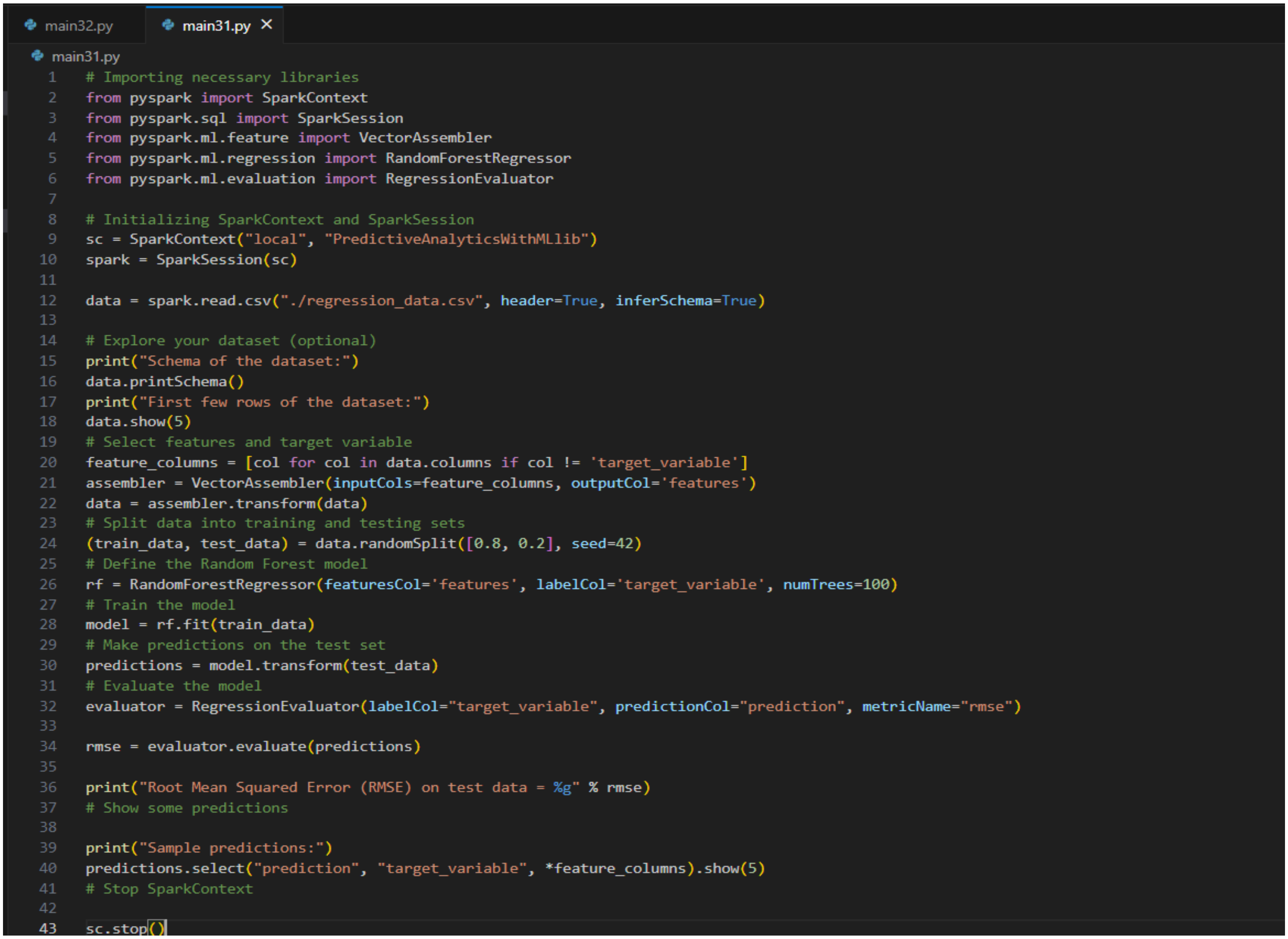
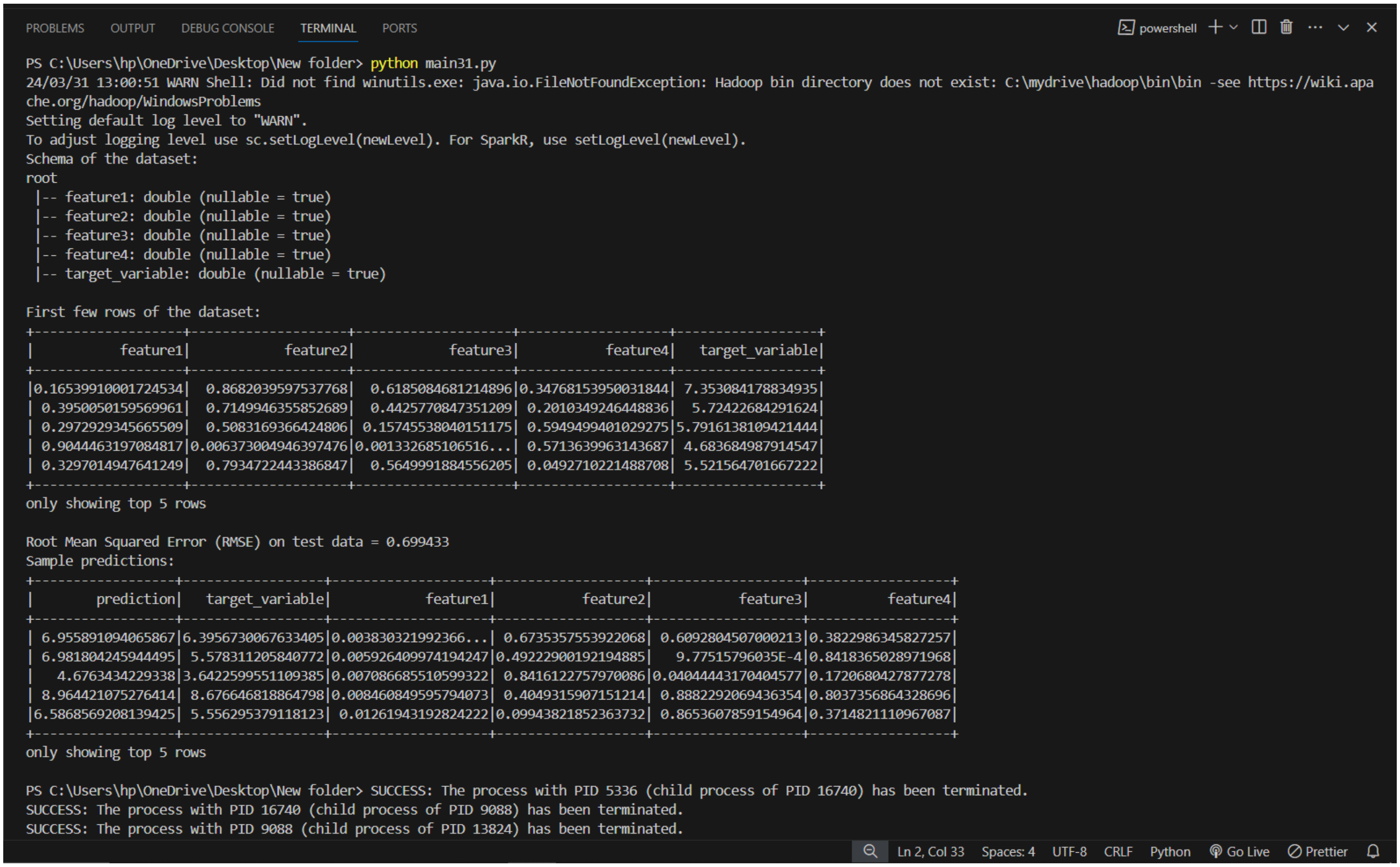


```python
# Importing necessary libraries
from pyspark import SparkContext
from pyspark.sql import SparkSession
from pyspark.ml.feature import VectorAssembler
from pyspark.ml.regression import RandomForestRegressor
from pyspark.ml.evaluation import RegressionEvaluator

# Initializing SparkContext and SparkSession
sc = SparkContext("local", "PredictiveAnalyticsWithMLlib")
spark = SparkSession(sc)

data = spark.read.csv("./regression_data.csv", header=True, inferSchema=True)

# Explore your dataset (optional)
print("Schema of the dataset:")
data.printSchema()
print("First few rows of the dataset:")
data.show(5)
# Select features and target variable
feature_columns = [col for col in data.columns if col != 'target_variable']
assembler = VectorAssembler(inputCols=feature_columns, outputCol='features')
data = assembler.transform(data)
# Split data into training and testing sets
(train_data, test_data) = data.randomSplit([0.8, 0.2], seed=42)
# Define the Random Forest model
rf = RandomForestRegressor(featuresCol='features', labelCol='target_variable', numTrees=100)
# Train the model
model = rf.fit(train_data)
# Make predictions on the test set
predictions = model.transform(test_data)
# Evaluate the model
evaluator = RegressionEvaluator(labelCol="target_variable", predictionCol="prediction", metricName="rmse")

rmse = evaluator.evaluate(predictions)

print("Root Mean Squared Error (RMSE) on test data = %g" % rmse)
# Show some predictions

print("Sample predictions:")
