



Assignment 3:

Latent Factor Model for Rating Prediction

Overview

- Programming
 - Latent factor model for rating prediction
- Requirements
 - Project report
 - Executable source code
 - Prediction results

Project Report

- Main content
 - Detailed implementation introduction
 - Key code explanation
 - Result analysis
 - Hyper-parameter influence
 - Table or plot
 - How to run your code
 - Environments (E.g., python 2.X or python 3.x, c++?)
 - Dependency libraries

Source Code

- Programming language
 - Python is preferred.
- Terminal command
 - E.g., `python2.7 xxx.py test.file output.file`
- Source code files
 - How the method is implemented

Data Format

- Train set
 - With rating scores

```
train.csv
1 user_id,business_id,date,stars
2 A2JGzkvNjckSmps_4FbKWw,Xg5qEQiB-7L6kGJ5F4K3bQ,2014-03-18 01:14:10,5.0
3 rypcWiSNGM0suWsiSLh9xA,4RoTEeqB_MNn6yaqZmlZHg,2015-08-29 18:32:15,4.0
4 Dgk0Wdoh7HPjhKQEPBU_jQ,ZOmF-3NN4Z59b2Fw6VAM7g,2015-09-14 16:33:03,3.0
5 FIk4lQQu1eTe2EpzQ4xhBA,HK2Ki-PvnNN-YMTlX1uSVA,2012-09-29 02:03:42,4.0
6 VizhcyMWWPz3UDXEBeix4w,UPIYuRaZvknINod1w8kqRQ,2011-06-10 20:35:42,3.0
7 2EuPAGalYnP7eSxPgFCNDg,E83nSU_y9zed0zQnkTjV1g,2017-08-14 18:56:00,2.0
8 WXlxViTwXHPBvhioljN9PQ,IRzY7yoBqoHaZNN08WiWQQ,2016-04-30 14:42:50,4.0
9 cMeTAiW60I5wE_vLfTxoJQ,DESv2ys6SjBKA4SyDtJvxw,2012-06-26 00:44:03,4.0
```

- Test set
 - Without rating scores

```
test.csv
1 user_id,business_id,date
2 PfpRvMAESbC2bC8FUIMdNg,Kbbm6Vd5UdbP10dwjBghRw,2018/10/15 0:52
3 oaaEXgQ3x51cXE3GTxrT1Q,2GmGT-7QjowR1ihup3FbVA,2011/11/27 9:12
4 yT_QCcnq-QGipWWuzIpvtw,pOEL97ld-FJMK08Ki8JmYg,2016/3/11 19:09
5 fRVNHA12RjosC67Y67G3cA,UklWme3kwg6L9rd4tCNB15w,2016/9/11 15:53
6 48vRThjhuhisQINQ2KV8Sw,LNGBEEelQx4zbfWnlc66cw,2011/3/23 3:22
7 q5FQmuXxzPEsvEtA_Mvd1w,fSBhe0A6Dfa8JCYccfpMog,2013/8/9 22:06
8 W0VE9M7Dikrpol8j1_QqyQ,3oTVApC-eUzpGjrOVxIr5g,2017/9/7 2:27
9 avmRUkWovTsaDqKiNKdivQ,wUKzaS1MHg94RGM6z8u9mw,2012/10/25 19:31
```

Prediction Results

- A file containing prediction results
 - Format: each line corresponds to one prediction tuple
 - User_id, Business_id, Rating_score
 - E.g., A2JGzkvNjckSmps_4FbKWw, Xg5qEQiB-7L6kGJ5F4K3bQ, 4.0

Submission

- File name
 - Zip/Rar file, named with 'ID'
- Submission URL
 - <http://xzc.cn/W7MwKi3wMy>

第三次作业

第三次作业

文件提交区域

* 提交人姓名:

→ 拖拽文件上传
(或点击)

最多上传20个文件, 单个文件最大20M

提交

How to evaluate?

- Code
 - Executable
 - Clear, easy-to-understood
- Prediction results
 - Effectiveness
 - How is the predicted score compared with ground-truth?
 - RMSE
- Report
 - Integrity
 - Readability
 - Highlights

Data and Deadline

- **Data**

- <https://pan.baidu.com/s/1nc63rbYsU58PyvoZJtLgBQ>
- 提取码: pgqe

- **Deadline**

- 24:00, 2020-5-30
- Submission after that time might be penalized