = spark.read.csv("heart_attack_data.csv", header=True,
inferSchema=True)

# Drop rows with missing values
df = df.dropna()

# Display the schema of the dataset
df.printSchema()

root
 |-- Age: integer (nullable = true)
 |-- Gender: string (nullable = true)
 |-- Cholesterol: integer (nullable = true)
 |-- BloodPressure: integer (nullable = true)
 |-- HeartRate: integer (nullable = true)
 |-- BMI: double (nullable = true)
 |-- Smoker: integer (nullable = true)
 |-- Diabetes: integer (nullable = true)
 |-- Hypertension: integer (nullable = true)
 |-- FamilyHistory: integer (nullable = true)
 |-- PhysicalActivity: integer (nullable = true)
 |-- AlcoholConsumption: integer (nullable = true)
 |-- Diet: string (nullable = true)
 |-- StressLevel: integer (nullable = true)
 |-- Ethnicity: string (nullable = true)
 |-- Income: integer (nullable = true)
 |-- EducationLevel: string (nullable = true)
 |-- Medication: string (nullable = true)
 |-- ChestPainType: string (nullable = true)
 |-- ECGResults: string (nullable = true)
 |-- MaxHeartRate: integer (nullable = true)
 |-- ST_Depression: double (nullable = true)
 |-- ExerciseInducedAngina: string (nullable = true)
 |-- Slope: string (nullable = true)
 |-- NumberOfMajorVessels: integer (nullable = true)
 |-- Thalassemia: string (nullable = true)
 |-- PreviousHeartAttack: integer (nullable = true)
 |-- StrokeHistory: integer (nullable = true)
 |-- Residence: string (nullable = true)
 |-- EmploymentStatus: string (nullable = true)
 |-- MaritalStatus: string (nullable = true)
 |-- Outcome: string (nullable = true)


# =========================
# Select Features
# =========================

# Define the target column and selected features
target_col = "Outcome"
```

```python
selected_numeric = [
    "MaxHeartRate", "PhysicalActivity", "Hypertension",
"FamilyHistory", "StressLevel",
    "Smoker", "AlcoholConsumption", "Cholesterol", "StrokeHistory",
"ST_Depression",
    "NumberOfMajorVessels", "BMI", "PreviousHeartAttack", "HeartRate",
"Diabetes"
]
selected_categorical = [
    "Gender", "Diet", "Ethnicity", "EducationLevel", "Medication",
"ChestPainType",
    "ECGResults", "ExerciseInducedAngina", "Slope", "Thalassemia",
"Residence", "MaritalStatus"
]

categorical_cols = selected_categorical
numeric_cols = selected_numeric

# =========================
# Preprocessing Pipeline
# =========================

# StringIndexer for categorical columns (including target)
indexers = [StringIndexer(inputCol=col, outputCol=col+"_idx",
handleInvalid="keep") for col in categorical_cols + [target_col]]

# OneHotEncoder for categorical features (not target)
encoder = OneHotEncoder(inputCols=[col+"_idx" for col in
categorical_cols],
                        outputCols=[col+"_ohe" for col in
categorical_cols])

# Assemble features
feature_cols = [col+"_ohe" for col in categorical_cols] + numeric_cols
assembler = VectorAssembler(inputCols=feature_cols,
outputCol="features_unscaled")

# StandardScaler
scaler = StandardScaler(inputCol="features_unscaled",
outputCol="features", withMean=True, withStd=True)

# Create the pipeline
pipeline = Pipeline(stages=indexers + [encoder, assembler, scaler])
pipeline_model = pipeline.fit(df)

# Transform the dataset
df_prep = pipeline_model.transform(df)

# Select final features and label
```

```python
df_final = df_prep.select("features",
col("Outcome_idx").alias("label"))

# ==========================
# Train/Test Split
# ==========================

# Split the data into training and testing sets
train, test = df_final.randomSplit([0.8, 0.2], seed=42)

# Display the sizes of the splits
print(f"Training set size: {train.count()}")
print(f"Testing set size: {test.count()}")
```

```
Training set size: 298307
Testing set size: 74667
```

```python
# ==========================
# Train Models
# ==========================

# Logistic Regression
lr = LogisticRegression(featuresCol="features", labelCol="label",
maxIter=100)
lr_model = lr.fit(train)

# Random Forest
rf = RandomForestClassifier(featuresCol="features", labelCol="label",
numTrees=100, seed=42)
rf_model = rf.fit(train)

# Multilayer Perceptron (ANN)
input_size = train.first()['features'].size
layers = [input_size, 16, 8, 2]  # 2 output classes
mlp = MultilayerPerceptronClassifier(featuresCol="features",
labelCol="label", maxIter=100, layers=layers, seed=42)
mlp_model = mlp.fit(train)

# ==========================
# Save Models and Pipeline
# ==========================

# Create a directory for saving models
os.makedirs("saved_models", exist_ok=True)

# Remove existing directories if they exist
for path in [
    "saved_models/preprocessing_pipeline",
    "saved_models/logistic_regression",
    "saved_models/random_forest",
    "saved_models/ann_mlp"
```

```python
]:
    if os.path.exists(path):
        shutil.rmtree(path)

# Save the pipeline and models
pipeline_model.save("saved_models/preprocessing_pipeline")
lr_model.save("saved_models/logistic_regression")
rf_model.save("saved_models/random_forest")
mlp_model.save("saved_models/ann_mlp")

print("Models and pipeline saved in 'saved_models/' directory.")
```

```
Models and pipeline saved in 'saved_models/' directory.
```

```python
# =========================
# Evaluate Models
# =========================

# Function to evaluate a model
def evaluate_model(model, test, name):
    predictions = model.transform(test)
    evaluator = MulticlassClassificationEvaluator(labelCol="label",
predictionCol="prediction", metricName="accuracy")
    acc = evaluator.evaluate(predictions)
    print(f"\n{name} Accuracy: {(acc):.3f}")

# Evaluate all models
evaluate_model(lr_model, test, "Logistic Regression")
evaluate_model(rf_model, test, "Random Forest")
evaluate_model(mlp_model, test, "Multilayer Perceptron (ANN)")
```

```
Logistic Regression Accuracy: 0.801

Random Forest Accuracy: 0.800

Multilayer Perceptron (ANN) Accuracy: 0.800
```

```python
# =========================
# Load Models and Pipeline
# =========================

# Load the saved pipeline and models
loaded_pipeline =
PipelineModel.load("saved_models/preprocessing_pipeline")
loaded_lr =
LogisticRegressionModel.load("saved_models/logistic_regression")
loaded_rf =
RandomForestClassificationModel.load("saved_models/random_forest")
loaded_mlp =
MultilayerPerceptronClassificationModel.load("saved_models/ann_mlp")
```

```python
print("Models and pipeline loaded successfully.")

Models and pipeline loaded successfully.

# =========================
# Inference on New Case
# =========================

# Function to predict a new case
def predict_new_case(new_case_dict):
    new_df = spark.createDataFrame([new_case_dict])
    new_df_prep = loaded_pipeline.transform(new_df)
    new_df_final = new_df_prep.select("features")
    print("\n--- Inference on New Case ---")
    for name, model in [("Logistic Regression", loaded_lr), ("Random
Forest", loaded_rf), ("ANN", loaded_mlp)]:
        pred =
model.transform(new_df_final).select("prediction").collect()[0][0]
        outcome = "Heart Attack" if pred == 1.0 else "No Heart Attack"
        print(f"{name}: {outcome}")

# Example new case
new_case = {
    'Gender': 'Male',
    'Diet': 'Moderate',
    'Ethnicity': 'White',
    'EducationLevel': 'College',
    'Medication': 'Yes',
    'ChestPainType': 'Typical',
    'ECGResults': 'Normal',
    'ExerciseInducedAngina': 'No',
    'Slope': 'Flat',
    'Thalassemia': 'Normal',
    'Residence': 'Urban',
    'MaritalStatus': 'Married',
    'MaxHeartRate': 150,
    'PhysicalActivity': 2,
    'Hypertension': 1,
    'FamilyHistory': 1,
    'StressLevel': 2,
    'Smoker': 1,
    'AlcoholConsumption': 1,
    'Cholesterol': 200,
    'StrokeHistory': 0,
    'ST_Depression': 1.2,
    'NumberOfMajorVessels': 0,
    'BMI': 27.5,
    'PreviousHeartAttack': 0,
    'HeartRate': 80,
```

```python
    'Diabetes': 0
}
predict_new_case(new_case)
```

```
--- Inference on New Case ---
Logistic Regression: Heart Attack
Random Forest: No Heart Attack
ANN: Heart Attack
```

```python
# Zip the preprocessing pipeline
!zip -r preprocessing_pipeline.zip saved_models/preprocessing_pipeline

