Load Packages

```
import pyspark
import pyspark.sql.functions as F
from pyspark.sql.functions import *
from pyspark.sql.types import IntegerType

from pyspark.ml.regression import *
from pyspark.ml.evaluation import *
from pyspark.ml.feature import VectorAssembler
from pyspark.ml.tuning import ParamGridBuilder, TrainValidationSplit
from pyspark.ml.tuning import *
import mlflow
import mlflow.spark

import pandas as pd
import numpy as np
```

Data Wrangling

```
df_results = spark.read.csv('s3://columbia-gr5069-main/raw/results.csv',header=True)
display(df_results)
```

6	18	6	3	8	13	6	6	6	3	57	\N	\N	50	14	1:29.639	21
7	18	7	5	14	17	7	7	7	2	55	\N	\N	22	12	1:29.534	21
8	18	8	6	1	15	8	8	8	1	53	\N	\N	20	4	1:27.903	21
9	18	9	2	4	2	\N	R	9	0	47	\N	\N	15	9	1:28.753	21
10	18	10	7	12	18	\N	R	10	0	43	\N	\N	23	13	1:29.558	21
11	18	11	8	18	19	\N	R	11	0	32	\N	\N	24	15	1:30.892	21
12	18	12	4	6	20	\N	R	12	0	30	\N	\N	20	16	1:31.384	20
13	18	13	6	2	4	\N	R	13	0	29	\N	\N	23	6	1:28.175	21
14	18	14	9	9	8	\N	R	14	0	25	\N	\N	21	11	1:29.502	21
15	18	15	7	11	6	\N	R	15	0	19	\N	\N	18	10	1:29.310	21

```
df = df_results.select('raceId','driverId','resultId','positionOrder','points','laps','fastestLap','fastestLapTime','fastestLapSpeed','statusId')
display(df)
```

raceId	driverId	resultId	positionOrder	points	laps	fastestLap	fastestLapTime	fastestLapSpeed	statusId
18	1	1	1	10	58	39	1:27.452	218.300	1
18	2	2	2	8	58	41	1:27.739	217.586	1
18	3	3	3	6	58	41	1:28.090	216.719	1
18	4	4	4	5	58	58	1:28.603	215.464	1
18	5	5	5	4	58	43	1:27.418	218.385	1
18	6	6	6	3	57	50	1:29.639	212.974	11
18	7	7	7	2	55	22	1:29.534	213.224	5
18	8	8	8	1	53	20	1:27.903	217.180	5
18	9	9	9	0	47	15	1:28.753	215.100	4
18	10	10	10	0	43	23	1:29.558	213.166	3

```
df2 = df.filter(df.fastestLapTime != '\\N' )
display(df2)
```

18	3	3	3	6	58	41	1:28.090	216.719	1
18	4	4	4	5	58	58	1:28.603	215.464	1
18	5	5	5	4	58	43	1:27.418	218.385	1
18	6	6	6	3	57	50	1:29.639	212.974	11
18	7	7	7	2	55	22	1:29.534	213.224	5
18	8	8	8	1	53	20	1:27.903	217.180	5
18	9	9	9	0	47	15	1:28.753	215.100	4
18	10	10	10	0	43	23	1:29.558	213.166	3
18	11	11	11	0	32	24	1:30.892	210.038	7
18	12	12	12	0	30	20	1:31.384	208.907	8
18	13	13	13	0	29	23	1:28.175	216.510	5

```
for col_name in df2.columns:
    df2 = df2.withColumn(col_name, col(col_name).cast("Integer"))
df2 = df2.drop("fastestLapTime")
display(df2)
```

raceId	driverId	resultId	positionOrder	points	laps	fastestLap	fastestLapSpeed	statusId
18	1	1	1	10	58	39	218	1
18	2	2	2	8	58	41	217	1
18	3	3	3	6	58	41	216	1
18	4	4	4	5	58	58	215	1
18	5	5	5	4	58	43	218	1
18	6	6	6	3	57	50	212	11
18	7	7	7	2	55	22	213	5
18	8	8	8	1	53	20	217	5
18	9	9	9	0	47	15	215	4
18	10	10	10	0	43	23	213	3

```
(df2.describe().select(
                                    "summary",
                                    F.round("raceId", 4).alias("raceId"),
                                    F.round("driverId", 4).alias("driverId"),
                                    F.round("resultId", 4).alias("resultId"),
                                    F.round("positionOrder", 4).alias("positionOrder"),
                                    F.round("points", 4).alias("points"),
                                    F.round("laps", 4).alias("laps"),
                                    F.round("fastestLap", 4).alias("fastestLap"),
                                    F.round("fastestLapSpeed", 4).alias("fastestLapSpeed"),
                                    F.round("statusId", 4).alias("statusId"))
```

```
                                    F.round("statusId", 4).alias("statusId"))
                                    .show())
```

```
+-------+--------+--------+--------+-------------+------+-------+----------+--------------+--------+
|summary|  raceId|driverId|resultId|positionOrder|points|   laps|fastestLap|fastestLapSpeed|statusId|
+-------+--------+--------+--------+-------------+------+-------+----------+--------------+--------+
|  count|  7379.0|  7379.0|  7379.0|       7379.0|7379.0| 7379.0|    7379.0|        7379.0|  7379.0|
|   mean|668.8462|363.4491|17093.185|       10.713| 4.116|54.9416|   42.5142|      202.5907|   8.509|
| stddev|418.2864|397.8858|9698.2234|       5.9932|6.4427|15.2391|   16.8357|      21.3601|  16.2741|
|    min|     1.0|     1.0|     1.0|          1.0|   0.0|    0.0|       2.0|          89.0|     1.0|
|    max|  1096.0|   856.0| 25845.0|         24.0|  50.0|   87.0|      85.0|         257.0|   141.0|
+-------+--------+--------+--------+-------------+------+-------+----------+--------------+--------+
```

```
df2.printSchema()
```

```
root
-- raceId: integer (nullable = true)
-- driverId: integer (nullable = true)
-- resultId: integer (nullable = true)
-- positionOrder: integer (nullable = true)
-- points: integer (nullable = true)
-- laps: integer (nullable = true)
-- fastestLap: integer (nullable = true)
-- fastestLapSpeed: integer (nullable = true)
-- statusId: integer (nullable = true)
```

Feature Extraction

```
featureCols = ["laps", "fastestLap", "fastestLapSpeed", "statusId", "driverId", "raceId"]
assembler = VectorAssembler(inputCols=featureCols, outputCol="features")
assembled_df = assembler.transform(df2)
display(assembled_df)
```

raceId	driverId	resultId	positionOrder	points	laps	fastestLap	fastestLapSpeed	statusId	features
18	1	1	1	10	58	39	218	1	List(1, 6, List(), List(58.0, 39.0, 218.0, 1.0, 1.0, 18.0))
18	2	2	2	8	58	41	217	1	List(1, 6, List(), List(58.0, 41.0, 217.0, 1.0, 2.0, 18.0))
18	3	3	3	6	58	41	216	1	List(1, 6, List(), List(58.0, 41.0, 216.0, 1.0, 3.0, 18.0))
18	4	4	4	5	58	58	215	1	List(1, 6, List(), List(58.0, 58.0, 215.0, 1.0, 4.0, 18.0))
18	5	5	5	4	58	43	218	1	List(1, 6, List(), List(58.0, 43.0, 218.0, 1.0, 5.0, 18.0))
18	6	6	6	3	57	50	212	11	List(1, 6, List(), List(57.0, 50.0, 212.0, 11.0, 6.0, 18.0))
18	7	7	7	2	55	22	213	5	List(1, 6, List(), List(55.0, 22.0, 213.0, 5.0, 7.0, 18.0))
18	8	8	8	1	53	20	217	5	List(1, 6, List(), List(53.0, 20.0, 217.0, 5.0, 8.0, 18.0))
18	9	9	9	0	47	15	215	4	List(1, 6, List(), List(47.0, 15.0, 215.0, 4.0, 9.0, 18.0))
18	10	10	10	0	43	23	213	3	List(1, 6, List(), List(43.0, 23.0, 213.0, 3.0, 10.0, 18.0))

Prepare training and test data

```
train, test = assembled_df.randomSplit([0.9, 0.1], seed=12345)
display(train)
```

raceId	driverId	resultId	positionOrder	points	laps	fastestLap	fastestLapSpeed	statusId	features
1	1	7573	20	0	58	39	214	2	List(1, 6, List(), List(58.0, 39.0, 214.0, 2.0, 1.0, 1.0))
1	2	7563	10	0	58	48	216	1	List(1, 6, List(), List(58.0, 48.0, 216.0, 1.0, 2.0, 1.0))
1	3	7559	6	3	58	48	217	1	List(1, 6, List(), List(58.0, 48.0, 217.0, 1.0, 3.0, 1.0))
1	4	7558	5	4	58	53	215	1	List(1, 6, List(), List(58.0, 53.0, 215.0, 1.0, 4.0, 1.0))
1	6	7571	18	0	17	6	212	3	List(1, 6, List(), List(17.0, 6.0, 212.0, 3.0, 6.0, 1.0))
1	7	7561	8	1	58	50	212	1	List(1, 6, List(), List(58.0, 50.0, 212.0, 1.0, 7.0, 1.0))
1	8	7568	15	0	55	35	215	24	List(1, 6, List(), List(55.0, 35.0, 215.0, 24.0, 8.0, 1.0))
1	9	7567	14	0	55	36	216	4	List(1, 6, List(), List(55.0, 36.0, 216.0, 4.0, 9.0, 1.0))
1	10	7557	4	5	58	53	215	1	List(1, 6, List(), List(58.0, 53.0, 215.0, 1.0, 10.0, 1.0))
1	12	7570	17	0	24	17	210	20	List(1, 6, List(), List(24.0, 17.0, 210.0, 20.0, 12.0, 1.0))

Fit Model

```
lr = LinearRegression(featuresCol="features",labelCol='positionOrder',predictionCol='predpositionOrder',maxIter=10, regParam=1, elasticNetParam=1,standardization=False)
model=lr.fit(train)
```

```
model.coefficients
```

```
Out[63]: DenseVector([-0.1665, 0.0134, -0.0475, 0.105, 0.003, -0.0019])
```

```
featureCols
```

```
Out[64]: ['laps', 'fastestLap', 'fastestLapSpeed', 'statusId', 'driverId', 'raceId']
```

```
model.intercept
```

```
Out[65]: 28.184234435518764
```

```
coeff_df = pd.DataFrame({"Feature": ["Intercept"] + featureCols, "Co-efficients": np.insert(model.coefficients.toArray(), 0, model.intercept)})
coeff_df = coeff_df[["Feature", "Co-efficients"]]
display(coeff_df)
```

Feature	Co-efficients
Intercept	28.184234435518764
laps	-0.16649838162466551
fastestLap	0.01335855165524685
fastestLapSpeed	-0.04747640140956104
statusId	0.1049845434326192
driverId	0.0030216930656948976
raceId	-0.0018627411946988339

Generate Predictions

```
predictions = model.transform(test)
```

```
result = RegressionEvaluator(labelCol='positionOrder',predictionCol='predpositionOrder').evaluate(predictions)
result
```

```
Out[68]: 4.63074975564234
```

```
predandlabels = predictions.select("positionOrder", "predpositionOrder").show()
```

```
+-------------+------------------+
positionOrder| predpositionOrder|
+-------------+------------------+
           10|13.829534212474647|
           14|14.834924938456977|
           16|16.158681951605523|
           16|13.021355773487052|
           17|14.478477466644637|
            2| 8.141295694810836|
            6|  8.44017290099757|
           15| 9.326167621690395|
           18|10.357346258441076|
            8| 9.047713972424358|
            3| 8.963661743395509|
            6| 8.419517868989146|
           14|10.190407267215761|
            7| 8.310623510578488|
           16|10.116119379166825|
            1| 8.516558498867706|
           13| 9.990548408904452|
           19| 17.03660175217184|
            8| 6.849110410436918|
            2| 9.938202415482085|
+-------------+------------------+
only showing top 20 rows
```

