

STAT 6390 Mini Project 3 | Brett Walker - 3/28/2023

Question 2.

- a. The data is a sparse matrix of movie ratings. The movies are from a wide array of genres, and users are both men and women.
- b. 563 users have over 50 ratings and 626 movies have over 50 ratings
- c. See code for implementation (basically, I followed the slides to implement a modified single-hidden-layer autoencoder which minimized the loss function given in the slides via gradient descent)
- d.

Using $k = 5$:

Recommendations for user 1: "317" - 5, "427" - 5, "402" - 4.54, "399" - 4.51, "483" - 4.47

Recommendations for user 2: "312" - 4.69, "412" - 4.61, "480" - 4.61, "435" - 4.57, "333" - 4.46

Using $k = 10$:

Recommendations for user 1: "427" - 4.71, "317" - 4.70, "399" - 4.30, "402" - 4.16, "483" - 4.13

Recommendations for user 2: "312" - 4.61, "480" - 4.56, "435" - 4.51, "412" - 4.50 "367" - 4.34

The top 5 movie recommendations are the same for user 1 between $k = 5$ and 10 (albeit in a different order), but they change a little for user 2. "333" which was user 2's #5 recommendation drops to #8. Predicted ratings are slightly lower going from $k = 5$ to 10.

- e. k can be tuned as a hyperparameter. Additionally, one can choose a k by performing PCA on the dataset then looking at the number of principal components to include that explains an arbitrarily decided level of variance.