predictions.select("prediction", "target_variable", *feature_columns).show(5)
# Stop SparkContext

sc.stop()
```



```
PS C:\Users\hp\OneDrive\Desktop\New folder> python main31.py
24/03/31 13:00:51 WARN Shell: Did not find winutils.exe: java.io.FileNotFoundException: Hadoop bin directory does not exist: C:\mydrive\hadoop\bin\bin -see https://wiki.apa
che.org/hadoop/WindowsProblems
Setting default log level to "WARN".
To adjust logging level use sc.setLogLevel(newLevel). For SparkR, use setLogLevel(newLevel).
Schema of the dataset:
root
 |-- feature1: double (nullable = true)
 |-- feature2: double (nullable = true)
 |-- feature3: double (nullable = true)
 |-- feature4: double (nullable = true)
 |-- target_variable: double (nullable = true)

First few rows of the dataset:
+-------------------+--------------------+--------------------+-------------------+------------------+
|           feature1|            feature2|            feature3|           feature4|   target_variable|
+-------------------+--------------------+--------------------+-------------------+------------------+
|0.16539910001724534|  0.8682039597537768|  0.6185084681214896|0.34768153950031844| 7.353084178834935|
| 0.3950050159569961|  0.7149946355852689|  0.4425770847351209| 0.2010349246448836|  5.72422684291624|
| 0.2972929345665509|  0.5083169366424806| 0.15745538040151175| 0.5949499401029275|5.7916138109421444|
| 0.9044463197084817|0.006373004946397476|0.001332685106516...| 0.5713639963143687| 4.683684987914547|
| 0.3297014947641249|  0.7934722443386847|  0.5649991884556205| 0.0492710221488708| 5.521564701667222|
