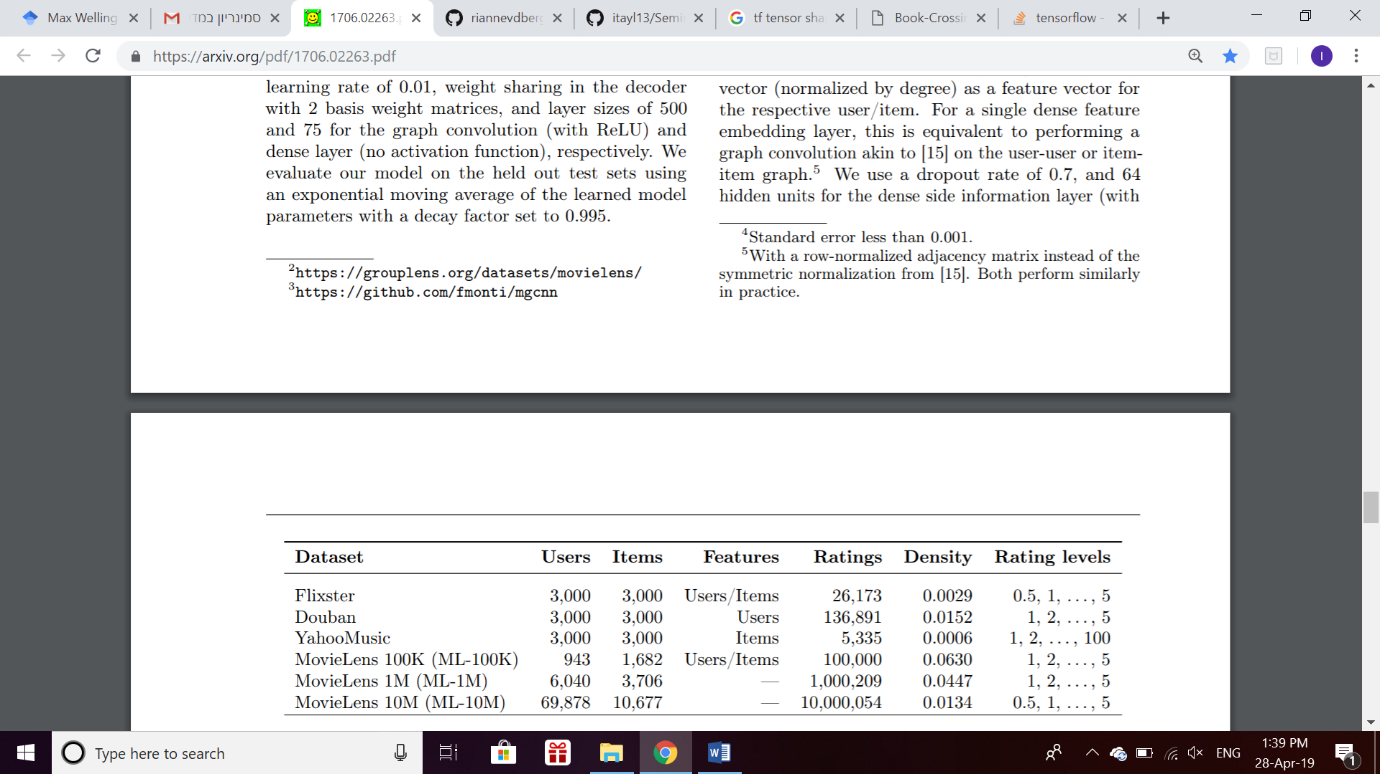
