1.(a) The submissions would be judged by the "root mean squared error (RMSE)" of the submitted model (Bell et al., 2010, p.24). The winner should make "the greatest improvement in root mean squared error (RMSE) over Netflix's internal algorithm, Cinematch" (Bell et al., 2010, p.24). The RMSE function is as follows:

$$RMSE = \sqrt{\frac{\sum_{i=1}^{n}(X_{obs,i} - X_{model,i})^2}{n}}$$

For cutoffs, a submission has to achieve over 10% improvement in RMSE to be qualified for winning, but I didn't see any cutoff for judging from the given material.

(b) At the start, the most commonly used method is "nearest neighbors", with which "the predicted rating for an item by a user might be a weighted average rating of similar items by the same user" (Bell et al., 2010, p.25).

(c) One model can improve the blend "if it was not highly correlated with the other component" (Bell et al., 2010, p.28).

2. (a) my Project Euler user name and friend key:
    txzheng: 1409422_NchOFhXkhMbDYyKIbqdhKsHSYzzb60o6

(b) I chose problem 2: Find the sum of all the even-valued terms in the Fibonacci sequence which do not exceed four million.

My python code attached below:

```python
def fibonacci(n):
    if n == 1:
        return 1
    if n == 2:
        return 2
    else:
        return fibonacci(n - 1) + fibonacci(n - 2)
sequence = []
n = 1
sum = 0
while fibonacci(n) < 4000000:
    sequence.append(fibonacci(n))
    n += 1
for num in sequence:
    if num %2 == 0:
        sum += num
print(sum)
```

My result is 4613732.

(c) The three awards I would most like to achieve are Big Game Hunter, One in A Hundred and Esteemed Journalist. I like Big Game Hunter as it requires solving twenty-five of the fifty hardest problems. An award like this encourages people to challenge themselves. To achieve One in A Hundred, people have to be the first one hundred to solve a problem. I like this one as it is challenging and at the same time not as unrealistic as Gold Medal or Chart Topper, which require people to be the first one or ten to solve a problem. To be "Esteemed Journalist", people have to receive one-hundred Kudos across all their permanent posts, which is quite hard as only 78 members achieved it. But I like this one as it encourages people to share their ideas to help others.

3.(a) The task I selected is "Find Entities in Tweets".
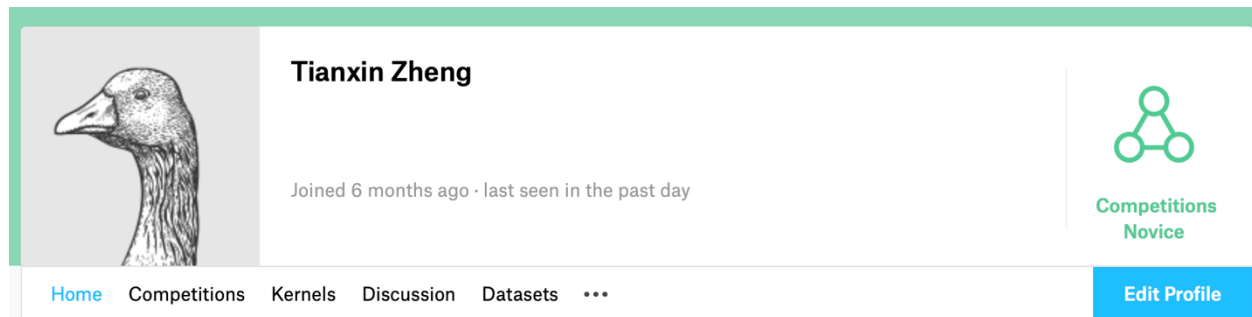(b) The reward is $0.05 per entity.
(c) The qualifications include: message-us-to-proceed has not been granted; External_Entities_Qual Test has been granted; External_Entity_Qualification has been granted; Masters has been granted; z-questions-answered is greater than 200.
(d) Time allotted is 10 minutes. I think I can do 60 in an hour. The implied hourly rate would be $3 per hour.
(e) The job expires on 1/24/2019.
(f) If one million people participated in and they each looked for one entity, it would cost the creator $50000.

4. (a) I've registered.

(b) The competition title is "Two Sigma: Using News to Predict Stock Movements". The sponsor of the competition is Two sigma. Two Sigma is an investment management company driven by technology. Submissions will be evaluated by the mean divided by the standard deviation of the submitted daily xt values. xt is calculated as the sum of the product of ŷti, rti and uti, which are the signed confidence value, the 10-day market-adjusted leading return for day t for instrument I and a 0/1 control variable that decides whether a particular asset is included in scoring on a particular day respectively. The prize structure is as follows: 1st place could win $25,000, 2nd place could win $20,000, 3rd place could win $15,000, 4th through 7th place could win $10,000 each. Important honor code issues include: participants cannot use information outside the dataset provided by the competition to improve the accuracy; participants cannot use the provided information in improper ways. The competition is divided into two stage – a

Submission period and a Scoring period. The Submission period is for participants to train their models and in the Scoring period the models will be evaluated against latest data. Important dates are as follows: the start date is 9/25/2018; submission deadline is 1/8/2019; end date is 7/15/2019. According to submission instructions, all submissions are made through Kaggle kernels environment and there is only one "valid" submission file that could span all the time periods.

(c) I think Two Sigma might use the winning model to make better investment strategy for their clients. They could make better investment by predicting the future stock movement based on existing news data and adjust their portfolio accordingly.