# Zip the random forest model
!zip -r random_forest.zip saved_models/random_forest
```

```
  adding: saved_models/preprocessing_pipeline/ (stored 0%)
  adding: saved_models/preprocessing_pipeline/stages/ (stored 0%)
  adding:
saved_models/preprocessing_pipeline/stages/08_StringIndexer_2995b821f4
a2/ (stored 0%)
  adding:
saved_models/preprocessing_pipeline/stages/08_StringIndexer_2995b821f4
a2/metadata/ (stored 0%)
  adding:
saved_models/preprocessing_pipeline/stages/08_StringIndexer_2995b821f4
a2/metadata/._SUCCESS.crc (stored 0%)
  adding:
saved_models/preprocessing_pipeline/stages/08_StringIndexer_2995b821f4
a2/metadata/.part-00000.crc (stored 0%)
  adding:
saved_models/preprocessing_pipeline/stages/08_StringIndexer_2995b821f4
a2/metadata/part-00000 (deflated 38%)
  adding:
saved_models/preprocessing_pipeline/stages/08_StringIndexer_2995b821f4
a2/metadata/_SUCCESS (stored 0%)
  adding:
saved_models/preprocessing_pipeline/stages/08_StringIndexer_2995b821f4
a2/data/ (stored 0%)
  adding:
saved_models/preprocessing_pipeline/stages/08_StringIndexer_2995b821f4
a2/data/.part-00000-f3bb3179-2210-4d33-a723-6e08ea5aa8aa-
c000.snappy.parquet.crc (stored 0%)
  adding:
saved_models/preprocessing_pipeline/stages/08_StringIndexer_2995b821f4
a2/data/._SUCCESS.crc (stored 0%)
  adding:
saved_models/preprocessing_pipeline/stages/08_StringIndexer_2995b821f4
```

```
a2/data/part-00000-f3bb3179-2210-4d33-a723-6e08ea5aa8aa-
c000.snappy.parquet (deflated 38%)
  adding:
saved_models/preprocessing_pipeline/stages/08_StringIndexer_2995b821f4
a2/data/_SUCCESS (stored 0%)
  adding:
saved_models/preprocessing_pipeline/stages/05_StringIndexer_d10f6c7407
22/ (stored 0%)
  adding:
saved_models/preprocessing_pipeline/stages/05_StringIndexer_d10f6c7407
22/metadata/ (stored 0%)
  adding:
saved_models/preprocessing_pipeline/stages/05_StringIndexer_d10f6c7407
22/metadata/._SUCCESS.crc (stored 0%)
  adding:
saved_models/preprocessing_pipeline/stages/05_StringIndexer_d10f6c7407
22/metadata/.part-00000.crc (stored 0%)
  adding:
saved_models/preprocessing_pipeline/stages/05_StringIndexer_d10f6c7407
22/metadata/part-00000 (deflated 40%)
  adding:
saved_models/preprocessing_pipeline/stages/05_StringIndexer_d10f6c7407
22/metadata/_SUCCESS (stored 0%)
  adding:
saved_models/preprocessing_pipeline/stages/05_StringIndexer_d10f6c7407
22/data/ (stored 0%)
  adding:
saved_models/preprocessing_pipeline/stages/05_StringIndexer_d10f6c7407
22/data/part-00000-9b4e9a84-7117-481e-b0a8-1272ec1ba435-
c000.snappy.parquet (deflated 36%)
  adding:
saved_models/preprocessing_pipeline/stages/05_StringIndexer_d10f6c7407
22/data/.part-00000-9b4e9a84-7117-481e-b0a8-1272ec1ba435-
c000.snappy.parquet.crc (stored 0%)
  adding:
saved_models/preprocessing_pipeline/stages/05_StringIndexer_d10f6c7407
22/data/._SUCCESS.crc (stored 0%)
  adding:
saved_models/preprocessing_pipeline/stages/05_StringIndexer_d10f6c7407
22/data/_SUCCESS (stored 0%)
  adding:
saved_models/preprocessing_pipeline/stages/10_StringIndexer_da353680d5
81/ (stored 0%)
  adding:
saved_models/preprocessing_pipeline/stages/10_StringIndexer_da353680d5
81/metadata/ (stored 0%)
  adding:
saved_models/preprocessing_pipeline/stages/10_StringIndexer_da353680d5
81/metadata/._SUCCESS.crc (stored 0%)
```

```
  adding:
saved_models/preprocessing_pipeline/stages/10_StringIndexer_da353680d5
81/metadata/.part-00000.crc (stored 0%)
  adding:
saved_models/preprocessing_pipeline/stages/10_StringIndexer_da353680d5
81/metadata/part-00000 (deflated 39%)
  adding:
saved_models/preprocessing_pipeline/stages/10_StringIndexer_da353680d5
81/metadata/_SUCCESS (stored 0%)
  adding:
saved_models/preprocessing_pipeline/stages/10_StringIndexer_da353680d5
81/data/ (stored 0%)
  adding:
saved_models/preprocessing_pipeline/stages/10_StringIndexer_da353680d5
81/data/part-00000-ab8360a7-49eb-4bf9-804b-86c13a7c7ab9-
c000.snappy.parquet (deflated 35%)
  adding:
saved_models/preprocessing_pipeline/stages/10_StringIndexer_da353680d5
81/data/._SUCCESS.crc (stored 0%)
  adding:
saved_models/preprocessing_pipeline/stages/10_StringIndexer_da353680d5
81/data/.part-00000-ab8360a7-49eb-4bf9-804b-86c13a7c7ab9-
c000.snappy.parquet.crc (stored 0%)
  adding:
saved_models/preprocessing_pipeline/stages/10_StringIndexer_da353680d5
81/data/_SUCCESS (stored 0%)
  adding:
saved_models/preprocessing_pipeline/stages/11_StringIndexer_1a1427dac2
d5/ (stored 0%)
  adding:
saved_models/preprocessing_pipeline/stages/11_StringIndexer_1a1427dac2
d5/metadata/ (stored 0%)
  adding:
saved_models/preprocessing_pipeline/stages/11_StringIndexer_1a1427dac2
d5/metadata/._SUCCESS.crc (stored 0%)
  adding:
saved_models/preprocessing_pipeline/stages/11_StringIndexer_1a1427dac2
d5/metadata/.part-00000.crc (stored 0%)
  adding:
saved_models/preprocessing_pipeline/stages/11_StringIndexer_1a1427dac2
d5/metadata/part-00000 (deflated 40%)
  adding:
saved_models/preprocessing_pipeline/stages/11_StringIndexer_1a1427dac2
d5/metadata/_SUCCESS (stored 0%)
  adding:
saved_models/preprocessing_pipeline/stages/11_StringIndexer_1a1427dac2
d5/data/ (stored 0%)
  adding:
saved_models/preprocessing_pipeline/stages/11_StringIndexer_1a1427dac2
```

```
d5/data/._SUCCESS.crc (stored 0%)
  adding:
saved_models/preprocessing_pipeline/stages/11_StringIndexer_1a1427dac2
d5/data/part-00000-a58520e8-230f-4f3c-9228-7eb450fdec46-
c000.snappy.parquet (deflated 36%)
  adding:
saved_models/preprocessing_pipeline/stages/11_StringIndexer_1a1427dac2
d5/data/.part-00000-a58520e8-230f-4f3c-9228-7eb450fdec46-
c000.snappy.parquet.crc (stored 0%)
  adding:
saved_models/preprocessing_pipeline/stages/11_StringIndexer_1a1427dac2
d5/data/_SUCCESS (stored 0%)
  adding:
saved_models/preprocessing_pipeline/stages/04_StringIndexer_afdd5bb570
2c/ (stored 0%)
  adding:
saved_models/preprocessing_pipeline/stages/04_StringIndexer_afdd5bb570
2c/metadata/ (stored 0%)
  adding:
saved_models/preprocessing_pipeline/stages/04_StringIndexer_afdd5bb570
2c/metadata/._SUCCESS.crc (stored 0%)
  adding:
saved_models/preprocessing_pipeline/stages/04_StringIndexer_afdd5bb570
2c/metadata/.part-00000.crc (stored 0%)
  adding:
saved_models/preprocessing_pipeline/stages/04_StringIndexer_afdd5bb570
2c/metadata/part-00000 (deflated 39%)
  adding:
saved_models/preprocessing_pipeline/stages/04_StringIndexer_afdd5bb570
2c/metadata/_SUCCESS (stored 0%)
  adding:
saved_models/preprocessing_pipeline/stages/04_StringIndexer_afdd5bb570
2c/data/ (stored 0%)
  adding:
saved_models/preprocessing_pipeline/stages/04_StringIndexer_afdd5bb570
2c/data/._SUCCESS.crc (stored 0%)
  adding:
saved_models/preprocessing_pipeline/stages/04_StringIndexer_afdd5bb570
2c/data/.part-00000-f85f1237-6d5a-48eb-974e-ba1a09026be8-
c000.snappy.parquet.crc (stored 0%)
  adding:
saved_models/preprocessing_pipeline/stages/04_StringIndexer_afdd5bb570
2c/data/_SUCCESS (stored 0%)
  adding:
saved_models/preprocessing_pipeline/stages/04_StringIndexer_afdd5bb570
2c/data/part-00000-f85f1237-6d5a-48eb-974e-ba1a09026be8-
c000.snappy.parquet (deflated 35%)
  adding:
saved_models/preprocessing_pipeline/stages/00_StringIndexer_1dd5e7eeba
```