Tuning hyperparameters Individually

```
def run_model(elasticNetParaminput):
  with mlflow.start_run(run_name="Elastic-Net-Regression-Model") as run:
    ### Feature Extraction
    featureCols = ["laps", "fastestLap", "fastestLapSpeed", "statusId", "driverId", "raceId"]
    assembler = VectorAssembler(inputCols=featureCols, outputCol="features")
    assembled_df = assembler.transform(df2)
    ### Prepare training and test data
    train, test = assembled_df.randomSplit([0.9, 0.1], seed=12345)
    ### Fit Model
    ### Generate Predictions
    regParam=1
    fitIntercept=True
    elasticNetParam=elasticNetParaminput
    lr = LinearRegression(featuresCol="features",labelCol='positionOrder',predictionCol='predpositionOrder',maxIter=10, regParam=1, elasticNetParam=elasticNetParaminput,standardization=False)
    ### Tuning hyperparameters Individually
    model1 = lr.fit(train)
    mlflow.spark.log_model(model1, "Elastic-Net-Regression-Model")
    predtest = model1.transform(test)
    evaluator = RegressionEvaluator(labelCol='positionOrder',predictionCol='predpositionOrder')
    rmse = evaluator.evaluate(predtest, {evaluator.metricName: "rmse"})
    mlflow.log_metric("rmse", rmse)
    mlflow.log_param("regParam",regParam)
    mlflow.log_param("fitIntercept",fitIntercept)
    mlflow.log_param("elasticNetParam",elasticNetParam)
    print("    rmse: {}".format(rmse))
    print("    regParam: {}".format(regParam))
    print("    elasticNetParam: {}".format(elasticNetParam))
    print("    fitIntercept: {}".format(fitIntercept))
    runID = run.info.run_uuid
    experimentID = run.info.experiment_id
    print("Inside MLflow Run with run_id {} and experiment_id {}".format(runID, experimentID))
  return (run.info.run_uuid, run.info.experiment_id)

run_model(elasticNetParaminput=0.1)
run_model(elasticNetParaminput=0.2)
run_model(elasticNetParaminput=0.3)
run_model(elasticNetParaminput=0.4)
run_model(elasticNetParaminput=0.5)
run_model(elasticNetParaminput=0.6)
run_model(elasticNetParaminput=0.7)
run_model(elasticNetParaminput=0.8)
run_model(elasticNetParaminput=0.9)
run_model(elasticNetParaminput=1.0)
run_model(elasticNetParaminput=0)
```

```
2023/05/03 03:40:32 WARNING mlflow.utils.autologging_utils: Encountered unexpected error during spark autologging: Exception while attempting to initialize JVM-side state for S
park datasource autologging. Please create a new Spark session and ensure you have the mlflow-spark JAR attached to your Spark session as described in http://mlflow.org/docs/la
test/tracking.html#automatic-logging-from-spark-experimental. Exception:
'JavaPackage' object is not callable
2023/05/03 03:40:32 WARNING mlflow.utils.autologging_utils: Encountered unexpected error during spark autologging: Exception while attempting to initialize JVM-side state for S
park datasource autologging. Please create a new Spark session and ensure you have the mlflow-spark JAR attached to your Spark session as described in http://mlflow.org/docs/la
test/tracking.html#automatic-logging-from-spark-experimental. Exception:
'JavaPackage' object is not callable
2023/05/03 03:40:32 WARNING mlflow.utils.autologging_utils: Encountered unexpected error during spark autologging: Exception while attempting to initialize JVM-side state for S
park datasource autologging. Please create a new Spark session and ensure you have the mlflow-spark JAR attached to your Spark session as described in http://mlflow.org/docs/la
test/tracking.html#automatic-logging-from-spark-experimental. Exception:
'JavaPackage' object is not callable
2023/05/03 03:40:32 WARNING mlflow.utils.autologging_utils: Encountered unexpected error during spark autologging: Exception while attempting to initialize JVM-side state for S
park datasource autologging. Please create a new Spark session and ensure you have the mlflow-spark JAR attached to your Spark session as described in http://mlflow.org/docs/la
test/tracking.html#automatic-logging-from-spark-experimental. Exception:
'JavaPackage' object is not callable
2023/05/03 03:40:32 WARNING mlflow.utils.autologging_utils: Encountered unexpected error during spark autologging: Exception while attempting to initialize JVM-side state for S
park datasource autologging. Please create a new Spark session and ensure you have the mlflow-spark JAR attached to your Spark session as described in http://mlflow.org/docs/la
test/tracking.html#automatic-logging-from-spark-experimental. Exception:
'JavaPackage' object is not callable
2023/05/03 03:40:32 WARNING mlflow.utils.autologging_utils: Encountered unexpected error during spark autologging: Exception while attempting to initialize JVM-side state for S
park datasource autologging. Please create a new Spark session and ensure you have the mlflow-spark JAR attached to your Spark session as described in http://mlflow.org/docs/la
test/tracking.html#automatic-logging-from-spark-experimental. Exception:
'JavaPackage' object is not callable
2023/05/03 03:40:33 WARNING mlflow.utils.autologging_utils: Encountered unexpected error during spark autologging: Exception while attempting to initialize JVM-side state for S
park datasource autologging. Please create a new Spark session and ensure you have the mlflow-spark JAR attached to your Spark session as described in http://mlflow.org/docs/la
test/tracking.html#automatic-logging-from-spark-experimental. Exception:
'JavaPackage' object is not callable
2023/05/03 03:40:33 WARNING mlflow.utils.autologging_utils: Encountered unexpected error during spark autologging: Exception while attempting to initialize JVM-side state for S
park datasource autologging. Please create a new Spark session and ensure you have the mlflow-spark JAR attached to your Spark session as described in http://mlflow.org/docs/la
test/tracking.html#automatic-logging-from-spark-experimental. Exception:
'JavaPackage' object is not callable
2023/05/03 03:40:33 WARNING mlflow.utils.autologging_utils: Encountered unexpected error during spark autologging: Exception while attempting to initialize JVM-side state for S
park datasource autologging. Please create a new Spark session and ensure you have the mlflow-spark JAR attached to your Spark session as described in http://mlflow.org/docs/la
test/tracking.html#automatic-logging-from-spark-experimental. Exception:
'JavaPackage' object is not callable
2023/05/03 03:40:33 WARNING mlflow.utils.autologging_utils: Encountered unexpected error during spark autologging: Exception while attempting to initialize JVM-side state for S
park datasource autologging. Please create a new Spark session and ensure you have the mlflow-spark JAR attached to your Spark session as described in http://mlflow.org/docs/la
test/tracking.html#automatic-logging-from-spark-experimental. Exception:
'JavaPackage' object is not callable
2023/05/03 03:40:37 WARNING mlflow.utils.autologging_utils: Encountered unexpected error during spark autologging: Exception while attempting to initialize JVM-side state for S
park datasource autologging. Please create a new Spark session and ensure you have the mlflow-spark JAR attached to your Spark session as described in http://mlflow.org/docs/la
test/tracking.html#automatic-logging-from-spark-experimental. Exception:
'JavaPackage' object is not callable
2023/05/03 03:40:47 WARNING mlflow.utils.environment: Encountered an unexpected error while inferring pip requirements (model URI: /tmp/tmpr6a1gbuq, flavor: spark), fall back t
o return ['pyspark==3.0.1']. Set logging level to DEBUG to see the full traceback.
2023/05/03 03:40:47 WARNING mlflow.utils.autologging_utils: Encountered unexpected error during spark autologging: Exception while attempting to initialize JVM-side state for S
park datasource autologging. Please create a new Spark session and ensure you have the mlflow-spark JAR attached to your Spark session as described in http://mlflow.org/docs/la
test/tracking.html#automatic-logging-from-spark-experimental. Exception:
'JavaPackage' object is not callable
2023/05/03 03:40:47 WARNING mlflow.utils.autologging_utils: Encountered unexpected error during spark autologging: Exception while attempting to initialize JVM-side state for S
park datasource autologging. Please create a new Spark session and ensure you have the mlflow-spark JAR attached to your Spark session as described in http://mlflow.org/docs/la
test/tracking.html#automatic-logging-from-spark-experimental. Exception:
'JavaPackage' object is not callable
2023/05/03 03:40:47 WARNING mlflow.utils.autologging_utils: Encountered unexpected error during spark autologging: Exception while attempting to initialize JVM-side state for S
park datasource autologging. Please create a new Spark session and ensure you have the mlflow-spark JAR attached to your Spark session as described in http://mlflow.org/docs/la
test/tracking.html#automatic-logging-from-spark-experimental. Exception:
'JavaPackage' object is not callable
2023/05/03 03:40:47 WARNING mlflow.utils.autologging_utils: Encountered unexpected error during spark autologging: Exception while attempting to initialize JVM-side state for S
park datasource autologging. Please create a new Spark session and ensure you have the mlflow-spark JAR attached to your Spark session as described in http://mlflow.org/docs/la
test/tracking.html#automatic-logging-from-spark-experimental. Exception:
'JavaPackage' object is not callable
2023/05/03 03:40:47 WARNING mlflow.utils.autologging_utils: Encountered unexpected error during spark autologging: Exception while attempting to initialize JVM-side state for S
park datasource autologging. Please create a new Spark session and ensure you have the mlflow-spark JAR attached to your Spark session as described in http://mlflow.org/docs/la
test/tracking.html#automatic-logging-from-spark-experimental. Exception:
'JavaPackage' object is not callable
2023/05/03 03:40:47 WARNING mlflow.utils.autologging_utils: Encountered unexpected error during spark autologging: Exception while attempting to initialize JVM-side state for S
park datasource autologging. Please create a new Spark session and ensure you have the mlflow-spark JAR attached to your Spark session as described in http://mlflow.org/docs/la
test/tracking.html#automatic-logging-from-spark-experimental. Exception:
'JavaPackage' object is not callable
2023/05/03 03:40:47 WARNING mlflow.utils.autologging_utils: Encountered unexpected error during spark autologging: Exception while attempting to initialize JVM-side state for S
park datasource autologging. Please create a new Spark session and ensure you have the mlflow-spark JAR attached to your Spark session as described in http://mlflow.org/docs/la
test/tracking.html#automatic-logging-from-spark-experimental. Exception:
'JavaPackage' object is not callable
2023/05/03 03:40:47 WARNING mlflow.utils.autologging_utils: Encountered unexpected error during spark autologging: Exception while attempting to initialize JVM-side state for S
park datasource autologging. Please create a new Spark session and ensure you have the mlflow-spark JAR attached to your Spark session as described in http://mlflow.org/docs/la
test/tracking.html#automatic-logging-from-spark-experimental. Exception:
'JavaPackage' object is not callable
2023/05/03 03:40:47 WARNING mlflow.utils.autologging_utils: Encountered unexpected error during spark autologging: Exception while attempting to initialize JVM-side state for S
park datasource autologging. Please create a new Spark session and ensure you have the mlflow-spark JAR attached to your Spark session as described in http://mlflow.org/docs/la
test/tracking.html#automatic-logging-from-spark-experimental. Exception:
'JavaPackage' object is not callable
2023/05/03 03:40:47 WARNING mlflow.utils.autologging_utils: Encountered unexpected error during spark autologging: Exception while attempting to initialize JVM-side state for S
park datasource autologging. Please create a new Spark session and ensure you have the mlflow-spark JAR attached to your Spark session as described in http://mlflow.org/docs/la
test/tracking.html#automatic-logging-from-spark-experimental. Exception:
'JavaPackage' object is not callable
2023/05/03 03:40:47 WARNING mlflow.utils.autologging_utils: Encountered unexpected error during spark autologging: Exception while attempting to initialize JVM-side state for S
park datasource autologging. Please create a new Spark session and ensure you have the mlflow-spark JAR attached to your Spark session as described in http://mlflow.org/docs/la
test/tracking.html#automatic-logging-from-spark-experimental. Exception:
'JavaPackage' object is not callable
2023/05/03 03:40:48 WARNING mlflow.utils.autologging_utils: Encountered unexpected error during spark autologging: Exception while attempting to initialize JVM-side state for S
park datasource autologging. Please create a new Spark session and ensure you have the mlflow-spark JAR attached to your Spark session as described in http://mlflow.org/docs/la
test/tracking.html#automatic-logging-from-spark-experimental. Exception:
'JavaPackage' object is not callable
2023/05/03 03:40:48 WARNING mlflow.utils.autologging_utils: Encountered unexpected error during spark autologging: Exception while attempting to initialize JVM-side state for S
park datasource autologging. Please create a new Spark session and ensure you have the mlflow-spark JAR attached to your Spark session as described in http://mlflow.org/docs/la
test/tracking.html#automatic-logging-from-spark-experimental. Exception:
'JavaPackage' object is not callable
2023/05/03 03:40:48 WARNING mlflow.utils.autologging_utils: Encountered unexpected error during spark autologging: Exception while attempting to initialize JVM-side state for S
park datasource autologging. Please create a new Spark session and ensure you have the mlflow-spark JAR attached to your Spark session as described in http://mlflow.org/docs/la
test/tracking.html#automatic-logging-from-spark-experimental. Exception:
'JavaPackage' object is not callable
2023/05/03 03:40:48 WARNING mlflow.utils.autologging_utils: Encountered unexpected error during spark autologging: Exception while attempting to initialize JVM-side state for S
park datasource autologging. Please create a new Spark session and ensure you have the mlflow-spark JAR attached to your Spark session as described in http://mlflow.org/docs/la
test/tracking.html#automatic-logging-from-spark-experimental. Exception:
'JavaPackage' object is not callable
2023/05/03 03:40:48 WARNING mlflow.utils.autologging_utils: Encountered unexpected error during spark autologging: Exception while attempting to initialize JVM-side state for S
park datasource autologging. Please create a new Spark session and ensure you have the mlflow-spark JAR attached to your Spark session as described in http://mlflow.org/docs/la
test/tracking.html#automatic-logging-from-spark-experimental. Exception:
'JavaPackage' object is not callable
2023/05/03 03:40:48 WARNING mlflow.utils.autologging_utils: Encountered unexpected error during spark autologging: Exception while attempting to initialize JVM-side state for S
park datasource autologging. Please create a new Spark session and ensure you have the mlflow-spark JAR attached to your Spark session as described in http://mlflow.org/docs/la
test/tracking.html#automatic-logging-from-spark-experimental. Exception:
'JavaPackage' object is not callable
2023/05/03 03:40:48 WARNING mlflow.utils.autologging_utils: Encountered unexpected error during spark autologging: Exception while attempting to initialize JVM-side state for S
park datasource autologging. Please create a new Spark session and ensure you have the mlflow-spark JAR attached to your Spark session as described in http://mlflow.org/docs/la
test/tracking.html#automatic-logging-from-spark-experimental. Exception:
'JavaPackage' object is not callable
2023/05/03 03:40:48 WARNING mlflow.utils.autologging_utils: Encountered unexpected error during spark autologging: Exception while attempting to initialize JVM-side state for S
park datasource autologging. Please create a new Spark session and ensure you have the mlflow-spark JAR attached to your Spark session as described in http://mlflow.org/docs/la
test/tracking.html#automatic-logging-from-spark-experimental. Exception:
'JavaPackage' object is not callable
2023/05/03 03:40:48 WARNING mlflow.utils.autologging_utils: Encountered unexpected error during spark autologging: Exception while attempting to initialize JVM-side state for S
park datasource autologging. Please create a new Spark session and ensure you have the mlflow-spark JAR attached to your Spark session as described in http://mlflow.org/docs/la
test/tracking.html#automatic-logging-from-spark-experimental. Exception:
```

```
'JavaPackage' object is not callable
2023/05/03 03:40:48 WARNING mlflow.utils.autologging_utils: Encountered unexpected error during spark autologging: Exception while attempting to initialize JVM-side state for S
park datasource autologging. Please create a new Spark session and ensure you have the mlflow-spark JAR attached to your Spark session as described in http://mlflow.org/docs/la
test/tracking.html#automatic-logging-from-spark-experimental. Exception:
'JavaPackage' object is not callable
2023/05/03 03:40:48 WARNING mlflow.utils.autologging_utils: Encountered unexpected error during spark autologging: Exception while attempting to initialize JVM-side state for S
park datasource autologging. Please create a new Spark session and ensure you have the mlflow-spark JAR attached to your Spark session as described in http://mlflow.org/docs/la
test/tracking.html#automatic-logging-from-spark-experimental. Exception:
'JavaPackage' object is not callable
2023/05/03 03:40:48 WARNING mlflow.utils.autologging_utils: Encountered unexpected error during spark autologging: Exception while attempting to initialize JVM-side state for S
park datasource autologging. Please create a new Spark session and ensure you have the mlflow-spark JAR attached to your Spark session as described in http://mlflow.org/docs/la
test/tracking.html#automatic-logging-from-spark-experimental. Exception:
'JavaPackage' object is not callable
2023/05/03 03:40:48 WARNING mlflow.utils.autologging_utils: Encountered unexpected error during spark autologging: Exception while attempting to initialize JVM-side state for S
park datasource autologging. Please create a new Spark session and ensure you have the mlflow-spark JAR attached to your Spark session as described in http://mlflow.org/docs/la
test/tracking.html#automatic-logging-from-spark-experimental. Exception:
'JavaPackage' object is not callable
2023/05/03 03:40:48 WARNING mlflow.utils.autologging_utils: Encountered unexpected error during spark autologging: Exception while attempting to initialize JVM-side state for S
park datasource autologging. Please create a new Spark session and ensure you have the mlflow-spark JAR attached to your Spark session as described in http://mlflow.org/docs/la