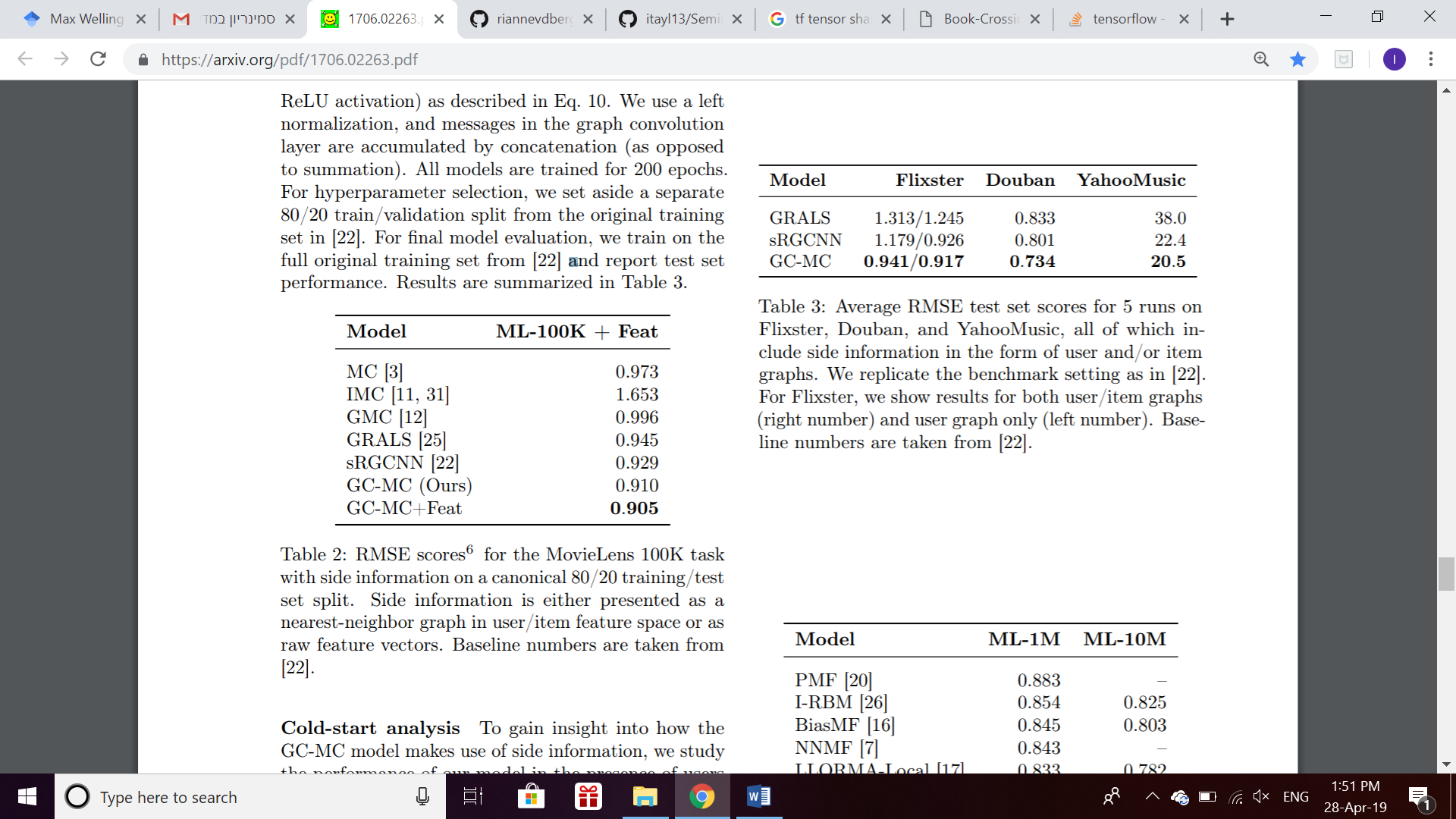