+-------------------+--------------------+--------------------+-------------------+------------------+
only showing top 5 rows

Root Mean Squared Error (RMSE) on test data = 0.699433
Sample predictions:
+------------------+------------------+--------------------+-------------------+-------------------+------------------+
|        prediction|   target_variable|            feature1|           feature2|           feature3|          feature4|
+------------------+------------------+--------------------+-------------------+-------------------+------------------+
| 6.955891094065867|6.3956730067633405|0.003830321992366...| 0.6735357553922068| 0.6092804507000213|0.3822986345827257|
| 6.981804245944495| 5.578311205840772|0.005926409974194247|0.49222900192194885|   9.77515796035E-4|0.8418365028971968|
|   4.6763434229338|3.6422599551109385|0.007086685510599322| 0.8416122757970086|0.04044443170404577|0.1720680427877278|
| 8.964421075276414| 8.676646818864798|0.008460849595794073| 0.4049315907151214| 0.8882292069436354|0.8037356864328696|
|6.5868569208139425| 5.556295379118123| 0.01261943192824222|0.09943821852363732| 0.8653607859154964|0.3714821110967087|
+------------------+------------------+--------------------+-------------------+-------------------+------------------+
only showing top 5 rows

PS C:\Users\hp\OneDrive\Desktop\New folder> SUCCESS: The process with PID 5336 (child process of PID 16740) has been terminated.
SUCCESS: The process with PID 16740 (child process of PID 9088) has been terminated.
SUCCESS: The process with PID 9088 (child process of PID 13824) has been terminated.
```