Author: Dhaval Kapgate
Email: dkapgat1@binghamton.edu

References:

- https://en.wikipedia.org/wiki/Recommender_system
- http://infolab.stanford.edu/~ullman/mmds/ch9.pdf
- http://www.cs.carleton.edu/cs_comps/0607/recommend/recommender/itembased.html

Formula used:

Cosine-base similarity:

$$sim(i,j) = \cos(\vec{i},\vec{j}) = \frac{\vec{i}\cdot\vec{j}}{\|\vec{i}\|_2 * \|\vec{j}\|_2}$$

Prediction:

$$P_{u,i} = \frac{\sum_{\text{all similar items, N}}(s_{i,N} * R_{u,N})}{\sum_{\text{all similar items, N}}(|s_{i,N}|)}$$

Algorithm:

- Read the data from the file and store in a “user x products” or utility matrix
- Create a “user x user” matrix (for calculating the cosine-based similarity for all the users against each users).
- Find the nearest neighbours (for each user find users with least cosine distance).
- Find the products for which there exist no rating by users (entries that are 0 or rated).
- calculate the final prediction by performing the weighted average of deviations from the neighbour’s mean.