```
a1/ (stored 0%)
  adding:
saved_models/preprocessing_pipeline/stages/00_StringIndexer_1dd5e7eeba
a1/metadata/ (stored 0%)
  adding:
saved_models/preprocessing_pipeline/stages/00_StringIndexer_1dd5e7eeba
a1/metadata/._SUCCESS.crc (stored 0%)
  adding:
saved_models/preprocessing_pipeline/stages/00_StringIndexer_1dd5e7eeba
a1/metadata/.part-00000.crc (stored 0%)
  adding:
saved_models/preprocessing_pipeline/stages/00_StringIndexer_1dd5e7eeba
a1/metadata/part-00000 (deflated 39%)
  adding:
saved_models/preprocessing_pipeline/stages/00_StringIndexer_1dd5e7eeba
a1/metadata/_SUCCESS (stored 0%)
  adding:
saved_models/preprocessing_pipeline/stages/00_StringIndexer_1dd5e7eeba
a1/data/ (stored 0%)
  adding:
saved_models/preprocessing_pipeline/stages/00_StringIndexer_1dd5e7eeba
a1/data/._SUCCESS.crc (stored 0%)
  adding:
saved_models/preprocessing_pipeline/stages/00_StringIndexer_1dd5e7eeba
a1/data/.part-00000-08e8421a-1568-48c8-a538-67b5717a98d8-
c000.snappy.parquet.crc (stored 0%)
  adding:
saved_models/preprocessing_pipeline/stages/00_StringIndexer_1dd5e7eeba
a1/data/_SUCCESS (stored 0%)
  adding:
saved_models/preprocessing_pipeline/stages/00_StringIndexer_1dd5e7eeba
a1/data/part-00000-08e8421a-1568-48c8-a538-67b5717a98d8-
c000.snappy.parquet (deflated 36%)
  adding:
saved_models/preprocessing_pipeline/stages/09_StringIndexer_26b0521e89
7b/ (stored 0%)
  adding:
saved_models/preprocessing_pipeline/stages/09_StringIndexer_26b0521e89
7b/metadata/ (stored 0%)
  adding:
saved_models/preprocessing_pipeline/stages/09_StringIndexer_26b0521e89
7b/metadata/._SUCCESS.crc (stored 0%)
  adding:
saved_models/preprocessing_pipeline/stages/09_StringIndexer_26b0521e89
7b/metadata/.part-00000.crc (stored 0%)
  adding:
saved_models/preprocessing_pipeline/stages/09_StringIndexer_26b0521e89
7b/metadata/part-00000 (deflated 39%)
  adding:
```

```
saved_models/preprocessing_pipeline/stages/09_StringIndexer_26b0521e89
7b/metadata/_SUCCESS (stored 0%)
  adding:
saved_models/preprocessing_pipeline/stages/09_StringIndexer_26b0521e89
7b/data/ (stored 0%)
  adding:
saved_models/preprocessing_pipeline/stages/09_StringIndexer_26b0521e89
7b/data/._SUCCESS.crc (stored 0%)
  adding:
saved_models/preprocessing_pipeline/stages/09_StringIndexer_26b0521e89
7b/data/part-00000-7ce6d08d-e0ae-4a0d-a177-9105690f1cdd-
c000.snappy.parquet (deflated 39%)
  adding:
saved_models/preprocessing_pipeline/stages/09_StringIndexer_26b0521e89
7b/data/.part-00000-7ce6d08d-e0ae-4a0d-a177-9105690f1cdd-
c000.snappy.parquet.crc (stored 0%)
  adding:
saved_models/preprocessing_pipeline/stages/09_StringIndexer_26b0521e89
7b/data/_SUCCESS (stored 0%)
  adding:
saved_models/preprocessing_pipeline/stages/15_StandardScaler_4e372ddab
86e/ (stored 0%)
  adding:
saved_models/preprocessing_pipeline/stages/15_StandardScaler_4e372ddab
86e/metadata/ (stored 0%)
  adding:
saved_models/preprocessing_pipeline/stages/15_StandardScaler_4e372ddab
86e/metadata/._SUCCESS.crc (stored 0%)
  adding:
saved_models/preprocessing_pipeline/stages/15_StandardScaler_4e372ddab
86e/metadata/.part-00000.crc (stored 0%)
  adding:
saved_models/preprocessing_pipeline/stages/15_StandardScaler_4e372ddab
86e/metadata/part-00000 (deflated 43%)
  adding:
saved_models/preprocessing_pipeline/stages/15_StandardScaler_4e372ddab
86e/metadata/_SUCCESS (stored 0%)
  adding:
saved_models/preprocessing_pipeline/stages/15_StandardScaler_4e372ddab
86e/data/ (stored 0%)
  adding:
saved_models/preprocessing_pipeline/stages/15_StandardScaler_4e372ddab
86e/data/part-00000-876d6125-bff3-4e3d-8488-dceebe2494a4-
c000.snappy.parquet (deflated 52%)
  adding:
saved_models/preprocessing_pipeline/stages/15_StandardScaler_4e372ddab
86e/data/._SUCCESS.crc (stored 0%)
  adding:
saved_models/preprocessing_pipeline/stages/15_StandardScaler_4e372ddab
```

```
86e/data/.part-00000-876d6125-bff3-4e3d-8488-dceebe2494a4-
c000.snappy.parquet.crc (stored 0%)
  adding:
saved_models/preprocessing_pipeline/stages/15_StandardScaler_4e372ddab
86e/data/_SUCCESS (stored 0%)
  adding:
saved_models/preprocessing_pipeline/stages/07_StringIndexer_d268d78651
9f/ (stored 0%)
  adding:
saved_models/preprocessing_pipeline/stages/07_StringIndexer_d268d78651
9f/metadata/ (stored 0%)
  adding:
saved_models/preprocessing_pipeline/stages/07_StringIndexer_d268d78651
9f/metadata/._SUCCESS.crc (stored 0%)
  adding:
saved_models/preprocessing_pipeline/stages/07_StringIndexer_d268d78651
9f/metadata/.part-00000.crc (stored 0%)
  adding:
saved_models/preprocessing_pipeline/stages/07_StringIndexer_d268d78651
9f/metadata/part-00000 (deflated 41%)
  adding:
saved_models/preprocessing_pipeline/stages/07_StringIndexer_d268d78651
9f/metadata/_SUCCESS (stored 0%)
  adding:
saved_models/preprocessing_pipeline/stages/07_StringIndexer_d268d78651
9f/data/ (stored 0%)
  adding:
saved_models/preprocessing_pipeline/stages/07_StringIndexer_d268d78651
9f/data/._SUCCESS.crc (stored 0%)
  adding:
saved_models/preprocessing_pipeline/stages/07_StringIndexer_d268d78651
9f/data/part-00000-e8e91440-0023-495d-ad16-9ef0c70c0e03-
c000.snappy.parquet (deflated 35%)
  adding:
saved_models/preprocessing_pipeline/stages/07_StringIndexer_d268d78651
9f/data/_SUCCESS (stored 0%)
  adding:
saved_models/preprocessing_pipeline/stages/07_StringIndexer_d268d78651
9f/data/.part-00000-e8e91440-0023-495d-ad16-9ef0c70c0e03-
c000.snappy.parquet.crc (stored 0%)
  adding:
saved_models/preprocessing_pipeline/stages/14_VectorAssembler_8c723eb6
3e86/ (stored 0%)
  adding:
saved_models/preprocessing_pipeline/stages/14_VectorAssembler_8c723eb6
3e86/metadata/ (stored 0%)
  adding:
saved_models/preprocessing_pipeline/stages/14_VectorAssembler_8c723eb6
3e86/metadata/._SUCCESS.crc (stored 0%)
```

```
  adding:
saved_models/preprocessing_pipeline/stages/14_VectorAssembler_8c723eb6
3e86/metadata/.part-00000.crc (stored 0%)
  adding:
saved_models/preprocessing_pipeline/stages/14_VectorAssembler_8c723eb6
3e86/metadata/part-00000 (deflated 42%)
  adding:
saved_models/preprocessing_pipeline/stages/14_VectorAssembler_8c723eb6
3e86/metadata/_SUCCESS (stored 0%)
  adding:
saved_models/preprocessing_pipeline/stages/12_StringIndexer_e6b137156f
46/ (stored 0%)
  adding:
saved_models/preprocessing_pipeline/stages/12_StringIndexer_e6b137156f
46/metadata/ (stored 0%)
  adding:
saved_models/preprocessing_pipeline/stages/12_StringIndexer_e6b137156f
46/metadata/._SUCCESS.crc (stored 0%)
  adding:
saved_models/preprocessing_pipeline/stages/12_StringIndexer_e6b137156f
46/metadata/.part-00000.crc (stored 0%)
  adding:
saved_models/preprocessing_pipeline/stages/12_StringIndexer_e6b137156f
46/metadata/part-00000 (deflated 39%)
  adding:
saved_models/preprocessing_pipeline/stages/12_StringIndexer_e6b137156f
46/metadata/_SUCCESS (stored 0%)
  adding:
saved_models/preprocessing_pipeline/stages/12_StringIndexer_e6b137156f
46/data/ (stored 0%)
  adding:
saved_models/preprocessing_pipeline/stages/12_StringIndexer_e6b137156f
46/data/part-00000-76a0b57e-711c-41f9-893a-737e84bc922a-
c000.snappy.parquet (deflated 39%)
  adding:
saved_models/preprocessing_pipeline/stages/12_StringIndexer_e6b137156f
46/data/.part-00000-76a0b57e-711c-41f9-893a-737e84bc922a-
c000.snappy.parquet.crc (stored 0%)
  adding:
saved_models/preprocessing_pipeline/stages/12_StringIndexer_e6b137156f
46/data/._SUCCESS.crc (stored 0%)
  adding:
saved_models/preprocessing_pipeline/stages/12_StringIndexer_e6b137156f
46/data/_SUCCESS (stored 0%)
  adding:
saved_models/preprocessing_pipeline/stages/06_StringIndexer_1b3e1eda27
43/ (stored 0%)
  adding:
saved_models/preprocessing_pipeline/stages/06_StringIndexer_1b3e1eda27
```