test/tracking.html#automatic-logging-from-spark-experimental. Exception:
'JavaPackage' object is not callable
2023/05/03 03:40:48 WARNING mlflow.utils.autologging_utils: Encountered unexpected error during spark autologging: Exception while attempting to initialize JVM-side state for S
park datasource autologging. Please create a new Spark session and ensure you have the mlflow-spark JAR attached to your Spark session as described in http://mlflow.org/docs/la
test/tracking.html#automatic-logging-from-spark-experimental. Exception:
'JavaPackage' object is not callable
  rmse: 4.626985266693593
  regParam: 1
  elasticNetParam: 0.1
  fitIntercept: True
Inside MLflow Run with run_id f6e2ac4e62b84f239e54c141b9713fe7 and experiment_id 5a58fa41f4ba496bbb8b860919a46b1a
2023/05/03 03:40:48 WARNING mlflow.utils.autologging_utils: Encountered unexpected error during spark autologging: Exception while attempting to initialize JVM-side state for S
park datasource autologging. Please create a new Spark session and ensure you have the mlflow-spark JAR attached to your Spark session as described in http://mlflow.org/docs/la
test/tracking.html#automatic-logging-from-spark-experimental. Exception:
'JavaPackage' object is not callable
2023/05/03 03:40:48 WARNING mlflow.utils.autologging_utils: Encountered unexpected error during spark autologging: Exception while attempting to initialize JVM-side state for S
park datasource autologging. Please create a new Spark session and ensure you have the mlflow-spark JAR attached to your Spark session as described in http://mlflow.org/docs/la
test/tracking.html#automatic-logging-from-spark-experimental. Exception:
'JavaPackage' object is not callable
2023/05/03 03:40:48 WARNING mlflow.utils.autologging_utils: Encountered unexpected error during spark autologging: Exception while attempting to initialize JVM-side state for S
park datasource autologging. Please create a new Spark session and ensure you have the mlflow-spark JAR attached to your Spark session as described in http://mlflow.org/docs/la
test/tracking.html#automatic-logging-from-spark-experimental. Exception:
'JavaPackage' object is not callable
2023/05/03 03:40:49 WARNING mlflow.utils.autologging_utils: Encountered unexpected error during spark autologging: Exception while attempting to initialize JVM-side state for S
park datasource autologging. Please create a new Spark session and ensure you have the mlflow-spark JAR attached to your Spark session as described in http://mlflow.org/docs/la
test/tracking.html#automatic-logging-from-spark-experimental. Exception:
'JavaPackage' object is not callable
2023/05/03 03:40:49 WARNING mlflow.utils.autologging_utils: Encountered unexpected error during spark autologging: Exception while attempting to initialize JVM-side state for S
park datasource autologging. Please create a new Spark session and ensure you have the mlflow-spark JAR attached to your Spark session as described in http://mlflow.org/docs/la
test/tracking.html#automatic-logging-from-spark-experimental. Exception:
'JavaPackage' object is not callable
2023/05/03 03:40:49 WARNING mlflow.utils.autologging_utils: Encountered unexpected error during spark autologging: Exception while attempting to initialize JVM-side state for S
park datasource autologging. Please create a new Spark session and ensure you have the mlflow-spark JAR attached to your Spark session as described in http://mlflow.org/docs/la
test/tracking.html#automatic-logging-from-spark-experimental. Exception:
'JavaPackage' object is not callable
2023/05/03 03:40:49 WARNING mlflow.utils.autologging_utils: Encountered unexpected error during spark autologging: Exception while attempting to initialize JVM-side state for S
park datasource autologging. Please create a new Spark session and ensure you have the mlflow-spark JAR attached to your Spark session as described in http://mlflow.org/docs/la
test/tracking.html#automatic-logging-from-spark-experimental. Exception:
'JavaPackage' object is not callable
2023/05/03 03:40:49 WARNING mlflow.utils.autologging_utils: Encountered unexpected error during spark autologging: Exception while attempting to initialize JVM-side state for S
park datasource autologging. Please create a new Spark session and ensure you have the mlflow-spark JAR attached to your Spark session as described in http://mlflow.org/docs/la
test/tracking.html#automatic-logging-from-spark-experimental. Exception:
'JavaPackage' object is not callable
2023/05/03 03:40:49 WARNING mlflow.utils.autologging_utils: Encountered unexpected error during spark autologging: Exception while attempting to initialize JVM-side state for S
park datasource autologging. Please create a new Spark session and ensure you have the mlflow-spark JAR attached to your Spark session as described in http://mlflow.org/docs/la
test/tracking.html#automatic-logging-from-spark-experimental. Exception:
'JavaPackage' object is not callable
2023/05/03 03:40:50 WARNING mlflow.utils.autologging_utils: Encountered unexpected error during spark autologging: Exception while attempting to initialize JVM-side state for S
park datasource autologging. Please create a new Spark session and ensure you have the mlflow-spark JAR attached to your Spark session as described in http://mlflow.org/docs/la
test/tracking.html#automatic-logging-from-spark-experimental. Exception:
'JavaPackage' object is not callable
2023/05/03 03:40:50 WARNING mlflow.utils.autologging_utils: Encountered unexpected error during spark autologging: Exception while attempting to initialize JVM-side state for S
park datasource autologging. Please create a new Spark session and ensure you have the mlflow-spark JAR attached to your Spark session as described in http://mlflow.org/docs/la
test/tracking.html#automatic-logging-from-spark-experimental. Exception:
'JavaPackage' object is not callable
2023/05/03 03:40:50 WARNING mlflow.utils.autologging_utils: Encountered unexpected error during spark autologging: Exception while attempting to initialize JVM-side state for S
park datasource autologging. Please create a new Spark session and ensure you have the mlflow-spark JAR attached to your Spark session as described in http://mlflow.org/docs/la
test/tracking.html#automatic-logging-from-spark-experimental. Exception:
'JavaPackage' object is not callable
2023/05/03 03:40:50 WARNING mlflow.utils.autologging_utils: Encountered unexpected error during spark autologging: Exception while attempting to initialize JVM-side state for S
park datasource autologging. Please create a new Spark session and ensure you have the mlflow-spark JAR attached to your Spark session as described in http://mlflow.org/docs/la
test/tracking.html#automatic-logging-from-spark-experimental. Exception:
'JavaPackage' object is not callable
2023/05/03 03:40:53 WARNING mlflow.utils.autologging_utils: Encountered unexpected error during spark autologging: Exception while attempting to initialize JVM-side state for S
park datasource autologging. Please create a new Spark session and ensure you have the mlflow-spark JAR attached to your Spark session as described in http://mlflow.org/docs/la
test/tracking.html#automatic-logging-from-spark-experimental. Exception:
'JavaPackage' object is not callable
2023/05/03 03:41:03 WARNING mlflow.utils.environment: Encountered an unexpected error while inferring pip requirements (model URI: /tmp/tmpkd3z3vqk, flavor: spark), fall back t
o return ['pyspark==3.0.1']. Set logging level to DEBUG to see the full traceback.
2023/05/03 03:41:03 WARNING mlflow.utils.autologging_utils: Encountered unexpected error during spark autologging: Exception while attempting to initialize JVM-side state for S
park datasource autologging. Please create a new Spark session and ensure you have the mlflow-spark JAR attached to your Spark session as described in http://mlflow.org/docs/la
test/tracking.html#automatic-logging-from-spark-experimental. Exception:
'JavaPackage' object is not callable
2023/05/03 03:41:03 WARNING mlflow.utils.autologging_utils: Encountered unexpected error during spark autologging: Exception while attempting to initialize JVM-side state for S
park datasource autologging. Please create a new Spark session and ensure you have the mlflow-spark JAR attached to your Spark session as described in http://mlflow.org/docs/la
test/tracking.html#automatic-logging-from-spark-experimental. Exception:
'JavaPackage' object is not callable
2023/05/03 03:41:03 WARNING mlflow.utils.autologging_utils: Encountered unexpected error during spark autologging: Exception while attempting to initialize JVM-side state for S
park datasource autologging. Please create a new Spark session and ensure you have the mlflow-spark JAR attached to your Spark session as described in http://mlflow.org/docs/la
test/tracking.html#automatic-logging-from-spark-experimental. Exception:

*** WARNING: skipped 153345 bytes of output ***

'JavaPackage' object is not callable
2023/05/03 03:43:16 WARNING mlflow.utils.autologging_utils: Encountered unexpected error during spark autologging: Exception while attempting to initialize JVM-side state for S
park datasource autologging. Please create a new Spark session and ensure you have the mlflow-spark JAR attached to your Spark session as described in http://mlflow.org/docs/la
test/tracking.html#automatic-logging-from-spark-experimental. Exception:
'JavaPackage' object is not callable
2023/05/03 03:43:16 WARNING mlflow.utils.autologging_utils: Encountered unexpected error during spark autologging: Exception while attempting to initialize JVM-side state for S
park datasource autologging. Please create a new Spark session and ensure you have the mlflow-spark JAR attached to your Spark session as described in http://mlflow.org/docs/la
test/tracking.html#automatic-logging-from-spark-experimental. Exception:
'JavaPackage' object is not callable
2023/05/03 03:43:16 WARNING mlflow.utils.autologging_utils: Encountered unexpected error during spark autologging: Exception while attempting to initialize JVM-side state for S
park datasource autologging. Please create a new Spark session and ensure you have the mlflow-spark JAR attached to your Spark session as described in http://mlflow.org/docs/la
test/tracking.html#automatic-logging-from-spark-experimental. Exception:
'JavaPackage' object is not callable
2023/05/03 03:43:16 WARNING mlflow.utils.autologging_utils: Encountered unexpected error during spark autologging: Exception while attempting to initialize JVM-side state for S
park datasource autologging. Please create a new Spark session and ensure you have the mlflow-spark JAR attached to your Spark session as described in http://mlflow.org/docs/la
test/tracking.html#automatic-logging-from-spark-experimental. Exception:
'JavaPackage' object is not callable
2023/05/03 03:43:16 WARNING mlflow.utils.autologging_utils: Encountered unexpected error during spark autologging: Exception while attempting to initialize JVM-side state for S
park datasource autologging. Please create a new Spark session and ensure you have the mlflow-spark JAR attached to your Spark session as described in http://mlflow.org/docs/la
```

```
test/tracking.html#automatic-logging-from-spark-experimental. Exception:
'JavaPackage' object is not callable
2023/05/03 03:43:16 WARNING mlflow.utils.autologging_utils: Encountered unexpected error during spark autologging: Exception while attempting to initialize JVM-side state for S
park datasource autologging. Please create a new Spark session and ensure you have the mlflow-spark JAR attached to your Spark session as described in http://mlflow.org/docs/la
test/tracking.html#automatic-logging-from-spark-experimental. Exception:
'JavaPackage' object is not callable
2023/05/03 03:43:16 WARNING mlflow.utils.autologging_utils: Encountered unexpected error during spark autologging: Exception while attempting to initialize JVM-side state for S
park datasource autologging. Please create a new Spark session and ensure you have the mlflow-spark JAR attached to your Spark session as described in http://mlflow.org/docs/la
test/tracking.html#automatic-logging-from-spark-experimental. Exception:
'JavaPackage' object is not callable
2023/05/03 03:43:16 WARNING mlflow.utils.autologging_utils: Encountered unexpected error during spark autologging: Exception while attempting to initialize JVM-side state for S
park datasource autologging. Please create a new Spark session and ensure you have the mlflow-spark JAR attached to your Spark session as described in http://mlflow.org/docs/la
test/tracking.html#automatic-logging-from-spark-experimental. Exception:
'JavaPackage' object is not callable
2023/05/03 03:43:16 WARNING mlflow.utils.autologging_utils: Encountered unexpected error during spark autologging: Exception while attempting to initialize JVM-side state for S
park datasource autologging. Please create a new Spark session and ensure you have the mlflow-spark JAR attached to your Spark session as described in http://mlflow.org/docs/la
test/tracking.html#automatic-logging-from-spark-experimental. Exception:
'JavaPackage' object is not callable
2023/05/03 03:43:16 WARNING mlflow.utils.autologging_utils: Encountered unexpected error during spark autologging: Exception while attempting to initialize JVM-side state for S
park datasource autologging. Please create a new Spark session and ensure you have the mlflow-spark JAR attached to your Spark session as described in http://mlflow.org/docs/la
test/tracking.html#automatic-logging-from-spark-experimental. Exception:
'JavaPackage' object is not callable
  rmse: 4.63074975564234
  regParam: 1
  elasticNetParam: 1.0
  fitIntercept: True
Inside MLflow Run with run_id 763450d095bb4a39b27f3ab61f4ca039 and experiment_id 5a58fa41f4ba496bbb8b860919a46b1a
2023/05/03 03:43:16 WARNING mlflow.utils.autologging_utils: Encountered unexpected error during spark autologging: Exception while attempting to initialize JVM-side state for S
park datasource autologging. Please create a new Spark session and ensure you have the mlflow-spark JAR attached to your Spark session as described in http://mlflow.org/docs/la
test/tracking.html#automatic-logging-from-spark-experimental. Exception:
'JavaPackage' object is not callable
2023/05/03 03:43:16 WARNING mlflow.utils.autologging_utils: Encountered unexpected error during spark autologging: Exception while attempting to initialize JVM-side state for S
park datasource autologging. Please create a new Spark session and ensure you have the mlflow-spark JAR attached to your Spark session as described in http://mlflow.org/docs/la
test/tracking.html#automatic-logging-from-spark-experimental. Exception:
'JavaPackage' object is not callable
2023/05/03 03:43:16 WARNING mlflow.utils.autologging_utils: Encountered unexpected error during spark autologging: Exception while attempting to initialize JVM-side state for S
park datasource autologging. Please create a new Spark session and ensure you have the mlflow-spark JAR attached to your Spark session as described in http://mlflow.org/docs/la
test/tracking.html#automatic-logging-from-spark-experimental. Exception:
'JavaPackage' object is not callable
2023/05/03 03:43:16 WARNING mlflow.utils.autologging_utils: Encountered unexpected error during spark autologging: Exception while attempting to initialize JVM-side state for S
park datasource autologging. Please create a new Spark session and ensure you have the mlflow-spark JAR attached to your Spark session as described in http://mlflow.org/docs/la
test/tracking.html#automatic-logging-from-spark-experimental. Exception:
'JavaPackage' object is not callable
2023/05/03 03:43:16 WARNING mlflow.utils.autologging_utils: Encountered unexpected error during spark autologging: Exception while attempting to initialize JVM-side state for S
park datasource autologging. Please create a new Spark session and ensure you have the mlflow-spark JAR attached to your Spark session as described in http://mlflow.org/docs/la
test/tracking.html#automatic-logging-from-spark-experimental. Exception:
'JavaPackage' object is not callable
2023/05/03 03:43:16 WARNING mlflow.utils.autologging_utils: Encountered unexpected error during spark autologging: Exception while attempting to initialize JVM-side state for S
park datasource autologging. Please create a new Spark session and ensure you have the mlflow-spark JAR attached to your Spark session as described in http://mlflow.org/docs/la
test/tracking.html#automatic-logging-from-spark-experimental. Exception:
'JavaPackage' object is not callable
2023/05/03 03:43:16 WARNING mlflow.utils.autologging_utils: Encountered unexpected error during spark autologging: Exception while attempting to initialize JVM-side state for S
park datasource autologging. Please create a new Spark session and ensure you have the mlflow-spark JAR attached to your Spark session as described in http://mlflow.org/docs/la
test/tracking.html#automatic-logging-from-spark-experimental. Exception:
'JavaPackage' object is not callable
2023/05/03 03:43:16 WARNING mlflow.utils.autologging_utils: Encountered unexpected error during spark autologging: Exception while attempting to initialize JVM-side state for S
park datasource autologging. Please create a new Spark session and ensure you have the mlflow-spark JAR attached to your Spark session as described in http://mlflow.org/docs/la
test/tracking.html#automatic-logging-from-spark-experimental. Exception:
'JavaPackage' object is not callable
2023/05/03 03:43:16 WARNING mlflow.utils.autologging_utils: Encountered unexpected error during spark autologging: Exception while attempting to initialize JVM-side state for S
park datasource autologging. Please create a new Spark session and ensure you have the mlflow-spark JAR attached to your Spark session as described in http://mlflow.org/docs/la
test/tracking.html#automatic-logging-from-spark-experimental. Exception:
'JavaPackage' object is not callable
2023/05/03 03:43:18 WARNING mlflow.utils.autologging_utils: Encountered unexpected error during spark autologging: Exception while attempting to initialize JVM-side state for S
park datasource autologging. Please create a new Spark session and ensure you have the mlflow-spark JAR attached to your Spark session as described in http://mlflow.org/docs/la
test/tracking.html#automatic-logging-from-spark-experimental. Exception:
'JavaPackage' object is not callable
2023/05/03 03:43:18 WARNING mlflow.utils.autologging_utils: Encountered unexpected error during spark autologging: Exception while attempting to initialize JVM-side state for S
park datasource autologging. Please create a new Spark session and ensure you have the mlflow-spark JAR attached to your Spark session as described in http://mlflow.org/docs/la
test/tracking.html#automatic-logging-from-spark-experimental. Exception:
'JavaPackage' object is not callable
```

