

**Recommenders Systems**

**Final Project Report**

**Predicting Movie Ratings**

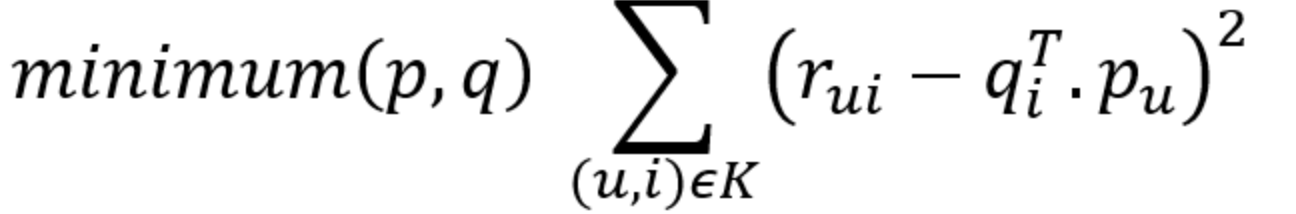
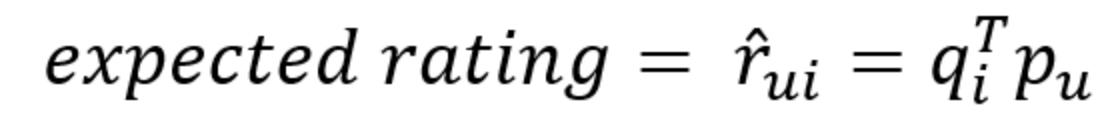
Group We’R, Ayşe Çetinel Yağşi, Engin Erdemli, Erçin Eldeleklioğlu, Yonus Kula

*Istanbul Technical University, MSc in Big Data and Business Analytics*

We installed a Python scikit library called Surprise to build a recommender systems and predict the movie ratings for the given user IDs in the test dataset. We first converted the datasets in the format requested by Surprise, then using cross validation we ran the matrix\_factorization.SVD algorithm.

SVD algorithm was popularized by Simon Funk during the Netflix Prize. SVD in the context of recommendation systems is used as a collaborative filtering (CF) algorithm. Indeed, SVD is a matrix factorization technique that is usually used to reduce the number of features of a data set by reducing space dimensions from N to K where K < N. For the purpose of the recommendation systems however, we are only interested in the matrix factorization part keeping same dimensionality. The matrix factorization is done on the user-item ratings matrix.

The prediction is set as:



Each movie can be represented by a vector `qi`. Similarly each user can be represented by a vector `pu` such that the dot product of those 2 vectors is the expected rating. `qi` and `pu` can be found in such a way that the square error difference between their dot product and the known rating in the user-movie matrix is minimum.

**Our prediction results yielded the average RMSE score of 0.93068 on training data.** For further information, you can refer to the following link for documentation of the algorithm: <http://surprise.readthedocs.io/en/stable/matrix_factorization.html>.