```
43/metadata/ (stored 0%)
  adding:
saved_models/preprocessing_pipeline/stages/06_StringIndexer_1b3e1eda27
43/metadata/._SUCCESS.crc (stored 0%)
  adding:
saved_models/preprocessing_pipeline/stages/06_StringIndexer_1b3e1eda27
43/metadata/.part-00000.crc (stored 0%)
  adding:
saved_models/preprocessing_pipeline/stages/06_StringIndexer_1b3e1eda27
43/metadata/part-00000 (deflated 40%)
  adding:
saved_models/preprocessing_pipeline/stages/06_StringIndexer_1b3e1eda27
43/metadata/_SUCCESS (stored 0%)
  adding:
saved_models/preprocessing_pipeline/stages/06_StringIndexer_1b3e1eda27
43/data/ (stored 0%)
  adding:
saved_models/preprocessing_pipeline/stages/06_StringIndexer_1b3e1eda27
43/data/._SUCCESS.crc (stored 0%)
  adding:
saved_models/preprocessing_pipeline/stages/06_StringIndexer_1b3e1eda27
43/data/.part-00000-97db43d9-f5f5-4a7f-91b2-8603bb114986-
c000.snappy.parquet.crc (stored 0%)
  adding:
saved_models/preprocessing_pipeline/stages/06_StringIndexer_1b3e1eda27
43/data/part-00000-97db43d9-f5f5-4a7f-91b2-8603bb114986-
c000.snappy.parquet (deflated 38%)
  adding:
saved_models/preprocessing_pipeline/stages/06_StringIndexer_1b3e1eda27
43/data/_SUCCESS (stored 0%)
  adding:
saved_models/preprocessing_pipeline/stages/13_OneHotEncoder_ac20700ca6
75/ (stored 0%)
  adding:
saved_models/preprocessing_pipeline/stages/13_OneHotEncoder_ac20700ca6
75/metadata/ (stored 0%)
  adding:
saved_models/preprocessing_pipeline/stages/13_OneHotEncoder_ac20700ca6
75/metadata/._SUCCESS.crc (stored 0%)
  adding:
saved_models/preprocessing_pipeline/stages/13_OneHotEncoder_ac20700ca6
75/metadata/.part-00000.crc (stored 0%)
  adding:
saved_models/preprocessing_pipeline/stages/13_OneHotEncoder_ac20700ca6
75/metadata/part-00000 (deflated 52%)
  adding:
saved_models/preprocessing_pipeline/stages/13_OneHotEncoder_ac20700ca6
75/metadata/_SUCCESS (stored 0%)
  adding:
```

```
  adding:
saved_models/preprocessing_pipeline/stages/13_OneHotEncoder_ac20700ca6
75/data/ (stored 0%)
  adding:
saved_models/preprocessing_pipeline/stages/13_OneHotEncoder_ac20700ca6
75/data/._SUCCESS.crc (stored 0%)
  adding:
saved_models/preprocessing_pipeline/stages/13_OneHotEncoder_ac20700ca6
75/data/part-00000-569f87e0-f972-4b58-80a7-4ff2c2ee6db3-
c000.snappy.parquet (deflated 27%)
  adding:
saved_models/preprocessing_pipeline/stages/13_OneHotEncoder_ac20700ca6
75/data/.part-00000-569f87e0-f972-4b58-80a7-4ff2c2ee6db3-
c000.snappy.parquet.crc (stored 0%)
  adding:
saved_models/preprocessing_pipeline/stages/13_OneHotEncoder_ac20700ca6
75/data/_SUCCESS (stored 0%)
  adding:
saved_models/preprocessing_pipeline/stages/02_StringIndexer_b77f1dca2d
7a/ (stored 0%)
  adding:
saved_models/preprocessing_pipeline/stages/02_StringIndexer_b77f1dca2d
7a/metadata/ (stored 0%)
  adding:
saved_models/preprocessing_pipeline/stages/02_StringIndexer_b77f1dca2d
7a/metadata/._SUCCESS.crc (stored 0%)
  adding:
saved_models/preprocessing_pipeline/stages/02_StringIndexer_b77f1dca2d
7a/metadata/.part-00000.crc (stored 0%)
  adding:
saved_models/preprocessing_pipeline/stages/02_StringIndexer_b77f1dca2d
7a/metadata/part-00000 (deflated 40%)
  adding:
saved_models/preprocessing_pipeline/stages/02_StringIndexer_b77f1dca2d
7a/metadata/_SUCCESS (stored 0%)
  adding:
saved_models/preprocessing_pipeline/stages/02_StringIndexer_b77f1dca2d
7a/data/ (stored 0%)
  adding:
saved_models/preprocessing_pipeline/stages/02_StringIndexer_b77f1dca2d
7a/data/part-00000-86029cb2-83a9-4033-8b2b-6ad4048c6fbd-
c000.snappy.parquet (deflated 35%)
  adding:
saved_models/preprocessing_pipeline/stages/02_StringIndexer_b77f1dca2d
7a/data/._SUCCESS.crc (stored 0%)
  adding:
saved_models/preprocessing_pipeline/stages/02_StringIndexer_b77f1dca2d
7a/data/.part-00000-86029cb2-83a9-4033-8b2b-6ad4048c6fbd-
c000.snappy.parquet.crc (stored 0%)
  adding:
```

```
saved_models/preprocessing_pipeline/stages/02_StringIndexer_b77f1dca2d
7a/data/_SUCCESS (stored 0%)
  adding:
saved_models/preprocessing_pipeline/stages/03_StringIndexer_41841b909e
a2/ (stored 0%)
  adding:
saved_models/preprocessing_pipeline/stages/03_StringIndexer_41841b909e
a2/metadata/ (stored 0%)
  adding:
saved_models/preprocessing_pipeline/stages/03_StringIndexer_41841b909e
a2/metadata/._SUCCESS.crc (stored 0%)
  adding:
saved_models/preprocessing_pipeline/stages/03_StringIndexer_41841b909e
a2/metadata/.part-00000.crc (stored 0%)
  adding:
saved_models/preprocessing_pipeline/stages/03_StringIndexer_41841b909e
a2/metadata/part-00000 (deflated 40%)
  adding:
saved_models/preprocessing_pipeline/stages/03_StringIndexer_41841b909e
a2/metadata/_SUCCESS (stored 0%)
  adding:
saved_models/preprocessing_pipeline/stages/03_StringIndexer_41841b909e
a2/data/ (stored 0%)
  adding:
saved_models/preprocessing_pipeline/stages/03_StringIndexer_41841b909e
a2/data/part-00000-279fe6d3-3884-4b5e-b5af-7cc5e485ae59-
c000.snappy.parquet (deflated 36%)
  adding:
saved_models/preprocessing_pipeline/stages/03_StringIndexer_41841b909e
a2/data/._SUCCESS.crc (stored 0%)
  adding:
saved_models/preprocessing_pipeline/stages/03_StringIndexer_41841b909e
a2/data/.part-00000-279fe6d3-3884-4b5e-b5af-7cc5e485ae59-
c000.snappy.parquet.crc (stored 0%)
  adding:
saved_models/preprocessing_pipeline/stages/03_StringIndexer_41841b909e
a2/data/_SUCCESS (stored 0%)
  adding:
saved_models/preprocessing_pipeline/stages/01_StringIndexer_4896fc3942
5a/ (stored 0%)
  adding:
saved_models/preprocessing_pipeline/stages/01_StringIndexer_4896fc3942
5a/metadata/ (stored 0%)
  adding:
saved_models/preprocessing_pipeline/stages/01_StringIndexer_4896fc3942
5a/metadata/._SUCCESS.crc (stored 0%)
  adding:
saved_models/preprocessing_pipeline/stages/01_StringIndexer_4896fc3942
5a/metadata/.part-00000.crc (stored 0%)
  adding:
```

