**Advanced Topics in Recommendation Systems – EX1**

Authors:

Maya Vilenko - 315339747

Lior Tondovski – 307870766

Ilan Vasilevski -

**Part 1 – Stochastic Gradient Descent:**

1. Before we write the explicit objective function, let's denote the components and their meaning:

The objective function:

2. The update step for each parameter:

1.

2.

3.

4.

3. for epoch in epochs:

For each :

Calculate prediction ()

Update the learnable parameters:

1.

2.

3.

4.

for each :

calculate the RMSE or other relevant metric over all samples in the validation dataset

*Note: It is possible to add an early stopping criterion as well. If you chose to do so:*

*check if the stopping criterion is met, if so, the training loop should be stopped, and if not, the training loop should continue to the next epoch. In our implementation the early stopping criterion is checking if the difference in the RMSE in the last two epochs is lower than 0.001.*

4. The hyperparameters that can be tuned are:

* Leaning Rate
* User and Item learnable vector size
* The early stopping criterion
* The number of epochs
* The value of the regularization’s parameters

5. In our validation set, the RMSE measure was calculated after each epoch (another relevant measure may be selected in this case). A decreasing RMSE during the epochs indicates that the model is learning and going towards a local minima. In addition, a sanity check can also be performed by calculating the RMSE of a naive model that always predicts the average rating in the dataset and checking if our SGD model has better performance.

6. we determined an early stopping criterion based on the RMSE calculated over all the samples in the validation set. We check whether the difference between the RMSE in the last two epochs is less than 0.001 as our stopping criterion. If the early stopping criterion met, we break the training loop and returning the trained model and now it can be used on the test set.

7. The chose to implement a class of SGD matrix factorization. The main components are:

1. The init function - A function to determine how many epochs to go through, the learning rate, and the dimensions of the learnable vectors, etc.

2. Run epoch function - iterates over all samples in the training data and updates the learnable vectors and parameters after each iteration.

3. Calculation of evaluation measures function - returns a dictionary of all the desired evaluation measures based on a dataset.

4. Fit function that first initiates the learnable vectors and parameters, iterates over all epochs, calls the run epoch function, and calculates evaluation measures right after, and then checks whether the early ending criteria is met.

8. What is the RMSE, MAE, R^2 and MPR of your model based on the validation set?

**Part 2 – Alternating Least Squares:**

1. The objective function is the same as in the SGD:

2. The update step for each parameter:

Let denote some additional marks:

Updating steps:

1.

2.

3.

4.

3. for epoch in epochs:

For user in M:

For items that user ranked:

Update the user vector and the user item bias parameters

For item in N:

For users that ranked item:

Update the item vector and the items bias parameters

for each :

calculate the RMSE or other relevant metric over all samples in the validation dataset

*Note: It is possible to add an early stopping criterion as well. If you chose to do so:*

*check if the stopping criterion is met, if so, the training loop should be stopped, and if not, the training loop should continue to the next epoch. In our implementation the early stopping criterion is checking if the difference in the RMSE in the last two epochs is lower than 0.001.*

4. The hyper-parameters are:

* User and Item learnable vector size
* The early stopping criterion
* The number of epochs
* The value of the regularization’s parameters

5. In our validation set, the RMSE measure was calculated after each epoch (another relevant measure may be selected in this case). A decreasing RMSE during the epochs indicates that the model is learning and going towards a local minima. In addition, a sanity check can also be performed by calculating the RMSE of a naive model that always predicts the average rating in the dataset and checking if our SGD model has better performance.

6. ALS matrix factorization is implemented as a class that inherits from SGD matrix factorization. The only thing we needed to code from scratch was the run epoch function, which updates the parameters according to the ALS algorithm.

7. What is the RMSE, MAE, R^2 and MPR of your model based on the validation set?

8. Compare the ALS and SGD solutions in terms of implementation, training and quality.