Based on the result of tuning pyperparameters, we choose elastic-net models that has elasticnet parameter value = 0 or 0.1 or 0.2, since they has the lowest RMSE value.

```
'JavaPackage' object is not callable
```

Grid Search and Tuning hyperparameters (MLflow)

```
with mlflow.start_run(run_name="Elastic-Net-Regression-Model") as run:
    ### Feature Extraction
    featureCols = ["laps", "fastestLap", "fastestLapSpeed", "statusId", "driverId", "raceId"]
    assembler = VectorAssembler(inputCols=featureCols, outputCol="features")
    assembled_df = assembler.transform(df2)
    ### Prepare training and test data
    train, test = assembled_df.randomSplit([0.9, 0.1], seed=12345)
    ### Fit Model
    ### Generate Predictions
    lr = LinearRegression(featuresCol="features",labelCol='positionOrder',predictionCol='predpositionOrder',maxIter=10, regParam=1, elasticNetParam=1,standardization=False)
    ### Grid Search and Tuning hyperparameters
    paramGrid = ParamGridBuilder()\
        .addGrid(lr.regParam, [0.01, 0.1, 1, 10]) \
        .addGrid(lr.fitIntercept, [False, True])\
        .addGrid(lr.elasticNetParam, [0.0,0.1,0.2,0.3,0.4,0.5,0.6,0.7,0.8,0.9,1.0])\
        .build()
    evaluator=RegressionEvaluator(labelCol='positionOrder',predictionCol='predpositionOrder')
    tvs = TrainValidationSplit(estimator=lr,
                               estimatorParamMaps=paramGrid,
                               evaluator=evaluator,
                               # 90% of the data will be used for training, 10% for validation.
                               trainRatio=0.9,
                               collectSubModels=True)

    model1 = tvs.fit(train)
    mlflow.spark.log_model(model1, "Elastic-Net-Regression-Model")
    predtest = model1.transform(test)
    evaluator = RegressionEvaluator(labelCol='positionOrder',predictionCol='predpositionOrder')
    rmse = evaluator.evaluate(predtest, {evaluator.metricName: "rmse"})
    mlflow.log_metric("rmse", rmse)
  # print("   paramGrid: {}".format(paramGrid))
    #mlflow.log_param("paramGrid",paramGrid)
    print("   rmse: {}".format(rmse))

    runID = run.info.run_uuid # uuid: the unique Id this runID is against
    experimentID = run.info.experiment_id
    print("Inside MLflow Run with run_id {} and experiment_id {}".format(runID, experimentID))
```

```
2023/05/03 03:21:26 WARNING mlflow.utils.autologging_utils: Encountered unexpected error during spark autologging: Exception while attempting to initialize JVM-side state for S
park datasource autologging. Please create a new Spark session and ensure you have the mlflow-spark JAR attached to your Spark session as described in http://mlflow.org/docs/la
test/tracking.html#automatic-logging-from-spark-experimental. Exception:
'JavaPackage' object is not callable
2023/05/03 03:21:26 WARNING mlflow.utils.autologging_utils: Encountered unexpected error during spark autologging: Exception while attempting to initialize JVM-side state for S
park datasource autologging. Please create a new Spark session and ensure you have the mlflow-spark JAR attached to your Spark session as described in http://mlflow.org/docs/la
test/tracking.html#automatic-logging-from-spark-experimental. Exception:
'JavaPackage' object is not callable
2023/05/03 03:21:26 WARNING mlflow.utils.autologging_utils: Encountered unexpected error during spark autologging: Exception while attempting to initialize JVM-side state for S
park datasource autologging. Please create a new Spark session and ensure you have the mlflow-spark JAR attached to your Spark session as described in http://mlflow.org/docs/la
test/tracking.html#automatic-logging-from-spark-experimental. Exception:
'JavaPackage' object is not callable
2023/05/03 03:21:26 WARNING mlflow.utils.autologging_utils: Encountered unexpected error during spark autologging: Exception while attempting to initialize JVM-side state for S
park datasource autologging. Please create a new Spark session and ensure you have the mlflow-spark JAR attached to your Spark session as described in http://mlflow.org/docs/la
test/tracking.html#automatic-logging-from-spark-experimental. Exception:
'JavaPackage' object is not callable
2023/05/03 03:21:26 WARNING mlflow.utils.autologging_utils: Encountered unexpected error during spark autologging: Exception while attempting to initialize JVM-side state for S
park datasource autologging. Please create a new Spark session and ensure you have the mlflow-spark JAR attached to your Spark session as described in http://mlflow.org/docs/la
test/tracking.html#automatic-logging-from-spark-experimental. Exception:
'JavaPackage' object is not callable
2023/05/03 03:21:26 WARNING mlflow.utils.autologging_utils: Encountered unexpected error during spark autologging: Exception while attempting to initialize JVM-side state for S
park datasource autologging. Please create a new Spark session and ensure you have the mlflow-spark JAR attached to your Spark session as described in http://mlflow.org/docs/la
test/tracking.html#automatic-logging-from-spark-experimental. Exception:
'JavaPackage' object is not callable
2023/05/03 03:21:27 WARNING mlflow.utils.autologging_utils: Encountered unexpected error during spark autologging: Exception while attempting to initialize JVM-side state for S
park datasource autologging. Please create a new Spark session and ensure you have the mlflow-spark JAR attached to your Spark session as described in http://mlflow.org/docs/la
test/tracking.html#automatic-logging-from-spark-experimental. Exception:
'JavaPackage' object is not callable
2023/05/03 03:21:27 WARNING mlflow.utils.autologging_utils: Encountered unexpected error during spark autologging: Exception while attempting to initialize JVM-side state for S
park datasource autologging. Please create a new Spark session and ensure you have the mlflow-spark JAR attached to your Spark session as described in http://mlflow.org/docs/la
test/tracking.html#automatic-logging-from-spark-experimental. Exception:
'JavaPackage' object is not callable
2023/05/03 03:21:27 WARNING mlflow.utils.autologging_utils: Encountered unexpected error during spark autologging: Exception while attempting to initialize JVM-side state for S
park datasource autologging. Please create a new Spark session and ensure you have the mlflow-spark JAR attached to your Spark session as described in http://mlflow.org/docs/la
test/tracking.html#automatic-logging-from-spark-experimental. Exception:
'JavaPackage' object is not callable
MLlib will automatically track trials in MLflow. After your tuning fit() call has completed, view the MLflow UI to see logged runs.
2023/05/03 03:21:30 WARNING mlflow.utils.autologging_utils: Encountered unexpected error during spark autologging: Exception while attempting to initialize JVM-side state for S
park datasource autologging. Please create a new Spark session and ensure you have the mlflow-spark JAR attached to your Spark session as described in http://mlflow.org/docs/la
test/tracking.html#automatic-logging-from-spark-experimental. Exception:
'JavaPackage' object is not callable
2023/05/03 03:21:30 WARNING mlflow.utils.autologging_utils: Encountered unexpected error during spark autologging: Exception while attempting to initialize JVM-side state for S
park datasource autologging. Please create a new Spark session and ensure you have the mlflow-spark JAR attached to your Spark session as described in http://mlflow.org/docs/la
test/tracking.html#automatic-logging-from-spark-experimental. Exception:
'JavaPackage' object is not callable
2023/05/03 03:21:30 WARNING mlflow.utils.autologging_utils: Encountered unexpected error during spark autologging: Exception while attempting to initialize JVM-side state for S
park datasource autologging. Please create a new Spark session and ensure you have the mlflow-spark JAR attached to your Spark session as described in http://mlflow.org/docs/la
test/tracking.html#automatic-logging-from-spark-experimental. Exception:
'JavaPackage' object is not callable
2023/05/03 03:21:31 WARNING mlflow.utils.autologging_utils: Encountered unexpected error during spark autologging: Exception while attempting to initialize JVM-side state for S
park datasource autologging. Please create a new Spark session and ensure you have the mlflow-spark JAR attached to your Spark session as described in http://mlflow.org/docs/la
test/tracking.html#automatic-logging-from-spark-experimental. Exception:
'JavaPackage' object is not callable
2023/05/03 03:21:31 WARNING mlflow.utils.autologging_utils: Encountered unexpected error during spark autologging: Exception while attempting to initialize JVM-side state for S
park datasource autologging. Please create a new Spark session and ensure you have the mlflow-spark JAR attached to your Spark session as described in http://mlflow.org/docs/la
test/tracking.html#automatic-logging-from-spark-experimental. Exception:
'JavaPackage' object is not callable
2023/05/03 03:21:31 WARNING mlflow.utils.autologging_utils: Encountered unexpected error during spark autologging: Exception while attempting to initialize JVM-side state for S
park datasource autologging. Please create a new Spark session and ensure you have the mlflow-spark JAR attached to your Spark session as described in http://mlflow.org/docs/la
test/tracking.html#automatic-logging-from-spark-experimental. Exception:
'JavaPackage' object is not callable
2023/05/03 03:21:31 WARNING mlflow.utils.autologging_utils: Encountered unexpected error during spark autologging: Exception while attempting to initialize JVM-side state for S
park datasource autologging. Please create a new Spark session and ensure you have the mlflow-spark JAR attached to your Spark session as described in http://mlflow.org/docs/la
test/tracking.html#automatic-logging-from-spark-experimental. Exception:
'JavaPackage' object is not callable
2023/05/03 03:21:31 WARNING mlflow.utils.autologging_utils: Encountered unexpected error during spark autologging: Exception while attempting to initialize JVM-side state for S
park datasource autologging. Please create a new Spark session and ensure you have the mlflow-spark JAR attached to your Spark session as described in http://mlflow.org/docs/la
test/tracking.html#automatic-logging-from-spark-experimental. Exception:
'JavaPackage' object is not callable
2023/05/03 03:21:31 WARNING mlflow.utils.autologging_utils: Encountered unexpected error during spark autologging: Exception while attempting to initialize JVM-side state for S
park datasource autologging. Please create a new Spark session and ensure you have the mlflow-spark JAR attached to your Spark session as described in http://mlflow.org/docs/la
test/tracking.html#automatic-logging-from-spark-experimental. Exception:
'JavaPackage' object is not callable
2023/05/03 03:21:31 WARNING mlflow.utils.autologging_utils: Encountered unexpected error during spark autologging: Exception while attempting to initialize JVM-side state for S
park datasource autologging. Please create a new Spark session and ensure you have the mlflow-spark JAR attached to your Spark session as described in http://mlflow.org/docs/la
test/tracking.html#automatic-logging-from-spark-experimental. Exception:
'JavaPackage' object is not callable
2023/05/03 03:21:31 WARNING mlflow.utils.autologging_utils: Encountered unexpected error during spark autologging: Exception while attempting to initialize JVM-side state for S
park datasource autologging. Please create a new Spark session and ensure you have the mlflow-spark JAR attached to your Spark session as described in http://mlflow.org/docs/la
test/tracking.html#automatic-logging-from-spark-experimental. Exception:
'JavaPackage' object is not callable
2023/05/03 03:21:31 WARNING mlflow.utils.autologging_utils: Encountered unexpected error during spark autologging: Exception while attempting to initialize JVM-side state for S
park datasource autologging. Please create a new Spark session and ensure you have the mlflow-spark JAR attached to your Spark session as described in http://mlflow.org/docs/la
test/tracking.html#automatic-logging-from-spark-experimental. Exception:
'JavaPackage' object is not callable
2023/05/03 03:21:31 WARNING mlflow.utils.autologging_utils: Encountered unexpected error during spark autologging: Exception while attempting to initialize JVM-side state for S
park datasource autologging. Please create a new Spark session and ensure you have the mlflow-spark JAR attached to your Spark session as described in http://mlflow.org/docs/la
test/tracking.html#automatic-logging-from-spark-experimental. Exception:
'JavaPackage' object is not callable
2023/05/03 03:21:31 WARNING mlflow.utils.autologging_utils: Encountered unexpected error during spark autologging: Exception while attempting to initialize JVM-side state for S
park datasource autologging. Please create a new Spark session and ensure you have the mlflow-spark JAR attached to your Spark session as described in http://mlflow.org/docs/la
test/tracking.html#automatic-logging-from-spark-experimental. Exception:
'JavaPackage' object is not callable
2023/05/03 03:21:31 WARNING mlflow.utils.autologging_utils: Encountered unexpected error during spark autologging: Exception while attempting to initialize JVM-side state for S
park datasource autologging. Please create a new Spark session and ensure you have the mlflow-spark JAR attached to your Spark session as described in http://mlflow.org/docs/la
test/tracking.html#automatic-logging-from-spark-experimental. Exception:
'JavaPackage' object is not callable
2023/05/03 03:21:31 WARNING mlflow.utils.autologging_utils: Encountered unexpected error during spark autologging: Exception while attempting to initialize JVM-side state for S
park datasource autologging. Please create a new Spark session and ensure you have the mlflow-spark JAR attached to your Spark session as described in http://mlflow.org/docs/la
test/tracking.html#automatic-logging-from-spark-experimental. Exception:
'JavaPackage' object is not callable
2023/05/03 03:21:31 WARNING mlflow.utils.autologging_utils: Encountered unexpected error during spark autologging: Exception while attempting to initialize JVM-side state for S
park datasource autologging. Please create a new Spark session and ensure you have the mlflow-spark JAR attached to your Spark session as described in http://mlflow.org/docs/la
test/tracking.html#automatic-logging-from-spark-experimental. Exception:
'JavaPackage' object is not callable
2023/05/03 03:21:31 WARNING mlflow.utils.autologging_utils: Encountered unexpected error during spark autologging: Exception while attempting to initialize JVM-side state for S
park datasource autologging. Please create a new Spark session and ensure you have the mlflow-spark JAR attached to your Spark session as described in http://mlflow.org/docs/la
test/tracking.html#automatic-logging-from-spark-experimental. Exception:
'JavaPackage' object is not callable
2023/05/03 03:21:31 WARNING mlflow.utils.autologging_utils: Encountered unexpected error during spark autologging: Exception while attempting to initialize JVM-side state for S
park datasource autologging. Please create a new Spark session and ensure you have the mlflow-spark JAR attached to your Spark session as described in http://mlflow.org/docs/la
test/tracking.html#automatic-logging-from-spark-experimental. Exception:
'JavaPackage' object is not callable
2023/05/03 03:21:31 WARNING mlflow.utils.autologging_utils: Encountered unexpected error during spark autologging: Exception while attempting to initialize JVM-side state for S
park datasource autologging. Please create a new Spark session and ensure you have the mlflow-spark JAR attached to your Spark session as described in http://mlflow.org/docs/la
test/tracking.html#automatic-logging-from-spark-experimental. Exception:
'JavaPackage' object is not callable
2023/05/03 03:21:31 WARNING mlflow.utils.autologging_utils: Encountered unexpected error during spark autologging: Exception while attempting to initialize JVM-side state for S
park datasource autologging. Please create a new Spark session and ensure you have the mlflow-spark JAR attached to your Spark session as described in http://mlflow.org/docs/la
test/tracking.html#automatic-logging-from-spark-experimental. Exception:
'JavaPackage' object is not callable
2023/05/03 03:21:31 WARNING mlflow.utils.autologging_utils: Encountered unexpected error during spark autologging: Exception while attempting to initialize JVM-side state for S
park datasource autologging. Please create a new Spark session and ensure you have the mlflow-spark JAR attached to your Spark session as described in http://mlflow.org/docs/la
test/tracking.html#automatic-logging-from-spark-experimental. Exception:
'JavaPackage' object is not callable
```

```
2023/05/03 03:21:31 WARNING mlflow.utils.autologging_utils: Encountered unexpected error during spark autologging: Exception while attempting to initialize JVM-side state for S
