MSDS 630
HW 3
PyTorch
Due: Feb 16, 2023 @ 11:59p

1 Logistic Matrix Factorization with PyTorch (15 pts)

Given a training and validation dataset with amazon ratings of books (`train_books_ratings.csv` and `valid_books_ratings.csv`, found in `data_hw3_2023.zip`, use PyTorch to implement the following model:

$$\hat{y}_{ij} = \sigma(u_i \cdot v_j + b_i + c_j),$$

where $u_i \cdot v_j$ represents a traditional matrix factorization model and $b_i,\ c_j$ are bias terms for the users and items.

Note that this is a binary dataset, and the loss function for this model is binary cross entropy.

Write your code based on the notebook: `MF_with_pytorch.ipynb`.

1.1 Instructions:

- Starting code can be found in `logistic_mf_start.py`. Some functions are pre-written according to their descriptions, others need to be written by you. These functions are indicated with sections like:

```
### BEGIN SOLUTION
### END SOLUTION
```

The specific directions for each function are found in the function docstring.

- Do not submit data to Github. Doing so will result in a loss of 5 points.

1.2 Deliverables:

Submit your completed functions as a Python script entitled "`logistic_mf.py`".

1.3 Evaluation:

Once you've finished coding, make sure you can run:

```
pytest test_logistic_mf.py
```

There are 3 tests to pass: one for the ML model itself, one for the validation function, and one for the one epoch training model. Each pass is worth 5 points. If you fail a test, but your code is on the right track, half credit will be awarded for that test.