COLLABORATIVE FILTERING; PEARSON FORMULA

compute for each user u mean and variance. Let N_u = number of movies rated by user u; R_{um} is the rating of user u for movie m

$$\mu_u = \frac{\sum_m R_{um}}{N_u}$$

$$\sigma_u = \frac{\sum_m R_{um}^2}{N_u} - \mu_u^2$$

normalize each ratings by substracting the user mean and dividing by user variance

$$\bar{r}_{um} = \frac{R_{um} - \mu_u}{\sigma_u}$$

compute user similarity between any two users u and v

$$\rho_{uv} = \frac{1}{\text{movies in common } m} \sum_m \bar{r}_{um} \cdot \bar{r}_{vm}$$

predict the rating for a new movie by accounting for all other users' v rating on the movie

$$predict(u, m) = \mu_u + \frac{\sum_v \rho_{uv} \cdot \bar{r}_{vm}}{\sum_v |\rho_{uv}|} \cdot \sigma_u$$