park datasource autologging. Please create a new Spark session and ensure you have the mlflow-spark JAR attached to your Spark session as described in http://mlflow.org/docs/la
test/tracking.html#automatic-logging-from-spark-experimental. Exception:
'JavaPackage' object is not callable
2023/05/03 03:21:31 WARNING mlflow.utils.autologging_utils: Encountered unexpected error during spark autologging: Exception while attempting to initialize JVM-side state for S
park datasource autologging. Please create a new Spark session and ensure you have the mlflow-spark JAR attached to your Spark session as described in http://mlflow.org/docs/la
test/tracking.html#automatic-logging-from-spark-experimental. Exception:
'JavaPackage' object is not callable
2023/05/03 03:21:31 WARNING mlflow.utils.autologging_utils: Encountered unexpected error during spark autologging: Exception while attempting to initialize JVM-side state for S
park datasource autologging. Please create a new Spark session and ensure you have the mlflow-spark JAR attached to your Spark session as described in http://mlflow.org/docs/la
test/tracking.html#automatic-logging-from-spark-experimental. Exception:
'JavaPackage' object is not callable
2023/05/03 03:21:31 WARNING mlflow.utils.autologging_utils: Encountered unexpected error during spark autologging: Exception while attempting to initialize JVM-side state for S
park datasource autologging. Please create a new Spark session and ensure you have the mlflow-spark JAR attached to your Spark session as described in http://mlflow.org/docs/la
test/tracking.html#automatic-logging-from-spark-experimental. Exception:
'JavaPackage' object is not callable
2023/05/03 03:21:31 WARNING mlflow.utils.autologging_utils: Encountered unexpected error during spark autologging: Exception while attempting to initialize JVM-side state for S
park datasource autologging. Please create a new Spark session and ensure you have the mlflow-spark JAR attached to your Spark session as described in http://mlflow.org/docs/la
test/tracking.html#automatic-logging-from-spark-experimental. Exception:
'JavaPackage' object is not callable
2023/05/03 03:21:31 WARNING mlflow.utils.autologging_utils: Encountered unexpected error during spark autologging: Exception while attempting to initialize JVM-side state for S
park datasource autologging. Please create a new Spark session and ensure you have the mlflow-spark JAR attached to your Spark session as described in http://mlflow.org/docs/la
test/tracking.html#automatic-logging-from-spark-experimental. Exception:
'JavaPackage' object is not callable
2023/05/03 03:21:32 WARNING mlflow.utils.autologging_utils: Encountered unexpected error during spark autologging: Exception while attempting to initialize JVM-side state for S
park datasource autologging. Please create a new Spark session and ensure you have the mlflow-spark JAR attached to your Spark session as described in http://mlflow.org/docs/la
test/tracking.html#automatic-logging-from-spark-experimental. Exception:
'JavaPackage' object is not callable
2023/05/03 03:21:32 WARNING mlflow.utils.autologging_utils: Encountered unexpected error during spark autologging: Exception while attempting to initialize JVM-side state for S
park datasource autologging. Please create a new Spark session and ensure you have the mlflow-spark JAR attached to your Spark session as described in http://mlflow.org/docs/la
test/tracking.html#automatic-logging-from-spark-experimental. Exception:
'JavaPackage' object is not callable
2023/05/03 03:21:32 WARNING mlflow.utils.autologging_utils: Encountered unexpected error during spark autologging: Exception while attempting to initialize JVM-side state for S
park datasource autologging. Please create a new Spark session and ensure you have the mlflow-spark JAR attached to your Spark session as described in http://mlflow.org/docs/la
test/tracking.html#automatic-logging-from-spark-experimental. Exception:
'JavaPackage' object is not callable
2023/05/03 03:21:32 WARNING mlflow.utils.autologging_utils: Encountered unexpected error during spark autologging: Exception while attempting to initialize JVM-side state for S
park datasource autologging. Please create a new Spark session and ensure you have the mlflow-spark JAR attached to your Spark session as described in http://mlflow.org/docs/la
test/tracking.html#automatic-logging-from-spark-experimental. Exception:
'JavaPackage' object is not callable
2023/05/03 03:21:32 WARNING mlflow.utils.autologging_utils: Encountered unexpected error during spark autologging: Exception while attempting to initialize JVM-side state for S
park datasource autologging. Please create a new Spark session and ensure you have the mlflow-spark JAR attached to your Spark session as described in http://mlflow.org/docs/la
test/tracking.html#automatic-logging-from-spark-experimental. Exception:
'JavaPackage' object is not callable
2023/05/03 03:21:32 WARNING mlflow.utils.autologging_utils: Encountered unexpected error during spark autologging: Exception while attempting to initialize JVM-side state for S
park datasource autologging. Please create a new Spark session and ensure you have the mlflow-spark JAR attached to your Spark session as described in http://mlflow.org/docs/la
test/tracking.html#automatic-logging-from-spark-experimental. Exception:
'JavaPackage' object is not callable
2023/05/03 03:21:32 WARNING mlflow.utils.autologging_utils: Encountered unexpected error during spark autologging: Exception while attempting to initialize JVM-side state for S
park datasource autologging. Please create a new Spark session and ensure you have the mlflow-spark JAR attached to your Spark session as described in http://mlflow.org/docs/la
test/tracking.html#automatic-logging-from-spark-experimental. Exception:
'JavaPackage' object is not callable
2023/05/03 03:21:32 WARNING mlflow.utils.autologging_utils: Encountered unexpected error during spark autologging: Exception while attempting to initialize JVM-side state for S
park datasource autologging. Please create a new Spark session and ensure you have the mlflow-spark JAR attached to your Spark session as described in http://mlflow.org/docs/la
test/tracking.html#automatic-logging-from-spark-experimental. Exception:
'JavaPackage' object is not callable
2023/05/03 03:21:32 WARNING mlflow.utils.autologging_utils: Encountered unexpected error during spark autologging: Exception while attempting to initialize JVM-side state for S
park datasource autologging. Please create a new Spark session and ensure you have the mlflow-spark JAR attached to your Spark session as described in http://mlflow.org/docs/la
test/tracking.html#automatic-logging-from-spark-experimental. Exception:
'JavaPackage' object is not callable
2023/05/03 03:21:32 WARNING mlflow.utils.autologging_utils: Encountered unexpected error during spark autologging: Exception while attempting to initialize JVM-side state for S
park datasource autologging. Please create a new Spark session and ensure you have the mlflow-spark JAR attached to your Spark session as described in http://mlflow.org/docs/la
test/tracking.html#automatic-logging-from-spark-experimental. Exception:
'JavaPackage' object is not callable
2023/05/03 03:21:32 WARNING mlflow.utils.autologging_utils: Encountered unexpected error during spark autologging: Exception while attempting to initialize JVM-side state for S
park datasource autologging. Please create a new Spark session and ensure you have the mlflow-spark JAR attached to your Spark session as described in http://mlflow.org/docs/la
test/tracking.html#automatic-logging-from-spark-experimental. Exception:
'JavaPackage' object is not callable
2023/05/03 03:21:32 WARNING mlflow.utils.autologging_utils: Encountered unexpected error during spark autologging: Exception while attempting to initialize JVM-side state for S
park datasource autologging. Please create a new Spark session and ensure you have the mlflow-spark JAR attached to your Spark session as described in http://mlflow.org/docs/la
test/tracking.html#automatic-logging-from-spark-experimental. Exception:
'JavaPackage' object is not callable
2023/05/03 03:21:32 WARNING mlflow.utils.autologging_utils: Encountered unexpected error during spark autologging: Exception while attempting to initialize JVM-side state for S
park datasource autologging. Please create a new Spark session and ensure you have the mlflow-spark JAR attached to your Spark session as described in http://mlflow.org/docs/la
test/tracking.html#automatic-logging-from-spark-experimental. Exception:
'JavaPackage' object is not callable
2023/05/03 03:21:32 WARNING mlflow.utils.autologging_utils: Encountered unexpected error during spark autologging: Exception while attempting to initialize JVM-side state for S
park datasource autologging. Please create a new Spark session and ensure you have the mlflow-spark JAR attached to your Spark session as described in http://mlflow.org/docs/la
test/tracking.html#automatic-logging-from-spark-experimental. Exception:
'JavaPackage' object is not callable
2023/05/03 03:21:32 WARNING mlflow.utils.autologging_utils: Encountered unexpected error during spark autologging: Exception while attempting to initialize JVM-side state for S
park datasource autologging. Please create a new Spark session and ensure you have the mlflow-spark JAR attached to your Spark session as described in http://mlflow.org/docs/la
test/tracking.html#automatic-logging-from-spark-experimental. Exception:
'JavaPackage' object is not callable
2023/05/03 03:21:32 WARNING mlflow.utils.autologging_utils: Encountered unexpected error during spark autologging: Exception while attempting to initialize JVM-side state for S
park datasource autologging. Please create a new Spark session and ensure you have the mlflow-spark JAR attached to your Spark session as described in http://mlflow.org/docs/la
test/tracking.html#automatic-logging-from-spark-experimental. Exception:
'JavaPackage' object is not callable
2023/05/03 03:21:32 WARNING mlflow.utils.autologging_utils: Encountered unexpected error during spark autologging: Exception while attempting to initialize JVM-side state for S
park datasource autologging. Please create a new Spark session and ensure you have the mlflow-spark JAR attached to your Spark session as described in http://mlflow.org/docs/la
test/tracking.html#automatic-logging-from-spark-experimental. Exception:
'JavaPackage' object is not callable
2023/05/03 03:21:32 WARNING mlflow.utils.autologging_utils: Encountered unexpected error during spark autologging: Exception while attempting to initialize JVM-side state for S
park datasource autologging. Please create a new Spark session and ensure you have the mlflow-spark JAR attached to your Spark session as described in http://mlflow.org/docs/la
test/tracking.html#automatic-logging-from-spark-experimental. Exception:

*** WARNING: skipped 70686 bytes of output ***

'JavaPackage' object is not callable
2023/05/03 03:21:39 WARNING mlflow.utils.autologging_utils: Encountered unexpected error during spark autologging: Exception while attempting to initialize JVM-side state for S
park datasource autologging. Please create a new Spark session and ensure you have the mlflow-spark JAR attached to your Spark session as described in http://mlflow.org/docs/la
test/tracking.html#automatic-logging-from-spark-experimental. Exception:
'JavaPackage' object is not callable
2023/05/03 03:21:40 WARNING mlflow.utils.autologging_utils: Encountered unexpected error during spark autologging: Exception while attempting to initialize JVM-side state for S
park datasource autologging. Please create a new Spark session and ensure you have the mlflow-spark JAR attached to your Spark session as described in http://mlflow.org/docs/la
test/tracking.html#automatic-logging-from-spark-experimental. Exception:
'JavaPackage' object is not callable
2023/05/03 03:21:40 WARNING mlflow.utils.autologging_utils: Encountered unexpected error during spark autologging: Exception while attempting to initialize JVM-side state for S
park datasource autologging. Please create a new Spark session and ensure you have the mlflow-spark JAR attached to your Spark session as described in http://mlflow.org/docs/la
test/tracking.html#automatic-logging-from-spark-experimental. Exception:
'JavaPackage' object is not callable
2023/05/03 03:21:40 WARNING mlflow.utils.autologging_utils: Encountered unexpected error during spark autologging: Exception while attempting to initialize JVM-side state for S
park datasource autologging. Please create a new Spark session and ensure you have the mlflow-spark JAR attached to your Spark session as described in http://mlflow.org/docs/la
test/tracking.html#automatic-logging-from-spark-experimental. Exception:
'JavaPackage' object is not callable
2023/05/03 03:21:40 WARNING mlflow.utils.autologging_utils: Encountered unexpected error during spark autologging: Exception while attempting to initialize JVM-side state for S
park datasource autologging. Please create a new Spark session and ensure you have the mlflow-spark JAR attached to your Spark session as described in http://mlflow.org/docs/la
test/tracking.html#automatic-logging-from-spark-experimental. Exception:
'JavaPackage' object is not callable
2023/05/03 03:21:40 WARNING mlflow.utils.autologging_utils: Encountered unexpected error during spark autologging: Exception while attempting to initialize JVM-side state for S
park datasource autologging. Please create a new Spark session and ensure you have the mlflow-spark JAR attached to your Spark session as described in http://mlflow.org/docs/la
```

```
test/tracking.html#automatic-logging-from-spark-experimental. Exception:
'JavaPackage' object is not callable
2023/05/03 03:21:40 WARNING mlflow.utils.autologging_utils: Encountered unexpected error during spark autologging: Exception while attempting to initialize JVM-side state for S
park datasource autologging. Please create a new Spark session and ensure you have the mlflow-spark JAR attached to your Spark session as described in http://mlflow.org/docs/la
test/tracking.html#automatic-logging-from-spark-experimental. Exception:
'JavaPackage' object is not callable
2023/05/03 03:21:40 WARNING mlflow.utils.autologging_utils: Encountered unexpected error during spark autologging: Exception while attempting to initialize JVM-side state for S
park datasource autologging. Please create a new Spark session and ensure you have the mlflow-spark JAR attached to your Spark session as described in http://mlflow.org/docs/la
test/tracking.html#automatic-logging-from-spark-experimental. Exception:
'JavaPackage' object is not callable
2023/05/03 03:21:40 WARNING mlflow.utils.autologging_utils: Encountered unexpected error during spark autologging: Exception while attempting to initialize JVM-side state for S
park datasource autologging. Please create a new Spark session and ensure you have the mlflow-spark JAR attached to your Spark session as described in http://mlflow.org/docs/la
test/tracking.html#automatic-logging-from-spark-experimental. Exception:
'JavaPackage' object is not callable
2023/05/03 03:21:40 WARNING mlflow.utils.autologging_utils: Encountered unexpected error during spark autologging: Exception while attempting to initialize JVM-side state for S
park datasource autologging. Please create a new Spark session and ensure you have the mlflow-spark JAR attached to your Spark session as described in http://mlflow.org/docs/la
test/tracking.html#automatic-logging-from-spark-experimental. Exception:
'JavaPackage' object is not callable
2023/05/03 03:21:40 WARNING mlflow.utils.autologging_utils: Encountered unexpected error during spark autologging: Exception while attempting to initialize JVM-side state for S
park datasource autologging. Please create a new Spark session and ensure you have the mlflow-spark JAR attached to your Spark session as described in http://mlflow.org/docs/la
test/tracking.html#automatic-logging-from-spark-experimental. Exception:
'JavaPackage' object is not callable
2023/05/03 03:21:40 WARNING mlflow.utils.autologging_utils: Encountered unexpected error during spark autologging: Exception while attempting to initialize JVM-side state for S
park datasource autologging. Please create a new Spark session and ensure you have the mlflow-spark JAR attached to your Spark session as described in http://mlflow.org/docs/la
test/tracking.html#automatic-logging-from-spark-experimental. Exception:
'JavaPackage' object is not callable
2023/05/03 03:21:40 WARNING mlflow.utils.autologging_utils: Encountered unexpected error during spark autologging: Exception while attempting to initialize JVM-side state for S
park datasource autologging. Please create a new Spark session and ensure you have the mlflow-spark JAR attached to your Spark session as described in http://mlflow.org/docs/la
test/tracking.html#automatic-logging-from-spark-experimental. Exception:
'JavaPackage' object is not callable
2023/05/03 03:21:40 WARNING mlflow.utils.autologging_utils: Encountered unexpected error during spark autologging: Exception while attempting to initialize JVM-side state for S
park datasource autologging. Please create a new Spark session and ensure you have the mlflow-spark JAR attached to your Spark session as described in http://mlflow.org/docs/la
test/tracking.html#automatic-logging-from-spark-experimental. Exception:
'JavaPackage' object is not callable
2023/05/03 03:21:40 WARNING mlflow.utils.autologging_utils: Encountered unexpected error during spark autologging: Exception while attempting to initialize JVM-side state for S
park datasource autologging. Please create a new Spark session and ensure you have the mlflow-spark JAR attached to your Spark session as described in http://mlflow.org/docs/la
test/tracking.html#automatic-logging-from-spark-experimental. Exception:
'JavaPackage' object is not callable
2023/05/03 03:21:40 WARNING mlflow.utils.autologging_utils: Encountered unexpected error during spark autologging: Exception while attempting to initialize JVM-side state for S
park datasource autologging. Please create a new Spark session and ensure you have the mlflow-spark JAR attached to your Spark session as described in http://mlflow.org/docs/la
test/tracking.html#automatic-logging-from-spark-experimental. Exception:
'JavaPackage' object is not callable
2023/05/03 03:21:40 WARNING mlflow.utils.autologging_utils: Encountered unexpected error during spark autologging: Exception while attempting to initialize JVM-side state for S
park datasource autologging. Please create a new Spark session and ensure you have the mlflow-spark JAR attached to your Spark session as described in http://mlflow.org/docs/la
test/tracking.html#automatic-logging-from-spark-experimental. Exception:
'JavaPackage' object is not callable
2023/05/03 03:21:40 WARNING mlflow.utils.autologging_utils: Encountered unexpected error during spark autologging: Exception while attempting to initialize JVM-side state for S
park datasource autologging. Please create a new Spark session and ensure you have the mlflow-spark JAR attached to your Spark session as described in http://mlflow.org/docs/la
test/tracking.html#automatic-logging-from-spark-experimental. Exception:
'JavaPackage' object is not callable
2023/05/03 03:21:40 WARNING mlflow.utils.autologging_utils: Encountered unexpected error during spark autologging: Exception while attempting to initialize JVM-side state for S
park datasource autologging. Please create a new Spark session and ensure you have the mlflow-spark JAR attached to your Spark session as described in http://mlflow.org/docs/la
test/tracking.html#automatic-logging-from-spark-experimental. Exception:
'JavaPackage' object is not callable
2023/05/03 03:21:40 WARNING mlflow.utils.autologging_utils: Encountered unexpected error during spark autologging: Exception while attempting to initialize JVM-side state for S
park datasource autologging. Please create a new Spark session and ensure you have the mlflow-spark JAR attached to your Spark session as described in http://mlflow.org/docs/la
test/tracking.html#automatic-logging-from-spark-experimental. Exception:
'JavaPackage' object is not callable
2023/05/03 03:21:40 WARNING mlflow.utils.autologging_utils: Encountered unexpected error during spark autologging: Exception while attempting to initialize JVM-side state for S
park datasource autologging. Please create a new Spark session and ensure you have the mlflow-spark JAR attached to your Spark session as described in http://mlflow.org/docs/la
test/tracking.html#automatic-logging-from-spark-experimental. Exception:
'JavaPackage' object is not callable
2023/05/03 03:21:40 WARNING mlflow.utils.autologging_utils: Encountered unexpected error during spark autologging: Exception while attempting to initialize JVM-side state for S
park datasource autologging. Please create a new Spark session and ensure you have the mlflow-spark JAR attached to your Spark session as described in http://mlflow.org/docs/la
test/tracking.html#automatic-logging-from-spark-experimental. Exception:
'JavaPackage' object is not callable
2023/05/03 03:21:58 WARNING mlflow.utils.autologging_utils: Encountered unexpected error during spark autologging: Exception while attempting to initialize JVM-side state for S
park datasource autologging. Please create a new Spark session and ensure you have the mlflow-spark JAR attached to your Spark session as described in http://mlflow.org/docs/la
test/tracking.html#automatic-logging-from-spark-experimental. Exception:
'JavaPackage' object is not callable
2023/05/03 03:22:08 WARNING mlflow.utils.environment: Encountered an unexpected error while inferring pip requirements (model URI: /tmp/tmp8c51mdbb, flavor: spark), fall back t
o return ['pyspark==3.0.1']. Set logging level to DEBUG to see the full traceback.
2023/05/03 03:22:08 WARNING mlflow.utils.autologging_utils: Encountered unexpected error during spark autologging: Exception while attempting to initialize JVM-side state for S
park datasource autologging. Please create a new Spark session and ensure you have the mlflow-spark JAR attached to your Spark session as described in http://mlflow.org/docs/la
test/tracking.html#automatic-logging-from-spark-experimental. Exception:
'JavaPackage' object is not callable
2023/05/03 03:22:08 WARNING mlflow.utils.autologging_utils: Encountered unexpected error during spark autologging: Exception while attempting to initialize JVM-side state for S
park datasource autologging. Please create a new Spark session and ensure you have the mlflow-spark JAR attached to your Spark session as described in http://mlflow.org/docs/la
test/tracking.html#automatic-logging-from-spark-experimental. Exception:
'JavaPackage' object is not callable
2023/05/03 03:22:08 WARNING mlflow.utils.autologging_utils: Encountered unexpected error during spark autologging: Exception while attempting to initialize JVM-side state for S
park datasource autologging. Please create a new Spark session and ensure you have the mlflow-spark JAR attached to your Spark session as described in http://mlflow.org/docs/la
test/tracking.html#automatic-logging-from-spark-experimental. Exception:
'JavaPackage' object is not callable
2023/05/03 03:22:08 WARNING mlflow.utils.autologging_utils: Encountered unexpected error during spark autologging: Exception while attempting to initialize JVM-side state for S
park datasource autologging. Please create a new Spark session and ensure you have the mlflow-spark JAR attached to your Spark session as described in http://mlflow.org/docs/la
test/tracking.html#automatic-logging-from-spark-experimental. Exception:
'JavaPackage' object is not callable
2023/05/03 03:22:08 WARNING mlflow.utils.autologging_utils: Encountered unexpected error during spark autologging: Exception while attempting to initialize JVM-side state for S
park datasource autologging. Please create a new Spark session and ensure you have the mlflow-spark JAR attached to your Spark session as described in http://mlflow.org/docs/la
test/tracking.html#automatic-logging-from-spark-experimental. Exception:
'JavaPackage' object is not callable
2023/05/03 03:22:08 WARNING mlflow.utils.autologging_utils: Encountered unexpected error during spark autologging: Exception while attempting to initialize JVM-side state for S
park datasource autologging. Please create a new Spark session and ensure you have the mlflow-spark JAR attached to your Spark session as described in http://mlflow.org/docs/la
test/tracking.html#automatic-logging-from-spark-experimental. Exception:
'JavaPackage' object is not callable
2023/05/03 03:22:08 WARNING mlflow.utils.autologging_utils: Encountered unexpected error during spark autologging: Exception while attempting to initialize JVM-side state for S
park datasource autologging. Please create a new Spark session and ensure you have the mlflow-spark JAR attached to your Spark session as described in http://mlflow.org/docs/la
test/tracking.html#automatic-logging-from-spark-experimental. Exception:
'JavaPackage' object is not callable
2023/05/03 03:22:08 WARNING mlflow.utils.autologging_utils: Encountered unexpected error during spark autologging: Exception while attempting to initialize JVM-side state for S
park datasource autologging. Please create a new Spark session and ensure you have the mlflow-spark JAR attached to your Spark session as described in http://mlflow.org/docs/la
test/tracking.html#automatic-logging-from-spark-experimental. Exception:
'JavaPackage' object is not callable
2023/05/03 03:22:08 WARNING mlflow.utils.autologging_utils: Encountered unexpected error during spark autologging: Exception while attempting to initialize JVM-side state for S
park datasource autologging. Please create a new Spark session and ensure you have the mlflow-spark JAR attached to your Spark session as described in http://mlflow.org/docs/la
test/tracking.html#automatic-logging-from-spark-experimental. Exception:
'JavaPackage' object is not callable
2023/05/03 03:22:08 WARNING mlflow.utils.autologging_utils: Encountered unexpected error during spark autologging: Exception while attempting to initialize JVM-side state for S
park datasource autologging. Please create a new Spark session and ensure you have the mlflow-spark JAR attached to your Spark session as described in http://mlflow.org/docs/la
test/tracking.html#automatic-logging-from-spark-experimental. Exception:
'JavaPackage' object is not callable
2023/05/03 03:22:08 WARNING mlflow.utils.autologging_utils: Encountered unexpected error during spark autologging: Exception while attempting to initialize JVM-side state for S
park datasource autologging. Please create a new Spark session and ensure you have the mlflow-spark JAR attached to your Spark session as described in http://mlflow.org/docs/la
test/tracking.html#automatic-logging-from-spark-experimental. Exception:
'JavaPackage' object is not callable
2023/05/03 03:22:08 WARNING mlflow.utils.autologging_utils: Encountered unexpected error during spark autologging: Exception while attempting to initialize JVM-side state for S
park datasource autologging. Please create a new Spark session and ensure you have the mlflow-spark JAR attached to your Spark session as described in http://mlflow.org/docs/la
test/tracking.html#automatic-logging-from-spark-experimental. Exception:
'JavaPackage' object is not callable
2023/05/03 03:22:08 WARNING mlflow.utils.autologging_utils: Encountered unexpected error during spark autologging: Exception while attempting to initialize JVM-side state for S
```

```
park datasource autologging. Please create a new Spark session and ensure you have the mlflow-spark JAR attached to your Spark session as described in http://mlflow.org/docs/la
test/tracking.html#automatic-logging-from-spark-experimental. Exception:
'JavaPackage' object is not callable
2023/05/03 03:22:09 WARNING mlflow.utils.autologging_utils: Encountered unexpected error during spark autologging: Exception while attempting to initialize JVM-side state for S
park datasource autologging. Please create a new Spark session and ensure you have the mlflow-spark JAR attached to your Spark session as described in http://mlflow.org/docs/la
test/tracking.html#automatic-logging-from-spark-experimental. Exception:
'JavaPackage' object is not callable
2023/05/03 03:22:09 WARNING mlflow.utils.autologging_utils: Encountered unexpected error during spark autologging: Exception while attempting to initialize JVM-side state for S
park datasource autologging. Please create a new Spark session and ensure you have the mlflow-spark JAR attached to your Spark session as described in http://mlflow.org/docs/la
test/tracking.html#automatic-logging-from-spark-experimental. Exception:
'JavaPackage' object is not callable
2023/05/03 03:22:09 WARNING mlflow.utils.autologging_utils: Encountered unexpected error during spark autologging: Exception while attempting to initialize JVM-side state for S
park datasource autologging. Please create a new Spark session and ensure you have the mlflow-spark JAR attached to your Spark session as described in http://mlflow.org/docs/la
test/tracking.html#automatic-logging-from-spark-experimental. Exception:
'JavaPackage' object is not callable
  paramGrid: [{Param(parent='LinearRegression_6544c87f2872', name='regParam', doc='regularization parameter (>= 0).'): 0.01, Param(parent='LinearRegression_6544c87f2872', name
='fitIntercept', doc='whether to fit an intercept term.'): False, Param(parent='LinearRegression_6544c87f2872', name='elasticNetParam', doc='the ElasticNet mixing parameter, in
range [0, 1]. For alpha = 0, the penalty is an L2 penalty. For alpha = 1, it is an L1 penalty.'): 1.0}, {Param(parent='LinearRegression_6544c87f2872', name='regParam', doc='reg
ularization parameter (>= 0).'): 0.01, Param(parent='LinearRegression_6544c87f2872', name='fitIntercept', doc='whether to fit an intercept term.'): True, Param(parent='LinearRe
gression_6544c87f2872', name='elasticNetParam', doc='the ElasticNet mixing parameter, in range [0, 1]. For alpha = 0, the penalty is an L2 penalty. For alpha = 1, it is an L1 p
enalty.'): 1.0}, {Param(parent='LinearRegression_6544c87f2872', name='regParam', doc='regularization parameter (>= 0).'): 0.1, Param(parent='LinearRegression_6544c87f2872', nam
e='fitIntercept', doc='whether to fit an intercept term.'): False, Param(parent='LinearRegression_6544c87f2872', name='elasticNetParam', doc='the ElasticNet mixing parameter, i
n range [0, 1]. For alpha = 0, the penalty is an L2 penalty. For alpha = 1, it is an L1 penalty.'): 1.0}, {Param(parent='LinearRegression_6544c87f2872', name='regParam', doc='r
egularization parameter (>= 0).'): 0.1, Param(parent='LinearRegression_6544c87f2872', name='fitIntercept', doc='whether to fit an intercept term.'): True, Param(parent='LinearR
egression_6544c87f2872', name='elasticNetParam', doc='the ElasticNet mixing parameter, in range [0, 1]. For alpha = 0, the penalty is an L2 penalty. For alpha = 1, it is an L1
penalty.'): 1.0}, {Param(parent='LinearRegression_6544c87f2872', name='regParam', doc='regularization parameter (>= 0).'): 1.0, Param(parent='LinearRegression_6544c87f2872', na
me='fitIntercept', doc='whether to fit an intercept term.'): False, Param(parent='LinearRegression_6544c87f2872', name='elasticNetParam', doc='the ElasticNet mixing parameter,
in range [0, 1]. For alpha = 0, the penalty is an L2 penalty. For alpha = 1, it is an L1 penalty.'): 1.0}, {Param(parent='LinearRegression_6544c87f2872', name='regParam', doc
='regularization parameter (>= 0).'): 1.0, Param(parent='LinearRegression_6544c87f2872', name='fitIntercept', doc='whether to fit an intercept term.'): True, Param(parent='Line
arRegression_6544c87f2872', name='elasticNetParam', doc='the ElasticNet mixing parameter, in range [0, 1]. For alpha = 0, the penalty is an L2 penalty. For alpha = 1, it is an
L1 penalty.'): 1.0}, {Param(parent='LinearRegression_6544c87f2872', name='regParam', doc='regularization parameter (>= 0).'): 10.0, Param(parent='LinearRegression_6544c87f287
2', name='fitIntercept', doc='whether to fit an intercept term.'): False, Param(parent='LinearRegression_6544c87f2872', name='elasticNetParam', doc='the ElasticNet mixing param
eter, in range [0, 1]. For alpha = 0, the penalty is an L2 penalty. For alpha = 1, it is an L1 penalty.'): 1.0}, {Param(parent='LinearRegression_6544c87f2872', name='regParam',
doc='regularization parameter (>= 0).'): 10.0, Param(parent='LinearRegression_6544c87f2872', name='fitIntercept', doc='whether to fit an intercept term.'): True, Param(parent
='LinearRegression_6544c87f2872', name='elasticNetParam', doc='the ElasticNet mixing parameter, in range [0, 1]. For alpha = 0, the penalty is an L2 penalty. For alpha = 1, it
is an L1 penalty.'): 1.0}]
2023/05/03 03:22:09 WARNING mlflow.utils.autologging_utils: Encountered unexpected error during spark autologging: Exception while attempting to initialize JVM-side state for S
park datasource autologging. Please create a new Spark session and ensure you have the mlflow-spark JAR attached to your Spark session as described in http://mlflow.org/docs/la
test/tracking.html#automatic-logging-from-spark-experimental. Exception:
'JavaPackage' object is not callable
2023/05/03 03:22:09 WARNING mlflow.utils.autologging_utils: Encountered unexpected error during spark autologging: Exception while attempting to initialize JVM-side state for S
park datasource autologging. Please create a new Spark session and ensure you have the mlflow-spark JAR attached to your Spark session as described in http://mlflow.org/docs/la
test/tracking.html#automatic-logging-from-spark-experimental. Exception:
'JavaPackage' object is not callable
```

