Assignment #6

MACS 30000, Dr. Evans

Due Monday, Nov. 19 at 11:30am

Ruixi Li

1. Netflix Prize and Bell, Koren, and Volinsky(2010)

a) The submissions to the Netflix Prize open call contest is judged on root mean squared error (RMSE). The criteria of this contest is to award the team which conducted "the greatest improvement in root mean squared error (RMSE) over Netflix's internal algorithm, Cinematch"(Bell et al., 2010). Teams whose submissions with a RMSE 10% lower than the Cinematch level (0.9514) was eligible for the judgment.
b) The most commonly used method for predicting ratings (stars) on movies was "nearest neighbors"(Bell et al., 2010) at the beginning of the Netflix Prize contest.
c) Two model are likely to be correlated with each other. (Bell et al., 2010) When combined, they would have a lower RMSE. That's why hybrids of multiple models (ensemble methods) won.

2. Collaborative problem solving: Project Euler

a) rxli: 1409405_W05caGc1h60IK0VgeEZ0dtoeoJEOeiXP
b) The problem I chose is Multiples of 3 and 5(https://projecteuler.net/problem=1). The answer is 233168.

```
In [1] sum = 0
       for x in range(1000):
           if x % 3 == 0 or x % 5 == 0:
               sum += x
       print(sum)
Out [1] 233168
```

c) Three awards that you would most aspire to achieving is Baby Steps, High Five, Master of Archives. First, Baby Steps is the beginning, which symbolize a great start and future. Second, High Five is in the middle which encourage me to plunge forward. Last, Master of Archives is the final stage, which gives credit to the efforts I had devoted.

3. Human computation projects on Amazon Mechanical Turk

a) Experiment: “Draw rectangles around tabular data (tables)”.

b) The reward is $0.02 for completing the whole survey.

c) There are three qualifications. One is that location is US. Second is that HIT approval rate is greater than 95. Third is that Masters has not been granted.

d) The expected time is 1 hours. Hourly rate is $0.02.

e) It expires on Nov 17th.

f) The most this project would cost the HIT experiment creator if 1 million people participated is 0.02 million dollars.

4. Kaggle open calls

a) I registered the Kaggle and my account name is Juliet Li.

b) The open movement I chose is "Two Sigma: Using News to Predict Stock Movements"(https://www.kaggle.com/c/two-sigma-financial-news#description).
Two Sigma sponsored this competition. Two Sigma is a scientifically driven investment institution, which applies technology and data science to financial forecasts for over 17 years. Their outstanding achievements in fields like big data, AI, and machine learning have advanced the development of the investment industry.
In this competition, teams must submit from Kaggle Kernels. Teams are evaluated on the preciseness on predicting the stock movements. Team in the 1st place has a prize of $25,000. Team in the 2nd place has a prize of $20,000. Team in the 3rd place has a prize of $15,000.

Teams in the places from 4th through 7th have $10,000 each. The competition has a list of honor code issues. Teams are not allowed to cheat, which includes using outside information, abusing the competition infrastructure. As for the timeline, the competition has two stage, that is Submission period and Scoring period. In the first period (September 25, 2018 to January 8, 2019), teams will use Kaggle Kernels to train their models. In the second period (January 8, 2019 to July 15, 2019), models last trained will be evaluated with regularly updated news and market data. For submission instructions, all submissions should occur through the Kernels. Only one valid submission per period. In the scoring period, Kaggle will only include valid time period.

c) Since Two Sigma is a scientifically driven investment institution, through this competition, they can find potential talents in computational finance, construct a well-performed stock predictive model and enhance their reputation. In turn, the competition provides opportunities for people ambitious in computer science. Equipped with the best answer, Two Sigma can put the model into use and make their prediction more precisely.

References

Bell, R. M., Koren, Y., & Volinsky, C. (2010). All together now: A perspective on the NETFLIX PRIZE. Chance, 23(1), 24-24. doi:10.1007/s00144-010-0005-2