1. a) The Netflix Prize open call context judges its participants by calculating how much improvement, in percentage, their models could make over their original recommendation system *CineMatch* (Bell, Koren & Volinsky 24). Their percentage improvements will be calculated in root mean squared error (RMSE) as the criterion function,

$$RMSE(\widehat{\theta}) = \sqrt{MSE(\widehat{\theta})} = \sqrt{E((\widehat{\theta} - \theta)^2)}$$

where $\widehat{\theta}$ is an estimator of an estimated parameter θ (Wikipedia contributors). The team who managed to improve the recommendation system the most by more than 10% would be awarded $1 million (25). According to the rules posted on the Netflix Prize website and the article by Bell, Koren, and Volinsky (25), once a team successfully registered and submitted their algorithms, their work will be judged. Thus, I do not think there are any cutoffs that would drop out a team's model from the competition. However, the $1 million prize would not be awarded to any team before the 10% improvement rate is reached. Each year, there would be a $50,000 progress prize granted to the team who achieves the highest percentage improvement, but it has to also be greater than the improvement rate for the previous year (Netflix).

b) The nearest neighbors method was the most commonly used at the beginning. Participants predict a viewer's rating for a movie by weighting and averaging his or her existing ratings for other similar movies (25). Therefore, predictions using this model could recommend viewers those movies that are similar to what they recently viewed and highly rated.

c) Bell et al. mentioned that most successful single models in the prize were matrix factorization models (26). When blending models together, averaging two models with similar RMSEs could improve the predictions more. Bell and colleagues' model consists of 107 prediction sets (28), including nearest neighbors and matrix factorization models.

2. a) username: delores9584
   Friend key: 1407410_ifVPgNO8eru2xAi14gHT3zd4vXdFdoT5

b) I selected problem 16, which asks the sum of the digits of the number 2^{1000}. The answer is 1366, and I attached the Python file.

c) I am very interested in the award The Journey Begins, Flawless fifty and High Five. The first one is achieved when one solves 25 problems and reaches level 1. It encourages people to really get involved in Project Euler after they are first introduced to it. Flawless fifty is to solve fifty consecutive questions. I like it because it promotes people to form a habit of solving Project Euler questions. For example, if one attempts to achieve this award by solving a problem every day, after fifty days of work, they could already be closely attached to this project and stick with this daily habit. Similarly, High Five could be earned to solve

the five most recent problems. It also encourages people to stick to the habit of solving problems on Project Euler.

3. a) I selected the HIT ScoutIt.

b) This HIT pays each participant $0.03.

c) The qualifications of this HIT includes: 1) participants' location must be within the United States, 2) a participant's HIT approval rate must not be less than 97%, and 3) the total approved HITs that a participant has completed is not less than 1000.

d) The allotted time is 20 mins. However, since a preview of this HIT tells me that one task simply involves looking at a receipt and select the possible business type of such receipt, I think that I could complete around 70-100 items in an hour. Therefore, my hourly payment will be about $2.1-$3.

e) This job will expire at 11/25/2018.

f) This job would cost 30000 if one million people participate in this task.

4. b) The title of the competition is "Quora Insincere Questions Classification". It is sponsored by Quora, a platform for people to ask questions about various issues and for others to respond and to post their thoughts. Its purpose is to help its users to effectively communicate, to solve each other's questions, and to share beliefs and knowledge. The goal of this current Kaggle competition is to look for models based on machine learning

techniques to identify Quora questions that are insincere, misleading and with low quality. All submissions will be made via a Kaggle Kernel output on F1 Score between their predictions and the observations. Therefore, each submission should include predictions, either sincere (0) or insincere (1), for the list of questions. The specific format of submission is provided online. The file must be named submission.csv and committed to an individual's or a team's Kernal. The competition provides datasets to competitors, and hence their performances will be judged by their places on the leaderboard. Later in the evaluation after the deadline, Quora will also apply a new dataset to the submitted models to finally decide a model's standing in the competition. Thus, the model wins the 1st place will be granted $12,000, the 2nd place will be granted $8,000, and the 3rd will be granted $5,000. The competition start date is Nov. 6, 2018, and the final submission deadline will be Feb. 5, 2019, while the entry deadline will be Jan. 29, 2019. Honor code rules for this competition involves 1) one participant should only have one account registered for this competition, 2) there is no private sharing outside teams and outside of forum posts, 3) a team should have no more than 8 members, 4) one could submit 5 entries per day at maximum, but only 2 final submissions should be selected for the final evaluation.

c) I think that Quora will combine several models together (possibly models that win the 1st, 2nd, and 3rd places) to, hopefully, create a better system that they could apply to their website in order to better filter out insincere questions.

References

Wikipedia contributors, "Root-mean-square deviation," Wikipedia, The Free Encyclopedia, https://en.wikipedia.org/w/index.php?title=Root-mean-square_deviation&oldid=856934073 (accessed November 19, 2018).

Netflix. "The Netflix Prize Rules." Netflix Prize: Home. Accessed November 19, 2018. https://www.netflixprize.com/rules.html.