Grid Search and Tuning hyperparameters (MLflow Autolog)

```
lr = LinearRegression(featuresCol="features",labelCol='positionOrder',predictionCol='predpositionOrder',maxIter=10, regParam=1, elasticNetParam=1,standardization=False)
```

```
paramGrid = ParamGridBuilder()\
       .addGrid(lr.regParam, [0.01, 0.1, 1, 10]) \
       .addGrid(lr.fitIntercept, [False, True])\
       .addGrid(lr.elasticNetParam, [0.0,0.1,0.2,0.3,0.4,0.5,0.6,0.7,0.8,0.9,1.0])\
       .build()
```

```
evaluator=RegressionEvaluator(labelCol='positionOrder',predictionCol='predpositionOrder')
```

```
# In this case the estimator is simply the linear regression.
# A TrainValidationSplit requires an Estimator, a set of Estimator ParamMaps, and an Evaluator.
tvs = TrainValidationSplit(estimator=lr,
                                              estimatorParamMaps=paramGrid,
                                              evaluator=evaluator,
                                              # 90% of the data will be used for training, 10% for validation.
                                              trainRatio=0.9)
```

```
model1 = tvs.fit(train)
```

```
2023/05/03 03:02:22 WARNING mlflow.utils.autologging_utils: Encountered unexpected error during spark autologging: Exception while attempting to initialize JVM-side state for S
park datasource autologging. Please create a new Spark session and ensure you have the mlflow-spark JAR attached to your Spark session as described in http://mlflow.org/docs/la
test/tracking.html#automatic-logging-from-spark-experimental. Exception:
'JavaPackage' object is not callable
2023/05/03 03:02:22 WARNING mlflow.utils.autologging_utils: Encountered unexpected error during spark autologging: Exception while attempting to initialize JVM-side state for S
park datasource autologging. Please create a new Spark session and ensure you have the mlflow-spark JAR attached to your Spark session as described in http://mlflow.org/docs/la
test/tracking.html#automatic-logging-from-spark-experimental. Exception:
'JavaPackage' object is not callable
2023/05/03 03:02:22 WARNING mlflow.utils.autologging_utils: Encountered unexpected error during spark autologging: Exception while attempting to initialize JVM-side state for S
park datasource autologging. Please create a new Spark session and ensure you have the mlflow-spark JAR attached to your Spark session as described in http://mlflow.org/docs/la
test/tracking.html#automatic-logging-from-spark-experimental. Exception:
'JavaPackage' object is not callable
MLlib will automatically track trials in MLflow. After your tuning fit() call has completed, view the MLflow UI to see logged runs.
2023/05/03 03:02:59 WARNING mlflow.utils.autologging_utils: Encountered unexpected error during spark autologging: Exception while attempting to initialize JVM-side state for S
park datasource autologging. Please create a new Spark session and ensure you have the mlflow-spark JAR attached to your Spark session as described in http://mlflow.org/docs/la
test/tracking.html#automatic-logging-from-spark-experimental. Exception:
'JavaPackage' object is not callable
2023/05/03 03:02:59 WARNING mlflow.utils.autologging_utils: Encountered unexpected error during spark autologging: Exception while attempting to initialize JVM-side state for S
park datasource autologging. Please create a new Spark session and ensure you have the mlflow-spark JAR attached to your Spark session as described in http://mlflow.org/docs/la
test/tracking.html#automatic-logging-from-spark-experimental. Exception:
'JavaPackage' object is not callable
2023/05/03 03:02:59 WARNING mlflow.utils.autologging_utils: Encountered unexpected error during spark autologging: Exception while attempting to initialize JVM-side state for S
park datasource autologging. Please create a new Spark session and ensure you have the mlflow-spark JAR attached to your Spark session as described in http://mlflow.org/docs/la
test/tracking.html#automatic-logging-from-spark-experimental. Exception:
'JavaPackage' object is not callable
2023/05/03 03:02:59 WARNING mlflow.utils.autologging_utils: Encountered unexpected error during spark autologging: Exception while attempting to initialize JVM-side state for S
park datasource autologging. Please create a new Spark session and ensure you have the mlflow-spark JAR attached to your Spark session as described in http://mlflow.org/docs/la
test/tracking.html#automatic-logging-from-spark-experimental. Exception:
'JavaPackage' object is not callable
2023/05/03 03:02:59 WARNING mlflow.utils.autologging_utils: Encountered unexpected error during spark autologging: Exception while attempting to initialize JVM-side state for S
park datasource autologging. Please create a new Spark session and ensure you have the mlflow-spark JAR attached to your Spark session as described in http://mlflow.org/docs/la
test/tracking.html#automatic-logging-from-spark-experimental. Exception:
'JavaPackage' object is not callable
2023/05/03 03:02:59 WARNING mlflow.utils.autologging_utils: Encountered unexpected error during spark autologging: Exception while attempting to initialize JVM-side state for S
park datasource autologging. Please create a new Spark session and ensure you have the mlflow-spark JAR attached to your Spark session as described in http://mlflow.org/docs/la
test/tracking.html#automatic-logging-from-spark-experimental. Exception:
'JavaPackage' object is not callable
2023/05/03 03:02:59 WARNING mlflow.utils.autologging_utils: Encountered unexpected error during spark autologging: Exception while attempting to initialize JVM-side state for S
park datasource autologging. Please create a new Spark session and ensure you have the mlflow-spark JAR attached to your Spark session as described in http://mlflow.org/docs/la
test/tracking.html#automatic-logging-from-spark-experimental. Exception:
'JavaPackage' object is not callable
2023/05/03 03:02:59 WARNING mlflow.utils.autologging_utils: Encountered unexpected error during spark autologging: Exception while attempting to initialize JVM-side state for S
park datasource autologging. Please create a new Spark session and ensure you have the mlflow-spark JAR attached to your Spark session as described in http://mlflow.org/docs/la
test/tracking.html#automatic-logging-from-spark-experimental. Exception:
'JavaPackage' object is not callable
2023/05/03 03:02:59 WARNING mlflow.utils.autologging_utils: Encountered unexpected error during spark autologging: Exception while attempting to initialize JVM-side state for S
park datasource autologging. Please create a new Spark session and ensure you have the mlflow-spark JAR attached to your Spark session as described in http://mlflow.org/docs/la
test/tracking.html#automatic-logging-from-spark-experimental. Exception:
```

```
# Make predictions on test data. model is the model with combination of parameters
# that performed best.
model1.transform(test)\
       .select("features", "positionOrder", "predpositionOrder").show()
```

```
2023/05/03 03:03:01 WARNING mlflow.utils.autologging_utils: Encountered unexpected error during spark autologging: Exception while attempting to initialize JVM-side state for S
park datasource autologging. Please create a new Spark session and ensure you have the mlflow-spark JAR attached to your Spark session as described in http://mlflow.org/docs/la
test/tracking.html#automatic-logging-from-spark-experimental. Exception:
'JavaPackage' object is not callable
2023/05/03 03:03:01 WARNING mlflow.utils.autologging_utils: Encountered unexpected error during spark autologging: Exception while attempting to initialize JVM-side state for S
park datasource autologging. Please create a new Spark session and ensure you have the mlflow-spark JAR attached to your Spark session as described in http://mlflow.org/docs/la
test/tracking.html#automatic-logging-from-spark-experimental. Exception:
'JavaPackage' object is not callable
2023/05/03 03:03:01 WARNING mlflow.utils.autologging_utils: Encountered unexpected error during spark autologging: Exception while attempting to initialize JVM-side state for S
park datasource autologging. Please create a new Spark session and ensure you have the mlflow-spark JAR attached to your Spark session as described in http://mlflow.org/docs/la
test/tracking.html#automatic-logging-from-spark-experimental. Exception:
'JavaPackage' object is not callable
+--------------------+-------------+------------------+
|            features|positionOrder| predpositionOrder|
+--------------------+-------------+------------------+
|[31.0,17.0,201.0,...|           10|13.829534212474647|
|[31.0,17.0,202.0,...|           14|14.834924938456977|
|[30.0,16.0,201.0,...|           16|16.158681951605523|
|[54.0,42.0,169.0,...|           16|13.021355773487052|
|[50.0,41.0,170.0,...|           17|14.478477466644637|
|[66.0,28.0,202.0,...|            2| 8.141295694810836|
|[78.0,48.0,158.0,...|            6|  8.44017290099757|
|[76.0,73.0,158.0,...|           15| 9.326167621690395|
|[57.0,53.0,215.0,...|           18|10.357346258441076|
|[58.0,56.0,218.0,...|            8| 9.047713972424358|
|[58.0,51.0,219.0,...|            3| 8.963661743395509|
|[70.0,63.0,191.0,...|            6| 8.419517868989146|
|[57.0,55.0,197.0,...|           14|10.190407267215761|
|[53.0,30.0,243.0,...|            7| 8.310623510578488|
|[51.0,50.0,242.0,...|           16|10.116119379166825|
|[53.0,48.0,245.0,...|            1| 8.516558498867706|
|[52.0,50.0,243.0,...|           13| 9.990548408904452|
|[11.0,5.0,219.0,8...|           19| 17.03660175217184|
|[71.0,20.0,209.0,...|            8| 6.849110410436918|
|[55.0,14.0,198.0,...|            2| 9.938202415482085|
+--------------------+-------------+------------------+
2023/05/03 03:03:00 WARNING mlflow.utils.autologging_utils: Encountered unexpected error during spark autologging: Exception while attempting to initialize JVM-side state for S
```