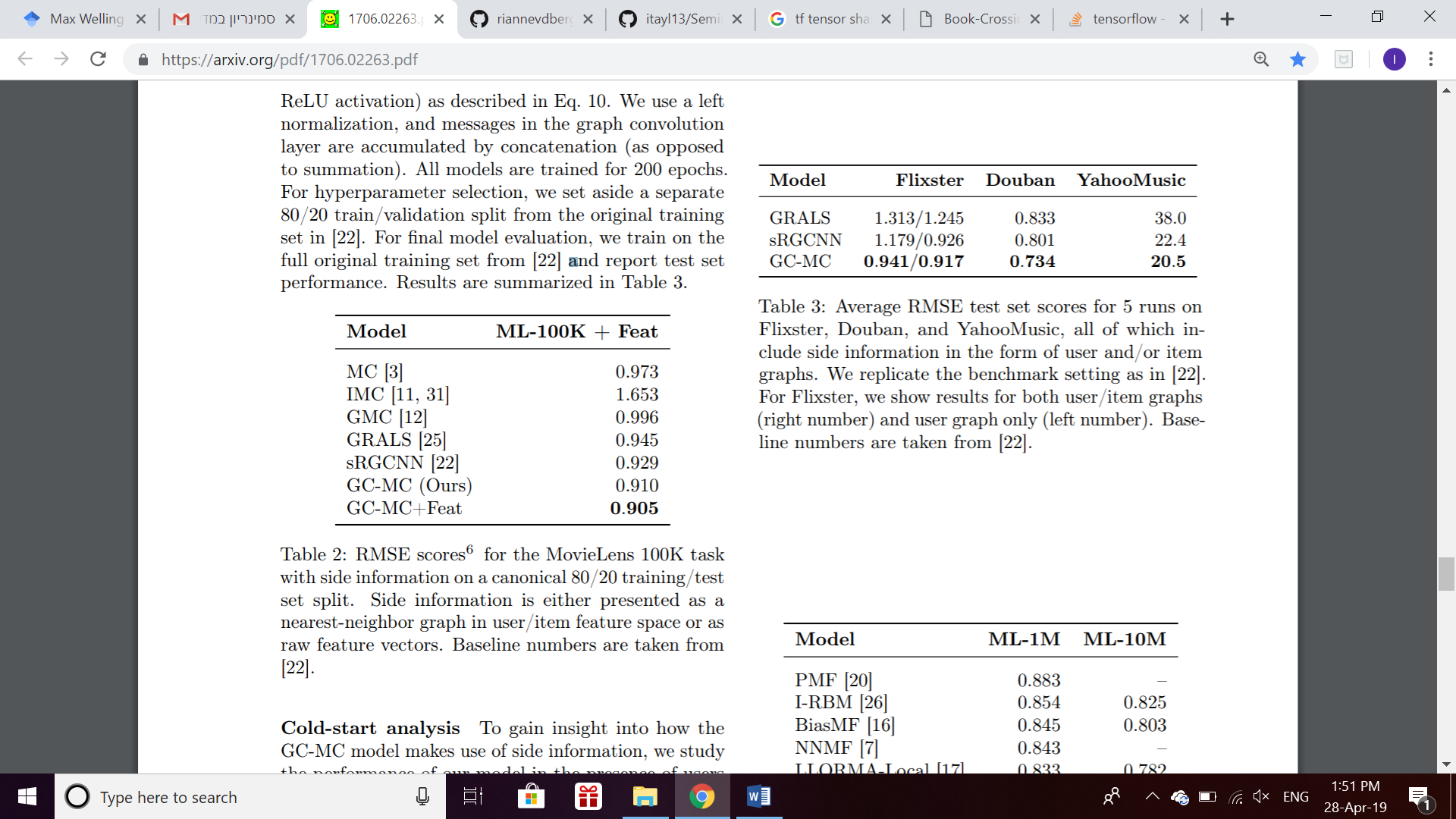
Results

Written by: Itay Levinas  
On the paper: Rianne van den Berg, Thomas N. Kipf, Max Welling, [Graph Convolutional Matrix Completion](https://arxiv.org/abs/1706.02263) (2017)  
Results are on the new dataset of [Book-Crossing](http://www2.informatik.uni-freiburg.de/~cziegler/BX/)

# Datasets and results from the paper

The table below summarizes the given datasets and their properties:

Results are measured using RMSE values. For MovieLens 100K dataset, the option of side-information is tested and does improve the results by a little:  
Note: Another result from the paper that I didn’t check was a cold-start analysis.

# The new dataset

Book-Crossing dataset is a dataset of book ratings, which contains 278,858 users (anonymized but with demographic information) providing 1,149,780 ratings (explicit / implicit) about 271,379 books.

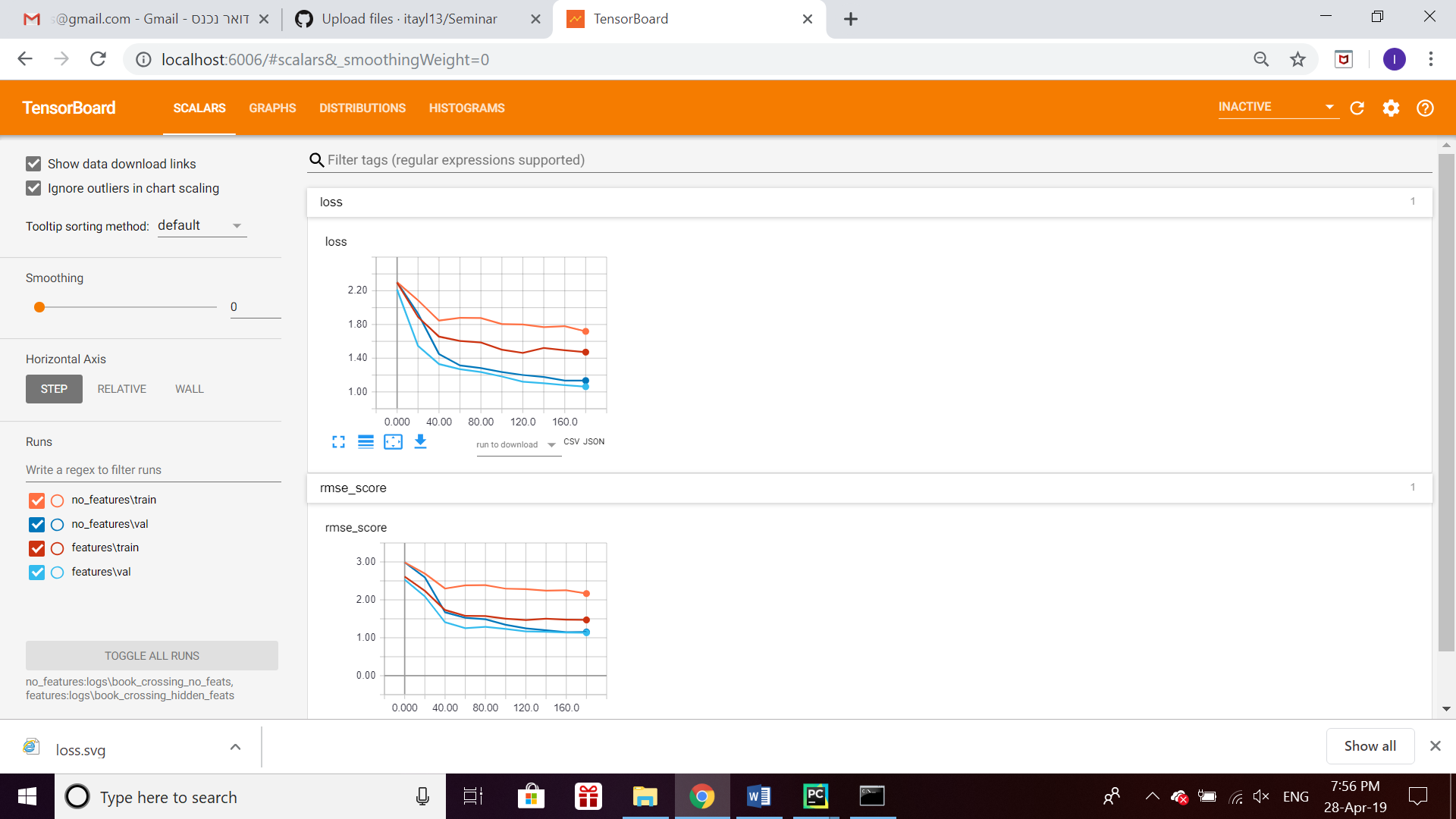
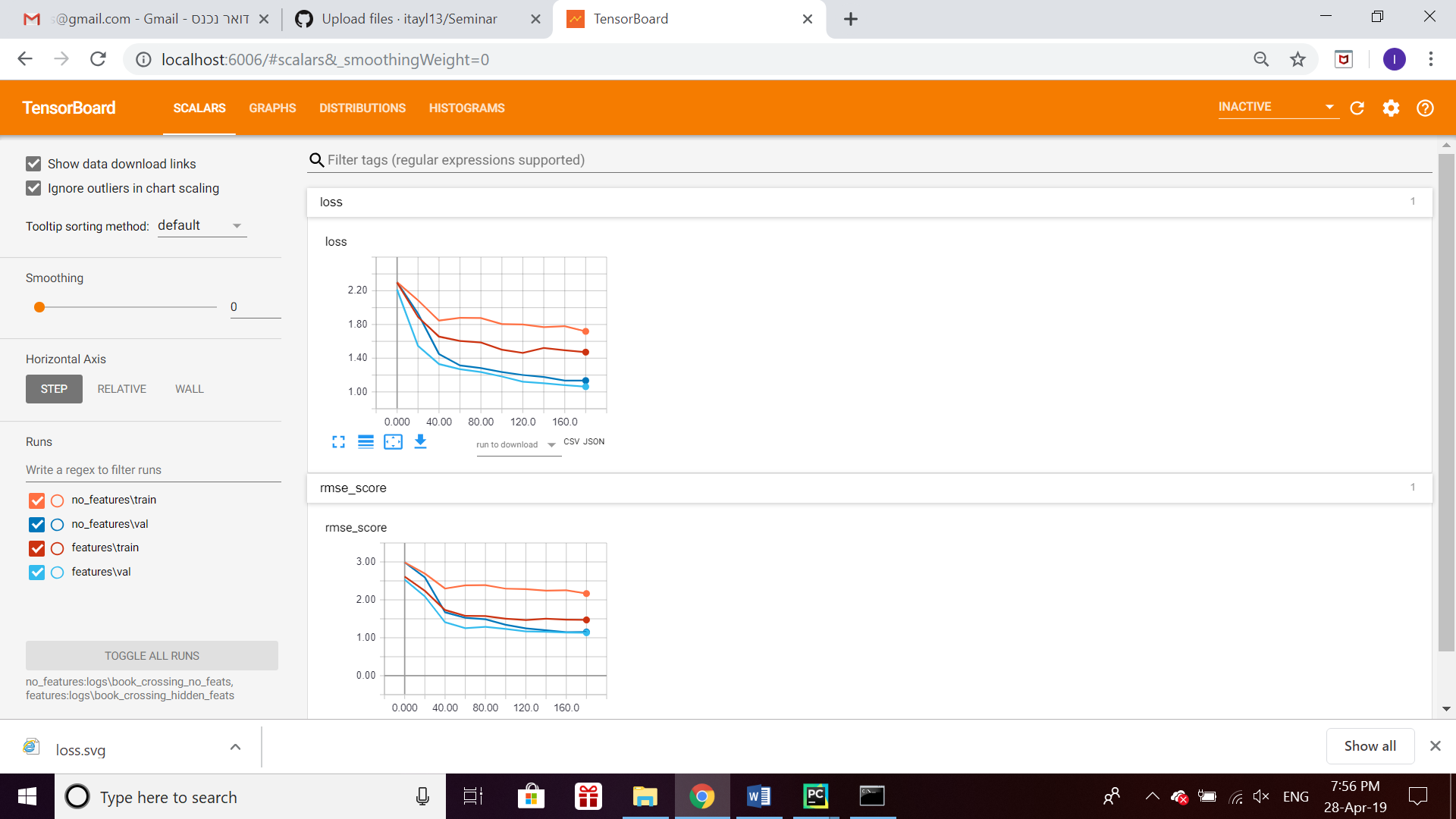
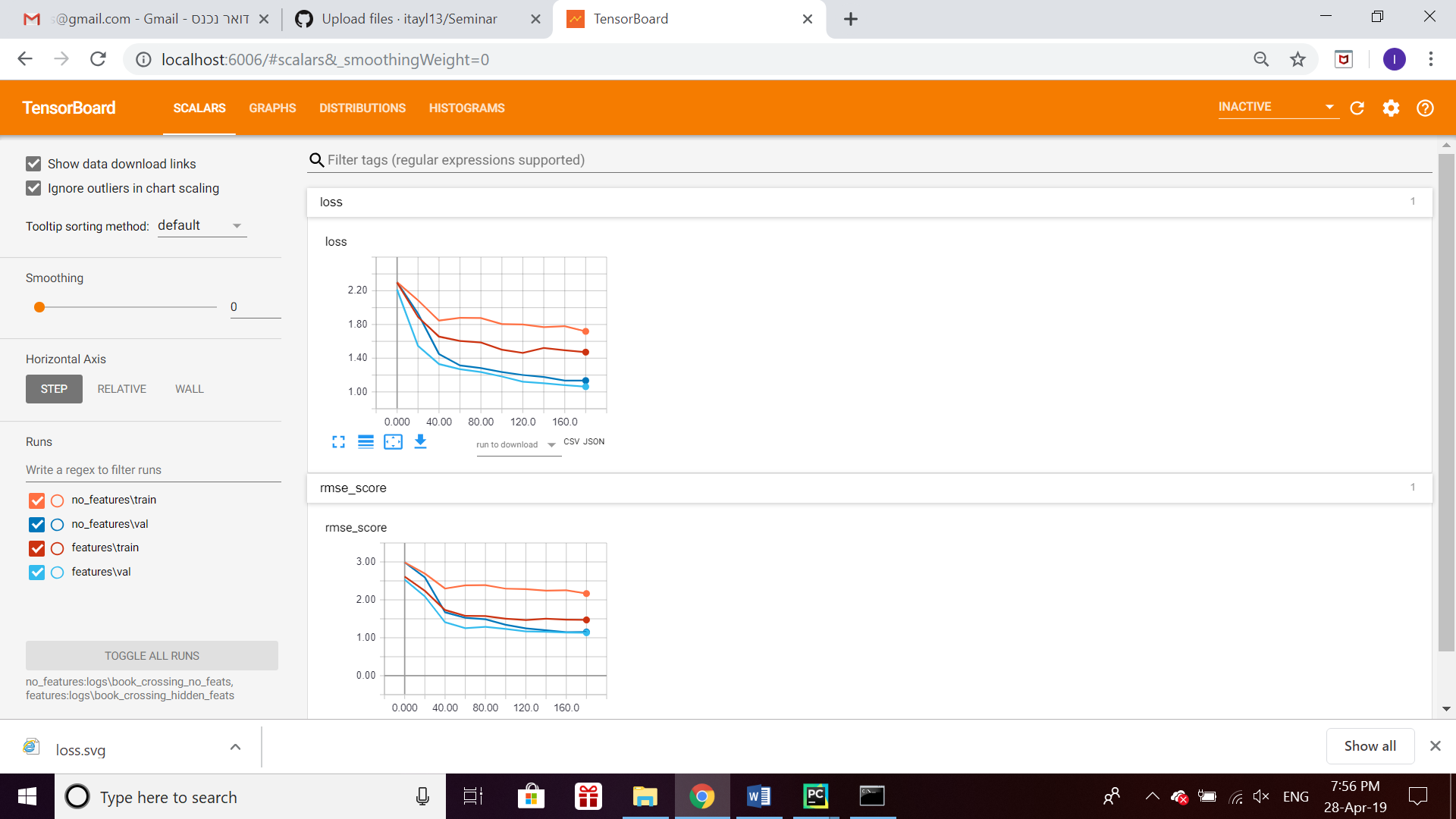
From the whole raw dataset, I had to do two main things to get the final dataset on which I implemented the method:

* User and item features: Some books were removed because they didn’t match the format I used (for example, a book name with symbols that weren’t able to be unicoded). Some Users were mentioning ages such as 200, so I bounded the ages between 2 and 100.   
  The final node features are age for users, author name and year of publication for books.
* Ratings and density: I removed the rating 0 (implicit rating), leaving only the ratings of 1-10. After that, I kept only the users that rated more than 0.00005 of the (valid) books. The final file has only those users and rated books.

To sum up, the data I used contains 3,565 users, 75,711 books and 132,021 ratings (density of approximately 0.0004).

# Results

The method was implemented on the books dataset in two versions: One with hidden features (age for users, author and year of publication for items) and one without them.

The results are presented below:

As seen, my (non-optimal) parameter choice caused overfitting.  
Still, one can see that the using side-information improves the results (The loss values are 1.062 with side-info vs 1.134 without and the RMSE values are 1.132 with vs 1.152 without).

Since I haven’t checked other models, I can’t conclude anything about the performance of this model comparing to others. The only thing I can conclude is that this model can learn this dataset (the loss decreases over the epochs).