Group member:

1155084561 DENG Weijun

1155084518 PENG Yibo

1155089134 GUAN Chenye

1155089150 Ou Kaishen

Group Email address: [1155089150@link.cuhk.edu.hk](mailto:1155089150@link.cuhk.edu.hk)

Proposal:

1. Our motivation is to build a movie recommendation system.
2. The most related parts between this course and our project are recommendation system, item similarity, map-reduce algorithm.
3. Input: A user with ratings for some sample movies. Output: Top N (e.g. 5) movies that this user may be interested in.
4. The dataset we plan to use is “MovieLens” (20M version, 190MB zip). Stable benchmark dataset. 20 million ratings and 465,000 tag applications applied to 27,000 movies by 138,000 users. Includes tag genome data with 12 million relevance scores across 1,100 tags.
5. 1. Contend-based recommendation; 2. Collaborative filtering recommendation; 3, LFM (Latent Factor Model, to be confirmed).
6. 1. Precision; 2. Call-back rate; 3. Popularity. More details of these criteria will be explained in the future report.
7. Existing work include the Netflix competition. Also, there are some movie recommendation sites like (IMDB).
8. A python prototype code (single machine version) of the contend-based recommendation system should be completed before 15th Nov.