```
  saved_models/preprocessing_pipeline/stages/01_StringIndexer_4896fc3942
5a/metadata/part-00000 (deflated 38%)
  adding:
saved_models/preprocessing_pipeline/stages/01_StringIndexer_4896fc3942
5a/metadata/_SUCCESS (stored 0%)
  adding:
saved_models/preprocessing_pipeline/stages/01_StringIndexer_4896fc3942
5a/data/ (stored 0%)
  adding:
saved_models/preprocessing_pipeline/stages/01_StringIndexer_4896fc3942
5a/data/part-00000-60c89ab9-8394-476b-8b3e-53c1f1fb8e14-
c000.snappy.parquet (deflated 37%)
  adding:
saved_models/preprocessing_pipeline/stages/01_StringIndexer_4896fc3942
5a/data/._SUCCESS.crc (stored 0%)
  adding:
saved_models/preprocessing_pipeline/stages/01_StringIndexer_4896fc3942
5a/data/_SUCCESS (stored 0%)
  adding:
saved_models/preprocessing_pipeline/stages/01_StringIndexer_4896fc3942
5a/data/.part-00000-60c89ab9-8394-476b-8b3e-53c1f1fb8e14-
c000.snappy.parquet.crc (stored 0%)
  adding: saved_models/preprocessing_pipeline/metadata/ (stored 0%)
  adding: saved_models/preprocessing_pipeline/metadata/._SUCCESS.crc
(stored 0%)
  adding: saved_models/preprocessing_pipeline/metadata/.part-00000.crc
(stored 0%)
  adding: saved_models/preprocessing_pipeline/metadata/part-00000
(deflated 48%)
  adding: saved_models/preprocessing_pipeline/metadata/_SUCCESS
(stored 0%)
  adding: saved_models/random_forest/ (stored 0%)
  adding: saved_models/random_forest/metadata/ (stored 0%)
  adding: saved_models/random_forest/metadata/._SUCCESS.crc (stored
0%)
  adding: saved_models/random_forest/metadata/.part-00000.crc (stored
0%)
  adding: saved_models/random_forest/metadata/part-00000 (deflated
46%)
  adding: saved_models/random_forest/metadata/_SUCCESS (stored 0%)
  adding: saved_models/random_forest/data/ (stored 0%)
  adding: saved_models/random_forest/data/part-00000-3e03ad8a-7e52-
4f4f-8a38-5cb10a2b37c3-c000.snappy.parquet (deflated 16%)
  adding: saved_models/random_forest/data/.part-00000-3e03ad8a-7e52-
4f4f-8a38-5cb10a2b37c3-c000.snappy.parquet.crc (stored 0%)
  adding: saved_models/random_forest/data/._SUCCESS.crc (stored 0%)
  adding: saved_models/random_forest/data/_SUCCESS (stored 0%)
  adding: saved_models/random_forest/treesMetadata/ (stored 0%)
  adding: saved_models/random_forest/treesMetadata/.part-00000-
```

```
4966da8d-bba1-47fc-930d-525798c559f5-c000.snappy.parquet.crc (stored
0%)
  adding: saved_models/random_forest/treesMetadata/._SUCCESS.crc
(stored 0%)
  adding: saved_models/random_forest/treesMetadata/part-00000-
4966da8d-bba1-47fc-930d-525798c559f5-c000.snappy.parquet (deflated
28%)
  adding: saved_models/random_forest/treesMetadata/_SUCCESS (stored
0%)
```

`!zip -r saved_models.zip saved_models`

```
  adding: saved_models/ (stored 0%)
  adding: saved_models/ann_mlp/ (stored 0%)
  adding: saved_models/ann_mlp/metadata/ (stored 0%)
  adding: saved_models/ann_mlp/metadata/._SUCCESS.crc (stored 0%)
  adding: saved_models/ann_mlp/metadata/.part-00000.crc (stored 0%)
  adding: saved_models/ann_mlp/metadata/part-00000 (deflated 46%)
  adding: saved_models/ann_mlp/metadata/_SUCCESS (stored 0%)
  adding: saved_models/ann_mlp/data/ (stored 0%)
  adding: saved_models/ann_mlp/data/.part-00000-7f847ddb-4f7c-4c91-
8133-e08ad3ac44a7-c000.snappy.parquet.crc (stored 0%)
  adding: saved_models/ann_mlp/data/._SUCCESS.crc (stored 0%)
  adding: saved_models/ann_mlp/data/_SUCCESS (stored 0%)
  adding: saved_models/ann_mlp/data/part-00000-7f847ddb-4f7c-4c91-
8133-e08ad3ac44a7-c000.snappy.parquet (deflated 11%)
  adding: saved_models/logistic_regression/ (stored 0%)
  adding: saved_models/logistic_regression/metadata/ (stored 0%)
  adding: saved_models/logistic_regression/metadata/._SUCCESS.crc
(stored 0%)
  adding: saved_models/logistic_regression/metadata/.part-00000.crc
(stored 0%)
  adding: saved_models/logistic_regression/metadata/part-00000
(deflated 45%)
  adding: saved_models/logistic_regression/metadata/_SUCCESS (stored
0%)
  adding: saved_models/logistic_regression/data/ (stored 0%)
  adding: saved_models/logistic_regression/data/._SUCCESS.crc (stored
0%)
  adding: saved_models/logistic_regression/data/.part-00000-1ab2c794-
ad66-4334-b1cb-806e9886170b-c000.snappy.parquet.crc (stored 0%)
  adding: saved_models/logistic_regression/data/part-00000-1ab2c794-
ad66-4334-b1cb-806e9886170b-c000.snappy.parquet (deflated 54%)
  adding: saved_models/logistic_regression/data/_SUCCESS (stored 0%)
  adding: saved_models/random_forest/ (stored 0%)
  adding: saved_models/random_forest/metadata/ (stored 0%)
  adding: saved_models/random_forest/metadata/._SUCCESS.crc (stored
0%)
```

```
  adding: saved_models/random_forest/metadata/.part-00000.crc (stored
0%)
  adding: saved_models/random_forest/metadata/part-00000 (deflated
46%)
  adding: saved_models/random_forest/metadata/_SUCCESS (stored 0%)
  adding: saved_models/random_forest/data/ (stored 0%)
  adding: saved_models/random_forest/data/part-00000-3e03ad8a-7e52-
4f4f-8a38-5cb10a2b37c3-c000.snappy.parquet (deflated 16%)
  adding: saved_models/random_forest/data/.part-00000-3e03ad8a-7e52-
4f4f-8a38-5cb10a2b37c3-c000.snappy.parquet.crc (stored 0%)
  adding: saved_models/random_forest/data/._SUCCESS.crc (stored 0%)
  adding: saved_models/random_forest/data/_SUCCESS (stored 0%)
  adding: saved_models/random_forest/treesMetadata/ (stored 0%)
  adding: saved_models/random_forest/treesMetadata/.part-00000-
4966da8d-bba1-47fc-930d-525798c559f5-c000.snappy.parquet.crc (stored
0%)
  adding: saved_models/random_forest/treesMetadata/._SUCCESS.crc
(stored 0%)
  adding: saved_models/random_forest/treesMetadata/part-00000-
4966da8d-bba1-47fc-930d-525798c559f5-c000.snappy.parquet (deflated
28%)
  adding: saved_models/random_forest/treesMetadata/_SUCCESS (stored
0%)
  adding: saved_models/preprocessing_pipeline/ (stored 0%)
  adding: saved_models/preprocessing_pipeline/stages/ (stored 0%)
  adding:
saved_models/preprocessing_pipeline/stages/08_StringIndexer_2995b821f4
a2/ (stored 0%)
  adding:
saved_models/preprocessing_pipeline/stages/08_StringIndexer_2995b821f4
a2/metadata/ (stored 0%)
  adding:
saved_models/preprocessing_pipeline/stages/08_StringIndexer_2995b821f4
a2/metadata/._SUCCESS.crc (stored 0%)
  adding:
saved_models/preprocessing_pipeline/stages/08_StringIndexer_2995b821f4
a2/metadata/.part-00000.crc (stored 0%)
  adding:
saved_models/preprocessing_pipeline/stages/08_StringIndexer_2995b821f4
a2/metadata/part-00000 (deflated 38%)
  adding:
saved_models/preprocessing_pipeline/stages/08_StringIndexer_2995b821f4
a2/metadata/_SUCCESS (stored 0%)
  adding:
saved_models/preprocessing_pipeline/stages/08_StringIndexer_2995b821f4
a2/data/ (stored 0%)
  adding:
saved_models/preprocessing_pipeline/stages/08_StringIndexer_2995b821f4
a2/data/.part-00000-f3bb3179-2210-4d33-a723-6e08ea5aa8aa-
```

