Course Project Quiz

LATEST SUBMISSION GRADE

100%

1. Please don't forget to import the assignment notebook before answering questions as instructed the learning item before. **1 / 1 point**

   https://github.com/IBM/coursera/blob/master/coursera_bd/week4/a6_w4_assignment.ipynb

   Please use the discussion forums if you have questions

   What is the correlation between HOURLYWindSpeed and HOURLYPressureTendency? Please use the already existing code for the colleration matrix and adjust the code respectivly.

   -0.01324305

   ✓ **Correct**

2. What is the RMSE metric obtained from the LinearRegression model (1st model in the notebook - cell has comment #LR1) **1 / 1 point**

   5.30775

   ✓ **Correct**

3. Please change #LR1 in order to use features_norm over features. What's the RMSE value you get now? **1 / 1 point**

   5.53835

   ✓ **Correct**

4. What's the RMSE value we obtain from cell #GBT1? **1 / 1 point**

   5.11807

   ✓ **Correct**

5. What is the accuracy you get from cell #LGReg1? **1 / 1 point**

   0.692922

   ✓ **Correct**

6. What is the accuracy you get from cell #RF1? **1 / 1 point**

   0.717352

   ✓ **Correct**

7. What is the accuracy you get from cell #GBT2? **1 / 1 point**

0.73105

✓ **Correct**

8. If you change the number of trees in cell #RF1 from 30 to 10, what's the new accuracy? **1 / 1 point**

0.71895

✓ **Correct**

9. What data storage format is the used? **1 / 1 point**

- ○ CSV, without header, columns separated by comma
- ◉ CSV, with header, columns separated by comma
- ○ PARQUET
- ○ JSON
- ○ CSV, without header, columns separated by semicolon
- ○ CSV, with header, columns separated by semicolon

✓ **Correct**
correct

10.What correlation methods are supported by the Correlation matrix function? 1 / 1 point

https://spark.apache.org/docs/latest/api/python/pyspark.ml.html#pyspark.ml.stat.Correlation

- [ ] MAE
- [ ] RMSE
- [x] Pearson

✓ **Correct**
correct

- [ ] R2
- [x] Spearman

✓ **Correct**
correct

11.Which Classification Model performs best in this notebook? 1 / 1 point

1. Logistic Regression
2. Random Forest
3. Gradient Boosted Tree

Possible answers: 1,2 or 3

3

✓ **Correct**

12.Which Regression Model performs best in this notebook? 1 / 1 point

1. Linear Regression
2. Gradient Boosted Tree Regressor

Possible answers: 1 or 2

2

✓ **Correct**