

The goal of this assignment is give you practice working with Matrix Factorization techniques.

Your task is implement a matrix factorization method—such as singular value decomposition (SVD) or Alternating Least Squares (ALS)—in the context of a recommender system.

You may approach this assignment in a number of ways. You are welcome to start with an existing recommender system written by yourself or someone else. Remember as always to cite your sources, so that you can be graded on what you added, not what you found.

SVD can be thought of as a pre-processing step for feature engineering. You might easily start with thousands or millions of items, and use SVD to create a much smaller set of “k” items (e.g. 20 or 70).

Notes/Limitations:

- SVD builds features that may or may not map neatly to items (such as movie genres or news topics). As in many areas of machine learning, the lack of explainability can be an issue).
- SVD requires that there are no missing values. There are various ways to handle this, including (1) imputation of missing values, (2) mean-centering values around 0, or (3) <advanced> using a more advance technique, such as stochastic gradient descent to simulate SVD in populating the factored matrices.
- Calculating the SVD matrices can be computationally expensive, although calculating ratings once the factorization is completed is very fast. You may need to create a subset of your data for SVD calculations to be successfully performed, especially on a machine with a small RAM footprint.

You may work in a small group. Please submit a link to your GitHub repository for your Jupyter notebook or RMarkdown file (and rpubs.com link).