```
c000.snappy.parquet.crc (stored 0%)
  adding:
saved_models/preprocessing_pipeline/stages/08_StringIndexer_2995b821f4
a2/data/._SUCCESS.crc (stored 0%)
  adding:
saved_models/preprocessing_pipeline/stages/08_StringIndexer_2995b821f4
a2/data/part-00000-f3bb3179-2210-4d33-a723-6e08ea5aa8aa-
c000.snappy.parquet (deflated 38%)
  adding:
saved_models/preprocessing_pipeline/stages/08_StringIndexer_2995b821f4
a2/data/_SUCCESS (stored 0%)
  adding:
saved_models/preprocessing_pipeline/stages/05_StringIndexer_d10f6c7407
22/ (stored 0%)
  adding:
saved_models/preprocessing_pipeline/stages/05_StringIndexer_d10f6c7407
22/metadata/ (stored 0%)
  adding:
saved_models/preprocessing_pipeline/stages/05_StringIndexer_d10f6c7407
22/metadata/._SUCCESS.crc (stored 0%)
  adding:
saved_models/preprocessing_pipeline/stages/05_StringIndexer_d10f6c7407
22/metadata/.part-00000.crc (stored 0%)
  adding:
saved_models/preprocessing_pipeline/stages/05_StringIndexer_d10f6c7407
22/metadata/part-00000 (deflated 40%)
  adding:
saved_models/preprocessing_pipeline/stages/05_StringIndexer_d10f6c7407
22/metadata/_SUCCESS (stored 0%)
  adding:
saved_models/preprocessing_pipeline/stages/05_StringIndexer_d10f6c7407
22/data/ (stored 0%)
  adding:
saved_models/preprocessing_pipeline/stages/05_StringIndexer_d10f6c7407
22/data/part-00000-9b4e9a84-7117-481e-b0a8-1272ec1ba435-
c000.snappy.parquet (deflated 36%)
  adding:
saved_models/preprocessing_pipeline/stages/05_StringIndexer_d10f6c7407
22/data/.part-00000-9b4e9a84-7117-481e-b0a8-1272ec1ba435-
c000.snappy.parquet.crc (stored 0%)
  adding:
saved_models/preprocessing_pipeline/stages/05_StringIndexer_d10f6c7407
22/data/._SUCCESS.crc (stored 0%)
  adding:
saved_models/preprocessing_pipeline/stages/05_StringIndexer_d10f6c7407
22/data/_SUCCESS (stored 0%)
  adding:
saved_models/preprocessing_pipeline/stages/10_StringIndexer_da353680d5
81/ (stored 0%)
```

```
  adding:
saved_models/preprocessing_pipeline/stages/10_StringIndexer_da353680d5
81/metadata/ (stored 0%)
  adding:
saved_models/preprocessing_pipeline/stages/10_StringIndexer_da353680d5
81/metadata/._SUCCESS.crc (stored 0%)
  adding:
saved_models/preprocessing_pipeline/stages/10_StringIndexer_da353680d5
81/metadata/.part-00000.crc (stored 0%)
  adding:
saved_models/preprocessing_pipeline/stages/10_StringIndexer_da353680d5
81/metadata/part-00000 (deflated 39%)
  adding:
saved_models/preprocessing_pipeline/stages/10_StringIndexer_da353680d5
81/metadata/_SUCCESS (stored 0%)
  adding:
saved_models/preprocessing_pipeline/stages/10_StringIndexer_da353680d5
81/data/ (stored 0%)
  adding:
saved_models/preprocessing_pipeline/stages/10_StringIndexer_da353680d5
81/data/part-00000-ab8360a7-49eb-4bf9-804b-86c13a7c7ab9-
c000.snappy.parquet (deflated 35%)
  adding:
saved_models/preprocessing_pipeline/stages/10_StringIndexer_da353680d5
81/data/._SUCCESS.crc (stored 0%)
  adding:
saved_models/preprocessing_pipeline/stages/10_StringIndexer_da353680d5
81/data/.part-00000-ab8360a7-49eb-4bf9-804b-86c13a7c7ab9-
c000.snappy.parquet.crc (stored 0%)
  adding:
saved_models/preprocessing_pipeline/stages/10_StringIndexer_da353680d5
81/data/_SUCCESS (stored 0%)
  adding:
saved_models/preprocessing_pipeline/stages/11_StringIndexer_1a1427dac2
d5/ (stored 0%)
  adding:
saved_models/preprocessing_pipeline/stages/11_StringIndexer_1a1427dac2
d5/metadata/ (stored 0%)
  adding:
saved_models/preprocessing_pipeline/stages/11_StringIndexer_1a1427dac2
d5/metadata/._SUCCESS.crc (stored 0%)
  adding:
saved_models/preprocessing_pipeline/stages/11_StringIndexer_1a1427dac2
d5/metadata/.part-00000.crc (stored 0%)
  adding:
saved_models/preprocessing_pipeline/stages/11_StringIndexer_1a1427dac2
d5/metadata/part-00000 (deflated 40%)
  adding:
saved_models/preprocessing_pipeline/stages/11_StringIndexer_1a1427dac2
```

```
d5/metadata/_SUCCESS (stored 0%)
  adding:
saved_models/preprocessing_pipeline/stages/11_StringIndexer_1a1427dac2
d5/data/ (stored 0%)
  adding:
saved_models/preprocessing_pipeline/stages/11_StringIndexer_1a1427dac2
d5/data/._SUCCESS.crc (stored 0%)
  adding:
saved_models/preprocessing_pipeline/stages/11_StringIndexer_1a1427dac2
d5/data/part-00000-a58520e8-230f-4f3c-9228-7eb450fdec46-
c000.snappy.parquet (deflated 36%)
  adding:
saved_models/preprocessing_pipeline/stages/11_StringIndexer_1a1427dac2
d5/data/.part-00000-a58520e8-230f-4f3c-9228-7eb450fdec46-
c000.snappy.parquet.crc (stored 0%)
  adding:
saved_models/preprocessing_pipeline/stages/11_StringIndexer_1a1427dac2
d5/data/_SUCCESS (stored 0%)
  adding:
saved_models/preprocessing_pipeline/stages/04_StringIndexer_afdd5bb570
2c/ (stored 0%)
  adding:
saved_models/preprocessing_pipeline/stages/04_StringIndexer_afdd5bb570
2c/metadata/ (stored 0%)
  adding:
saved_models/preprocessing_pipeline/stages/04_StringIndexer_afdd5bb570
2c/metadata/._SUCCESS.crc (stored 0%)
  adding:
saved_models/preprocessing_pipeline/stages/04_StringIndexer_afdd5bb570
2c/metadata/.part-00000.crc (stored 0%)
  adding:
saved_models/preprocessing_pipeline/stages/04_StringIndexer_afdd5bb570
2c/metadata/part-00000 (deflated 39%)
  adding:
saved_models/preprocessing_pipeline/stages/04_StringIndexer_afdd5bb570
2c/metadata/_SUCCESS (stored 0%)
  adding:
saved_models/preprocessing_pipeline/stages/04_StringIndexer_afdd5bb570
2c/data/ (stored 0%)
  adding:
saved_models/preprocessing_pipeline/stages/04_StringIndexer_afdd5bb570
2c/data/._SUCCESS.crc (stored 0%)
  adding:
saved_models/preprocessing_pipeline/stages/04_StringIndexer_afdd5bb570
2c/data/.part-00000-f85f1237-6d5a-48eb-974e-ba1a09026be8-
c000.snappy.parquet.crc (stored 0%)
  adding:
saved_models/preprocessing_pipeline/stages/04_StringIndexer_afdd5bb570
2c/data/_SUCCESS (stored 0%)
```

```
  adding:
saved_models/preprocessing_pipeline/stages/04_StringIndexer_afdd5bb570
2c/data/part-00000-f85f1237-6d5a-48eb-974e-ba1a09026be8-
c000.snappy.parquet (deflated 35%)
  adding:
saved_models/preprocessing_pipeline/stages/00_StringIndexer_1dd5e7eeba
a1/ (stored 0%)
  adding:
saved_models/preprocessing_pipeline/stages/00_StringIndexer_1dd5e7eeba
a1/metadata/ (stored 0%)
  adding:
saved_models/preprocessing_pipeline/stages/00_StringIndexer_1dd5e7eeba
a1/metadata/._SUCCESS.crc (stored 0%)
  adding:
saved_models/preprocessing_pipeline/stages/00_StringIndexer_1dd5e7eeba
a1/metadata/.part-00000.crc (stored 0%)
  adding:
saved_models/preprocessing_pipeline/stages/00_StringIndexer_1dd5e7eeba
a1/metadata/part-00000 (deflated 39%)
  adding:
saved_models/preprocessing_pipeline/stages/00_StringIndexer_1dd5e7eeba
a1/metadata/_SUCCESS (stored 0%)
  adding:
saved_models/preprocessing_pipeline/stages/00_StringIndexer_1dd5e7eeba
a1/data/ (stored 0%)
  adding:
saved_models/preprocessing_pipeline/stages/00_StringIndexer_1dd5e7eeba
a1/data/._SUCCESS.crc (stored 0%)
  adding:
saved_models/preprocessing_pipeline/stages/00_StringIndexer_1dd5e7eeba
a1/data/.part-00000-08e8421a-1568-48c8-a538-67b5717a98d8-
c000.snappy.parquet.crc (stored 0%)
  adding:
saved_models/preprocessing_pipeline/stages/00_StringIndexer_1dd5e7eeba
a1/data/_SUCCESS (stored 0%)
  adding:
saved_models/preprocessing_pipeline/stages/00_StringIndexer_1dd5e7eeba
a1/data/part-00000-08e8421a-1568-48c8-a538-67b5717a98d8-
c000.snappy.parquet (deflated 36%)
  adding:
saved_models/preprocessing_pipeline/stages/09_StringIndexer_26b0521e89
7b/ (stored 0%)
  adding:
saved_models/preprocessing_pipeline/stages/09_StringIndexer_26b0521e89
7b/metadata/ (stored 0%)
  adding:
saved_models/preprocessing_pipeline/stages/09_StringIndexer_26b0521e89
7b/metadata/._SUCCESS.crc (stored 0%)
  adding:
```