Other Codes

```
#from pyspark.ml.evaluation import RegressionEvaluator
#from pyspark.ml.regression import LinearRegression
#from pyspark.sql.functions import datediff, current_date, avg
#from mlflow import pyfunc
#import mlflow.tensorflow
```

```
#df_laptimes = spark.read.csv('s3://columbia-gr5069-main/raw/lap_times.csv',header=True)
#display(df_laptimes)
#df_drivers = spark.read.csv('s3://columbia-gr5069-main/raw/drivers.csv',header=True)
#display(df_drivers)
#df_pitstops = spark.read.csv('s3://columbia-gr5069-main/raw/pit_stops.csv',header=True)
#display(df_pitstops)
#df_lap_drivers = df_drivers.select('DriverId','driverRef','code','forename','surname','nationality','age').join(df_laptimes,on=['DriverId'])
#assembled_df.show(10, truncate=False)

# do join in pyspark, and do select in pyspark. The order is processed starting on left hand side, similar as in SQL.
```

```
#df3=df2.withColumn('fastestLapTime',to_timestamp(col('fastestLapTime'))).withColumn('DiffInSeconds',col('fastestLapTime').cast("long"))
#display(df3)
#df3=df2.withColumn('fastestLapTime',from_unixtime(unix_timestamp("fastestLapTime",  "m:ss.SSS"),"mm:ss.SSSSSS"))
#spark.conf.set("spark.sql.session.timeZone",  "America/Los_Angeles")
#df3=df2.select(unix_timestamp('fastestLapTime',  'yyyy-MM-dd').alias('unix_time')).collect()
#spark.conf.unset("spark.sql.session.timeZone")
#display(df3)
#df3=df2.withColumn("fastestLapTime",to_timestamp(col("fastestLapTime"),"mm:ss.SSS"))
#df3=df2.withColumn("fastestLapTime",unix_timestamp(col("fastestLapTime"),"mm:ss.SSS"))
#df3=df2.withColumn("fastestLapTime",unix_timestamp(col("fastestLapTime"),"mm:ss.SSS"))
#df3=df2.withColumn("fastestLapTime",df2['fastestLapTime'].cast(IntegerType()))
#display(df3)
#display(df3)
#df  =  spark.createDataFrame(data=dates,  schema=["id","input_timestamp"])

#from  pyspark.sql.functions  import  *

#Calculate  Time  difference  in  Seconds
#df3=df2.withColumn('fastestLapTime',to_timestamp(col('fastestLapTime')))

#\
  #  .withColumn('end_timestamp',  current_timestamp())\
#    .withColumn('DiffInSeconds',col("end_timestamp").cast("long")  -  col('fastestLapTime').cast("long"))
#df2.show(truncate=False)
```

```
#df2  =  pyspark.sql.DataFrame.dropna(df,'any')
#display(df2)
#df2.withColumn("raceId",col("raceId").cast("Integer")).withColumn("driverId",col("driverId").cast("Integer"))
#df2=df2.select(col("positionOrder").alias("label"),col("laps").alias("feature1"),col("fastestLap").alias("feature2"),col("fastestLapSpeed").alias("feature3"),col("statusId").alias("feature4"),col("driverId").alias("featu
```

```
#training  =  df
#lr  =  LinearRegression(maxIter=10)
#  Fit  Lasso  model  (λ  =  1,  α  =  1)  to  training  data
#regression  =  LinearRegression(labelCol='statusId',  regParam=1,  elasticNetParam=1).fit(train)

#  Calculate  the  RMSE  on  testing  data
##rmse  =  RegressionEvaluator(labelCol='duration').evaluate(regression.transform(flights_test))
#print("The  test  RMSE  is",  rmse)

#  Look  at  the  model  coefficients
#coeffs  =  regression.coefficients
#print(coeffs)

#  Number  of  zero  coefficients
#zero_coeff  =  sum([beta  ==  0  for  beta  in  regression.coefficients])
#print("Number  of  coefficients  equal  to  0:",  zero_coeff)

#lr  =  LinearRegression(maxIter=10,  regParam=0.3,  elasticNetParam=1)

#  Fit  the  model
#lrModel  =  lr.fit(training)

#  Print  the  weights  and  intercept  for  linear  regression
#print("Weights:  "  +  str(lrModel.weights))
#print("Intercept:  "  +  str(lrModel.intercept))
```

```
#regression  =  LinearRegression(labelCol='positionOrder',fitIntercept=True)
#regression=regression.fit(train)
#params  =  ParamGridBuilder()
#params  =  params.addGrid(regression.fitIntercept,  [True,False]).addGrid(regression.regParam,[0.001,0.01,0.1,1,10]).addGrid(regression.elasticNetParam,[0,0.25,0.5,0.75,1])
#params  =  params.build()
#print('Number  of  models  to  be  tested',len(params))
#evaluator=  RegressionEvaluator(labelCol='positionOrder',predictionCol='predpositionOrder')
#cv=  CrossValidator(estimator=  model,estimatorParamMaps=params,evaluator=evaluator)
#cv=cv.setNumFolds(10).setSeed(12345).fit(train)
#cv=cv.fit(train)
#cv.avgMetrics
#cv.bestModel
#predictions  =  cv.transform(test)
#cv.bestModel.explainParam('fitIntercept')
```

```
#model=lr.fit(train)
#  We  use  a  ParamGridBuilder  to  construct  a  grid  of  parameters  to  search  over.
#  TrainValidationSplit  will  try  all  combinations  of  values  and  determine  best  model  using
#  the  evaluator.

#    mlflow.spark.log_model(model,  "myModel")
#lr2  =  LinearRegression(maxIter=10,featuresCol="features",labelCol='positionOrder',predictionCol='predpositionOrder')

#  We  use  a  ParamGridBuilder  to  construct  a  grid  of  parameters  to  search  over.
#  TrainValidationSplit  will  try  all  combinations  of  values  and  determine  best  model  using
#  the  evaluator.
#paramGrid  =  ParamGridBuilder()\
#        .addGrid(lr2.regParam,  [0.1,  0.01])  \
#        .addGrid(lr2.fitIntercept,  [False,  True])\
#        .addGrid(lr2.elasticNetParam,  [0.0,  0.5,  1.0])\
#        .build()

    #  In  this  case  the  estimator  is  simply  the  linear  regression.
#  A  TrainValidationSplit  requires  an  Estimator,  a  set  of  Estimator  ParamMaps,  and  an  Evaluator.
#tvs  =  TrainValidationSplit(estimator=lr,
                                                #estimatorParamMaps=paramGrid,
                                                #evaluator=RegressionEvaluator(labelCol='positionOrder',#predictionCol='predpositionOrder'),
                                              #  80%  of  the  data  will  be  used  for  training,  20%  for  validation.
#                                                 trainRatio=0.9)
#  Run  TrainValidationSplit,  and  choose  the  best  set  of  parameters.
#model2  =  tvs.fit(train)

#  Make  predictions  on  test  data.  model  is  the  model  with  combination  of  parameters
#  that  performed  best.
#model.transform(test)\
  #      .select("features",  "label",  "prediction")\
#    .show()
```

```
#mlflow.spark.log_model(model1,  "myModel1")
```

```
#paramGrid = ParamGridBuilder()\
        #.addGrid(lr.regParam, [0.01, 0.1, 1, 10]) \
        #.addGrid(lr.fitIntercept, [False, True])\
        #.addGrid(lr.elasticNetParam, [0.0,0.1,0.2,0.3,0.4,0.5,0.6,0.7,0.8,0.9,1.0])\
        #.build()
#print('Number of models to be tested',len(paramGrid))
#SparkSession.builder.config("spark.jars.packages", "org.mlflow.mlflow-spark")
#mlflow.spark.autolog()
#model1 = tvs.fit(train)
#paramGrid = ParamGridBuilder()\
        #.addGrid(lr.regParam, [0.01, 0.1, 1, 10]) \
      #  .addGrid(lr.elasticNetParam, [0.0,0.3,0.6,0.9,1.0])\
        #.build()
#print('Number of models to be tested',len(paramGrid))
```

```
#with mlflow.start_run(run_name="Elastic Net Regression Model") as run:
    #vecAssembler = VectorAssembler(inputCols = ["bedrooms", "bathrooms"], outputCol = "features")

    #vecTrainDF = vecAssembler.transform(trainDF)

    #lr = LinearRegression(featuresCol = "features", labelCol = "price") # a spark model

    #lrModel = lr.fit(vecTrainDF)
```

```
#with mlflow.start_run():
        #model1 = tvs.fit(train)
        #mlflow.log_param("rmse",tvs.)
        #mlflow.sklearn.log_model(tvs,"model")

        #test_metric = evaluator.evaluate(model1.transform(test))
        #mlflow.log_metric('test_' + evaluator.getMetricName(), test_metric)
```

```
    #print("  r2: {}".format(r2))
#  Explicitly create a new run.
#  This allows this cell to be run multiple times.
#  If you omit mlflow.start_run(), then this cell could run once, but a second run would hit conflicts when attempting to overwrite the first run.

#with mlflow.start_run():
    #  Run the cross validation on the training dataset. The tvs.fit() call returns the best model it found.
    #model1 = tvs.fit(train)

    #  Evaluate the best model's performance on the test dataset and log the result.
    #test_metric = evaluator.evaluate(model1.transform(test))
    #mlflow.log_metric('test_' + evaluator.getMetricName(), test_metric)

    #  Log the best model.
    #mlflow.spark.log_model(spark_model=model1.bestModel, artifact_path='best-model')
#MLlib will automatically track trials in MLflow. After your tuning fit() call has completed, view the MLflow UI to see logged runs.
```

```
#vecAssembler = VectorAssembler(inputCols = ["bedrooms"], outputCol = "features")

#vecTrainDF = vecAssembler.transform(trainDF)

#lr = LinearRegression(featuresCol = "features", labelCol = "price")
#lrModel = lr.fit(vecTrainDF)
#vecAssembler = VectorAssembler(inputCols = ["bedrooms"], outputCol = "features")

#vecTrainDF = vecAssembler.transform(trainDF)
    #vecAssembler = VectorAssembler(inputCols = ["bedrooms", "bathrooms"], outputCol = "features")
    #vecTrainDF = vecAssembler.transform(train)
    #vecTrainDF = vecAssembler.transform(trainDF)
  #evaluator = RegressionEvaluator(predictionCol="prediction", labelCol="price")
    #lr = LinearRegression(featuresCol = "features", labelCol = "price") # a spark model
    #lrModel = lr.fit(vecTrainDF)
        #r2 = evaluator.evaluate(predtest, {evaluator.metricName: "r2"})
```