Hyperparameter Testing

To start we chose to try two different distributions to generate priors. Bernoulli and Gaussian.

While looking at the results for multiple users, It looks like Gaussian has a similar accuracy to Bernoulli, but Bernoulli has less total error. Even though Gaussian made the exact correct prediction more often than Bernoulli, Bernoulli has way less total error so it is closer in general to the correct ratings. Therefore, for our recommender system we will implement Bernoulli!

For a brief description of Grid-Search, what you generally feed in is an algorithm which would want to optimize, and which parameters you want to optimize on. For our case we also added in a custom scorer for it to prioritize. This was done so that rather than just counting whether the classification is wrong or right, it uses a loss function to tell how far away it is from its goal. We set the grid search to minimize this loss function to optimize our results.

We use Grid search to find an optimal alpha value for our Bernoulli distribution. Using a scorer that minimized our distance from the actual values, the grid search gave a new alpha value of 9. This change slightly reduces our total error and slightly raises our accuracy. Although it does seem that the probability of getting 100% accuracy goes down by a small margin when calculating n-users, the total error also decreases, meaning that the overall accuracy is improved.\

Single User:

Logo

Description automatically generatedIcon

Description automatically generated

N-Users

Logo

Description automatically generated Icon

Description automatically generated