```
  adding:
saved_models/preprocessing_pipeline/stages/09_StringIndexer_26b0521e89
7b/metadata/.part-00000.crc (stored 0%)
  adding:
saved_models/preprocessing_pipeline/stages/09_StringIndexer_26b0521e89
7b/metadata/part-00000 (deflated 39%)
  adding:
saved_models/preprocessing_pipeline/stages/09_StringIndexer_26b0521e89
7b/metadata/_SUCCESS (stored 0%)
  adding:
saved_models/preprocessing_pipeline/stages/09_StringIndexer_26b0521e89
7b/data/ (stored 0%)
  adding:
saved_models/preprocessing_pipeline/stages/09_StringIndexer_26b0521e89
7b/data/._SUCCESS.crc (stored 0%)
  adding:
saved_models/preprocessing_pipeline/stages/09_StringIndexer_26b0521e89
7b/data/part-00000-7ce6d08d-e0ae-4a0d-a177-9105690f1cdd-
c000.snappy.parquet (deflated 39%)
  adding:
saved_models/preprocessing_pipeline/stages/09_StringIndexer_26b0521e89
7b/data/.part-00000-7ce6d08d-e0ae-4a0d-a177-9105690f1cdd-
c000.snappy.parquet.crc (stored 0%)
  adding:
saved_models/preprocessing_pipeline/stages/09_StringIndexer_26b0521e89
7b/data/_SUCCESS (stored 0%)
  adding:
saved_models/preprocessing_pipeline/stages/15_StandardScaler_4e372ddab
86e/ (stored 0%)
  adding:
saved_models/preprocessing_pipeline/stages/15_StandardScaler_4e372ddab
86e/metadata/ (stored 0%)
  adding:
saved_models/preprocessing_pipeline/stages/15_StandardScaler_4e372ddab
86e/metadata/._SUCCESS.crc (stored 0%)
  adding:
saved_models/preprocessing_pipeline/stages/15_StandardScaler_4e372ddab
86e/metadata/.part-00000.crc (stored 0%)
  adding:
saved_models/preprocessing_pipeline/stages/15_StandardScaler_4e372ddab
86e/metadata/part-00000 (deflated 43%)
  adding:
saved_models/preprocessing_pipeline/stages/15_StandardScaler_4e372ddab
86e/metadata/_SUCCESS (stored 0%)
  adding:
saved_models/preprocessing_pipeline/stages/15_StandardScaler_4e372ddab
86e/data/ (stored 0%)
  adding:
saved_models/preprocessing_pipeline/stages/15_StandardScaler_4e372ddab
86e/data/part-00000-876d6125-bff3-4e3d-8488-dceebe2494a4-
```

```
c000.snappy.parquet (deflated 52%)
  adding:
saved_models/preprocessing_pipeline/stages/15_StandardScaler_4e372ddab
86e/data/._SUCCESS.crc (stored 0%)
  adding:
saved_models/preprocessing_pipeline/stages/15_StandardScaler_4e372ddab
86e/data/.part-00000-876d6125-bff3-4e3d-8488-dceebe2494a4-
c000.snappy.parquet.crc (stored 0%)
  adding:
saved_models/preprocessing_pipeline/stages/15_StandardScaler_4e372ddab
86e/data/_SUCCESS (stored 0%)
  adding:
saved_models/preprocessing_pipeline/stages/07_StringIndexer_d268d78651
9f/ (stored 0%)
  adding:
saved_models/preprocessing_pipeline/stages/07_StringIndexer_d268d78651
9f/metadata/ (stored 0%)
  adding:
saved_models/preprocessing_pipeline/stages/07_StringIndexer_d268d78651
9f/metadata/._SUCCESS.crc (stored 0%)
  adding:
saved_models/preprocessing_pipeline/stages/07_StringIndexer_d268d78651
9f/metadata/.part-00000.crc (stored 0%)
  adding:
saved_models/preprocessing_pipeline/stages/07_StringIndexer_d268d78651
9f/metadata/part-00000 (deflated 41%)
  adding:
saved_models/preprocessing_pipeline/stages/07_StringIndexer_d268d78651
9f/metadata/_SUCCESS (stored 0%)
  adding:
saved_models/preprocessing_pipeline/stages/07_StringIndexer_d268d78651
9f/data/ (stored 0%)
  adding:
saved_models/preprocessing_pipeline/stages/07_StringIndexer_d268d78651
9f/data/._SUCCESS.crc (stored 0%)
  adding:
saved_models/preprocessing_pipeline/stages/07_StringIndexer_d268d78651
9f/data/part-00000-e8e91440-0023-495d-ad16-9ef0c70c0e03-
c000.snappy.parquet (deflated 35%)
  adding:
saved_models/preprocessing_pipeline/stages/07_StringIndexer_d268d78651
9f/data/_SUCCESS (stored 0%)
  adding:
saved_models/preprocessing_pipeline/stages/07_StringIndexer_d268d78651
9f/data/.part-00000-e8e91440-0023-495d-ad16-9ef0c70c0e03-
c000.snappy.parquet.crc (stored 0%)
  adding:
saved_models/preprocessing_pipeline/stages/14_VectorAssembler_8c723eb6
3e86/ (stored 0%)
```

```
  adding:
saved_models/preprocessing_pipeline/stages/14_VectorAssembler_8c723eb6
3e86/metadata/ (stored 0%)
  adding:
saved_models/preprocessing_pipeline/stages/14_VectorAssembler_8c723eb6
3e86/metadata/._SUCCESS.crc (stored 0%)
  adding:
saved_models/preprocessing_pipeline/stages/14_VectorAssembler_8c723eb6
3e86/metadata/.part-00000.crc (stored 0%)
  adding:
saved_models/preprocessing_pipeline/stages/14_VectorAssembler_8c723eb6
3e86/metadata/part-00000 (deflated 42%)
  adding:
saved_models/preprocessing_pipeline/stages/14_VectorAssembler_8c723eb6
3e86/metadata/_SUCCESS (stored 0%)
  adding:
saved_models/preprocessing_pipeline/stages/12_StringIndexer_e6b137156f
46/ (stored 0%)
  adding:
saved_models/preprocessing_pipeline/stages/12_StringIndexer_e6b137156f
46/metadata/ (stored 0%)
  adding:
saved_models/preprocessing_pipeline/stages/12_StringIndexer_e6b137156f
46/metadata/._SUCCESS.crc (stored 0%)
  adding:
saved_models/preprocessing_pipeline/stages/12_StringIndexer_e6b137156f
46/metadata/.part-00000.crc (stored 0%)
  adding:
saved_models/preprocessing_pipeline/stages/12_StringIndexer_e6b137156f
46/metadata/part-00000 (deflated 39%)
  adding:
saved_models/preprocessing_pipeline/stages/12_StringIndexer_e6b137156f
46/metadata/_SUCCESS (stored 0%)
  adding:
saved_models/preprocessing_pipeline/stages/12_StringIndexer_e6b137156f
46/data/ (stored 0%)
  adding:
saved_models/preprocessing_pipeline/stages/12_StringIndexer_e6b137156f
46/data/part-00000-76a0b57e-711c-41f9-893a-737e84bc922a-
c000.snappy.parquet (deflated 39%)
  adding:
saved_models/preprocessing_pipeline/stages/12_StringIndexer_e6b137156f
46/data/.part-00000-76a0b57e-711c-41f9-893a-737e84bc922a-
c000.snappy.parquet.crc (stored 0%)
  adding:
saved_models/preprocessing_pipeline/stages/12_StringIndexer_e6b137156f
46/data/._SUCCESS.crc (stored 0%)
  adding:
saved_models/preprocessing_pipeline/stages/12_StringIndexer_e6b137156f
```

