Task-5-printSchema

July 28, 2020

```
[4]: import pandas as pd
[30]: shoesPredictions = spark.read.parquet("hdfs:///data/exercise/shoes_predictions.
      →parquet")
[31]: shoesPredictions.printSchema()
    root
     |-- created_at: string (nullable = true)
     |-- name: string (nullable = true)
     |-- prediction: double (nullable = true)
     |-- prediction_text: string (nullable = true)
[37]: pd.DataFrame(shoesPredictions.select('created_at', 'name', 'prediction', __
      [37]:
                           created_at \
         Mon Jul 27 16:07:18 +0000 2020
        Mon Jul 27 16:01:28 +0000 2020
     1
         Mon Jul 27 16:13:00 +0000 2020
     2
         Mon Jul 27 16:11:12 +0000 2020
     3
        Mon Jul 27 16:03:41 +0000 2020
     5
         Mon Jul 27 16:09:02 +0000 2020
        Mon Jul 27 16:03:19 +0000 2020
         Mon Jul 27 16:03:18 +0000 2020
        Mon Jul 27 16:05:53 +0000 2020
        Mon Jul 27 16:07:15 +0000 2020
     10 Mon Jul 27 16:12:35 +0000 2020
     11 Mon Jul 27 16:10:20 +0000 2020
     12 Mon Jul 27 16:10:28 +0000 2020
        Mon Jul 27 16:08:02 +0000 2020
     14 Mon Jul 27 16:08:13 +0000 2020
                                         original_tweet category \
         RT @DoctorKhanTweet: "
     0
                                   "\n\nI am no...
                                                      4.0
         "We're so much more capable than what people t...
```

```
2.0
2
    RT @PeteJennie: @BrandHubNg Hello! We have a ...
3
    RT @SapienCassidy: I recently opened Cassidy's...
                                                            2.0
4
    Channahmmalcolm Same shop. Same size. Vary the...
                                                            3.0
    Obviously I made impulsive purchases. Everyone...
                                                            2.0
5
6
    These shoes give a playful twist to the design...
                                                            2.0
     Royal Geo 2 Mini Women's Mini Skirt
7
                                             $27.49...
                                                           0.0
8
   RT @Nike: Shoes, hoodies, shorts, and more to ...
                                                            2.0
9
    RT @LordC4rl: @crimlawuk In my town shopping m...
                                                            4.0
10 RT @ArleasMommy: @HBCUfessions On top of all t...
                                                            4.0
   RT @FindomShawn: Back from a sweet Gym session...
                                                            2.0
12
   RT @Nike: Save big on next season's shoes, hoo...
                                                            1.0
    Upto 70% off on top brands like Puma, Liberty ...
                                                            0.0
    @TheDukeNGS And I completely respect and suppo...
                                                            1.0
```

category_text 0 shoes-athletic 1 mens-lace-up-shoes 2 mens-lace-up-shoes 3 mens-lace-up-shoes mens-shoes-athletic mens-lace-up-shoes 5 6 mens-lace-up-shoes 7 girls-shoes 8 mens-lace-up-shoes 9 shoes-athletic 10 shoes-athletic mens-lace-up-shoes 11 12 boys-shoes 13 girls-shoes 14 boys-shoes

[]: