**//Complete program of ALS prediction with images in dataframe Python - pyspark**

from pyspark.ml.evaluation import RegressionEvaluator

from pyspark.ml.recommendation import ALS

from pyspark.mllib.recommendation import Rating

from pyspark.sql import Row

#Craeted RDD dataset of Ratings data

ratingdata = spark.read.text("/user/Test/ratings.csv").rdd

#Splitting the data as "," delimited CSV

ratingspldata = ratingdata.map(lambda row: row.value.split(","))

#Remove header or filtered

header = ratingspldata.first()

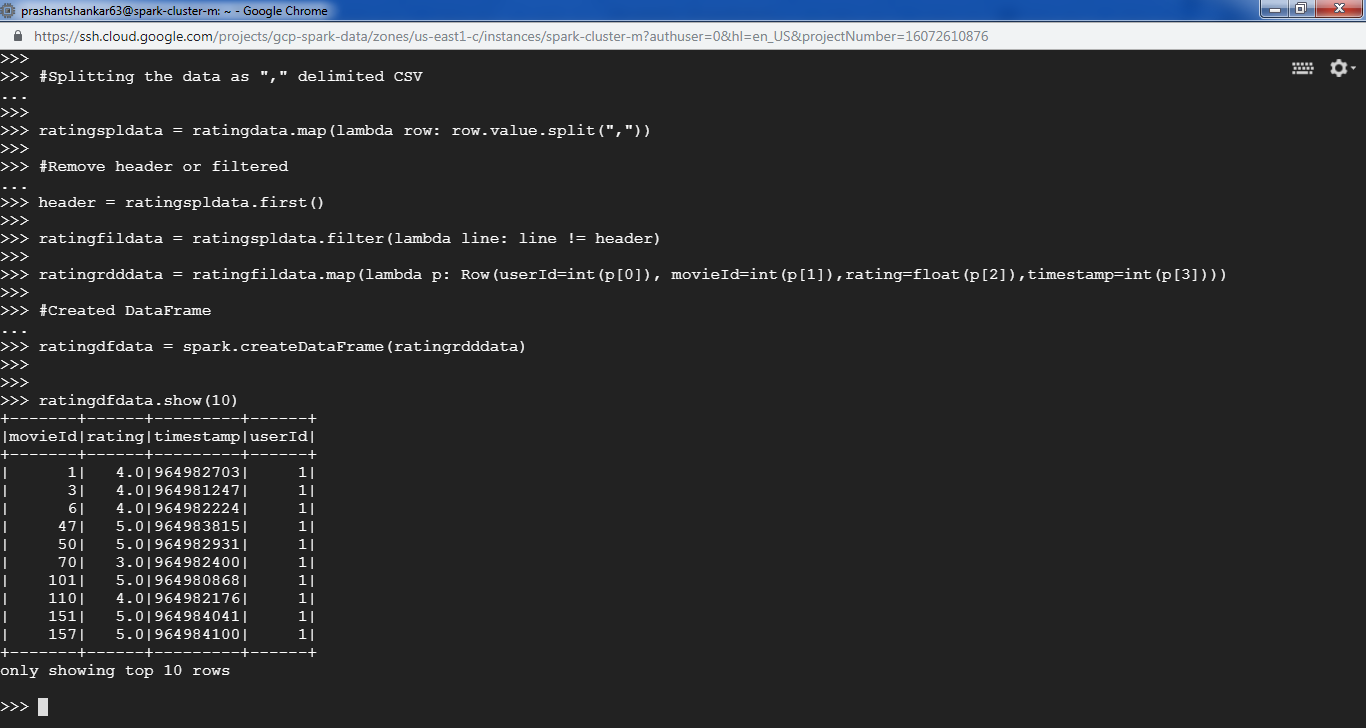
ratingfildata = ratingspldata.filter(lambda line: line != header)

ratingrdddata = ratingfildata.map(lambda p: Row(userId=int(p[0]), movieId=int(p[1]),rating=float(p[2]),timestamp=int(p[3])))

#Created DataFrame

ratingdfdata = spark.createDataFrame(ratingrdddata)

ratingdfdata.show(5)



# splitting dataframe into two parts

(training, test) = ratingdfdata.randomSplit([0.8, 0.2])

# Build the recommendation model using ALS on the training data

als = ALS(maxIter=5, regParam=0.01, userCol="userId", itemCol="movieId", ratingCol="rating")

model = als.fit(training)

# Evaluate the model by computing the RMSE on the test data

predictions = model.transform(test)

predictions.show(10)