```
46/data/_SUCCESS (stored 0%)
  adding:
saved_models/preprocessing_pipeline/stages/06_StringIndexer_1b3e1eda27
43/ (stored 0%)
  adding:
saved_models/preprocessing_pipeline/stages/06_StringIndexer_1b3e1eda27
43/metadata/ (stored 0%)
  adding:
saved_models/preprocessing_pipeline/stages/06_StringIndexer_1b3e1eda27
43/metadata/._SUCCESS.crc (stored 0%)
  adding:
saved_models/preprocessing_pipeline/stages/06_StringIndexer_1b3e1eda27
43/metadata/.part-00000.crc (stored 0%)
  adding:
saved_models/preprocessing_pipeline/stages/06_StringIndexer_1b3e1eda27
43/metadata/part-00000 (deflated 40%)
  adding:
saved_models/preprocessing_pipeline/stages/06_StringIndexer_1b3e1eda27
43/metadata/_SUCCESS (stored 0%)
  adding:
saved_models/preprocessing_pipeline/stages/06_StringIndexer_1b3e1eda27
43/data/ (stored 0%)
  adding:
saved_models/preprocessing_pipeline/stages/06_StringIndexer_1b3e1eda27
43/data/._SUCCESS.crc (stored 0%)
  adding:
saved_models/preprocessing_pipeline/stages/06_StringIndexer_1b3e1eda27
43/data/.part-00000-97db43d9-f5f5-4a7f-91b2-8603bb114986-
c000.snappy.parquet.crc (stored 0%)
  adding:
saved_models/preprocessing_pipeline/stages/06_StringIndexer_1b3e1eda27
43/data/part-00000-97db43d9-f5f5-4a7f-91b2-8603bb114986-
c000.snappy.parquet (deflated 38%)
  adding:
saved_models/preprocessing_pipeline/stages/06_StringIndexer_1b3e1eda27
43/data/_SUCCESS (stored 0%)
  adding:
saved_models/preprocessing_pipeline/stages/13_OneHotEncoder_ac20700ca6
75/ (stored 0%)
  adding:
saved_models/preprocessing_pipeline/stages/13_OneHotEncoder_ac20700ca6
75/metadata/ (stored 0%)
  adding:
saved_models/preprocessing_pipeline/stages/13_OneHotEncoder_ac20700ca6
75/metadata/._SUCCESS.crc (stored 0%)
  adding:
saved_models/preprocessing_pipeline/stages/13_OneHotEncoder_ac20700ca6
75/metadata/.part-00000.crc (stored 0%)
  adding:
```

```
  saved_models/preprocessing_pipeline/stages/13_OneHotEncoder_ac20700ca6
75/metadata/part-00000 (deflated 52%)
  adding:
saved_models/preprocessing_pipeline/stages/13_OneHotEncoder_ac20700ca6
75/metadata/_SUCCESS (stored 0%)
  adding:
saved_models/preprocessing_pipeline/stages/13_OneHotEncoder_ac20700ca6
75/data/ (stored 0%)
  adding:
saved_models/preprocessing_pipeline/stages/13_OneHotEncoder_ac20700ca6
75/data/._SUCCESS.crc (stored 0%)
  adding:
saved_models/preprocessing_pipeline/stages/13_OneHotEncoder_ac20700ca6
75/data/part-00000-569f87e0-f972-4b58-80a7-4ff2c2ee6db3-
c000.snappy.parquet (deflated 27%)
  adding:
saved_models/preprocessing_pipeline/stages/13_OneHotEncoder_ac20700ca6
75/data/.part-00000-569f87e0-f972-4b58-80a7-4ff2c2ee6db3-
c000.snappy.parquet.crc (stored 0%)
  adding:
saved_models/preprocessing_pipeline/stages/13_OneHotEncoder_ac20700ca6
75/data/_SUCCESS (stored 0%)
  adding:
saved_models/preprocessing_pipeline/stages/02_StringIndexer_b77f1dca2d
7a/ (stored 0%)
  adding:
saved_models/preprocessing_pipeline/stages/02_StringIndexer_b77f1dca2d
7a/metadata/ (stored 0%)
  adding:
saved_models/preprocessing_pipeline/stages/02_StringIndexer_b77f1dca2d
7a/metadata/._SUCCESS.crc (stored 0%)
  adding:
saved_models/preprocessing_pipeline/stages/02_StringIndexer_b77f1dca2d
7a/metadata/.part-00000.crc (stored 0%)
  adding:
saved_models/preprocessing_pipeline/stages/02_StringIndexer_b77f1dca2d
7a/metadata/part-00000 (deflated 40%)
  adding:
saved_models/preprocessing_pipeline/stages/02_StringIndexer_b77f1dca2d
7a/metadata/_SUCCESS (stored 0%)
  adding:
saved_models/preprocessing_pipeline/stages/02_StringIndexer_b77f1dca2d
7a/data/ (stored 0%)
  adding:
saved_models/preprocessing_pipeline/stages/02_StringIndexer_b77f1dca2d
7a/data/part-00000-86029cb2-83a9-4033-8b2b-6ad4048c6fbd-
c000.snappy.parquet (deflated 35%)
  adding:
saved_models/preprocessing_pipeline/stages/02_StringIndexer_b77f1dca2d
7a/data/._SUCCESS.crc (stored 0%)
```

```
  adding:
saved_models/preprocessing_pipeline/stages/02_StringIndexer_b77f1dca2d
7a/data/.part-00000-86029cb2-83a9-4033-8b2b-6ad4048c6fbd-
c000.snappy.parquet.crc (stored 0%)
  adding:
saved_models/preprocessing_pipeline/stages/02_StringIndexer_b77f1dca2d
7a/data/_SUCCESS (stored 0%)
  adding:
saved_models/preprocessing_pipeline/stages/03_StringIndexer_41841b909e
a2/ (stored 0%)
  adding:
saved_models/preprocessing_pipeline/stages/03_StringIndexer_41841b909e
a2/metadata/ (stored 0%)
  adding:
saved_models/preprocessing_pipeline/stages/03_StringIndexer_41841b909e
a2/metadata/._SUCCESS.crc (stored 0%)
  adding:
saved_models/preprocessing_pipeline/stages/03_StringIndexer_41841b909e
a2/metadata/.part-00000.crc (stored 0%)
  adding:
saved_models/preprocessing_pipeline/stages/03_StringIndexer_41841b909e
a2/metadata/part-00000 (deflated 40%)
  adding:
saved_models/preprocessing_pipeline/stages/03_StringIndexer_41841b909e
a2/metadata/_SUCCESS (stored 0%)
  adding:
saved_models/preprocessing_pipeline/stages/03_StringIndexer_41841b909e
a2/data/ (stored 0%)
  adding:
saved_models/preprocessing_pipeline/stages/03_StringIndexer_41841b909e
a2/data/part-00000-279fe6d3-3884-4b5e-b5af-7cc5e485ae59-
c000.snappy.parquet (deflated 36%)
  adding:
saved_models/preprocessing_pipeline/stages/03_StringIndexer_41841b909e
a2/data/._SUCCESS.crc (stored 0%)
  adding:
saved_models/preprocessing_pipeline/stages/03_StringIndexer_41841b909e
a2/data/.part-00000-279fe6d3-3884-4b5e-b5af-7cc5e485ae59-
c000.snappy.parquet.crc (stored 0%)
  adding:
saved_models/preprocessing_pipeline/stages/03_StringIndexer_41841b909e
a2/data/_SUCCESS (stored 0%)
  adding:
saved_models/preprocessing_pipeline/stages/01_StringIndexer_4896fc3942
5a/ (stored 0%)
  adding:
saved_models/preprocessing_pipeline/stages/01_StringIndexer_4896fc3942
5a/metadata/ (stored 0%)
  adding:
```

```
saved_models/preprocessing_pipeline/stages/01_StringIndexer_4896fc3942
5a/metadata/._SUCCESS.crc (stored 0%)
  adding:
saved_models/preprocessing_pipeline/stages/01_StringIndexer_4896fc3942
5a/metadata/.part-00000.crc (stored 0%)
  adding:
saved_models/preprocessing_pipeline/stages/01_StringIndexer_4896fc3942
5a/metadata/part-00000 (deflated 38%)
  adding:
saved_models/preprocessing_pipeline/stages/01_StringIndexer_4896fc3942
5a/metadata/_SUCCESS (stored 0%)
  adding:
saved_models/preprocessing_pipeline/stages/01_StringIndexer_4896fc3942
5a/data/ (stored 0%)
  adding:
saved_models/preprocessing_pipeline/stages/01_StringIndexer_4896fc3942
5a/data/part-00000-60c89ab9-8394-476b-8b3e-53c1f1fb8e14-
c000.snappy.parquet (deflated 37%)
  adding:
saved_models/preprocessing_pipeline/stages/01_StringIndexer_4896fc3942
5a/data/._SUCCESS.crc (stored 0%)
  adding:
saved_models/preprocessing_pipeline/stages/01_StringIndexer_4896fc3942
5a/data/_SUCCESS (stored 0%)
  adding:
saved_models/preprocessing_pipeline/stages/01_StringIndexer_4896fc3942
5a/data/.part-00000-60c89ab9-8394-476b-8b3e-53c1f1fb8e14-
c000.snappy.parquet.crc (stored 0%)
  adding: saved_models/preprocessing_pipeline/metadata/ (stored 0%)
  adding: saved_models/preprocessing_pipeline/metadata/._SUCCESS.crc
(stored 0%)
  adding: saved_models/preprocessing_pipeline/metadata/.part-00000.crc
(stored 0%)
  adding: saved_models/preprocessing_pipeline/metadata/part-00000
(deflated 48%)
  adding: saved_models/preprocessing_pipeline/metadata/_SUCCESS
(stored 0%)
```