Project 3

1. Task: Rating prediction

2. Data:
- User ratings for some businesses
- Extra user information
- Extra business information

3. Evaluation: RMSE

Project 3

1. Task: Rating prediction

2. Data:
- User ratings
- Extra user information
- Extra business information

3. Evaluation: RMSE

Project 3

User ratings:

	user_id	business_id	stars
0	ec8f38aa91755dcf5837020d022ad384	ecaa90564e18dca1c7b653038f71d6bf	1.0
1	64fe4dd0a489c9b96a3e8d7fbd337888	ef118bb0ae1fc369e1f47d1b34f6acee	5.0
2	a49909b39426ebb3538aa837b5b88840	e8b182a923810d52981aa02d56dde799	5.0
3	a56726d5676d647e42e2aca54f21b075	250040e979eae9ef5912aa5a1d285e4e	5.0
4	3e19d8260e655ba87bea0922bac92266	e02880faf4d42fe1df7bd370fb1c787b	4.0

Project 3

1. Task: Rating prediction

2. Data:
   - User ratings
   - Extra user information
   - Extra business information

3. Evaluation: RMSE

Project 3

Extra user information

{ ⊖
    "average_stars":3.63,
    "compliment_cool":1,
    "compliment_cute":0,
    "compliment_funny":1,
    "compliment_hot":1,
    "compliment_list":0,
    "compliment_more":0,
    "compliment_note":0,
    "compliment_photos":0,
    "compliment_plain":0,
    "compliment_profile":0,
    "compliment_writer":0,
    "cool":16,
    "elite":"",
    "fans":4,
    "funny":22,
    "name":"Jenna",
    "review_count":33,
    "useful":48,
    "user_id":"88422913727e71e88611fdfe3512fa03",
    "yelping_since":"2013-02-21 22:29:06"
}

Project 3

1. Task: Rating prediction

2. Data:
- User ratings
- Extra user information
- Extra business information

3. Evaluation: RMSE

Project 3

Extra business information

{ ⊟
    "address":"4075 S Durango Dr, Ste 105B",
    "attributes":{ ⊞ },
    "business_id":"c7d693d13177b9839d89f277e5280315",
    "categories":"Mobile Phones, Mobile Phone Repair, Shopping,
    "city":"Las Vegas",
    "hours":{ ⊞ },
    "is_open":1,
    "latitude":36.115305,
    "longitude":-115.280737,
    "name":"Computer Doctor BG",
    "postal_code":"89147",
    "review_count":211,
    "stars":5.0,
    "state":"NV"
}

Project 3

1. Task: Rating prediction

2. Data:
- User ratings
- Extra user information
- Extra business information

3. Evaluation: RMSE

Project 3

We provide:

- Rating data ('train.csv', 'valid.csv', 'test.csv'), 'stars' column in 'test.csv' is set to 0.0 by default
- User informations ('user.json')
- Business informations ('business.json')
- Sample code (Please replace STUDENT_ID by your *own student id* in the code)

Marking threshold based on the valid data will be released in tutorial 8.

You need to submit:

(1) The prediction file with the 'stars' column filled.
(2) A zip file that contains all the code for us to duplicate your result.
(3) A brief report.

PS: Code for generating prediction files are included in the sample code

Project 3

Other information:

1. Due time: Apr. 24th 23:59 HK time

2. You are welcome to use any methods to make the prediction.

3. The methods taught in the class/tutorial (including previous ones) + some parameter tuning + some feature engineering are enough for you to get the full marks.

4. Late submission policy is same as project 1.

5. Your final grade is automatically calculated based on the RMSE score of your submission file with the gold mention labels. So please do not change the order.