**Collaborative Filtering: Data Set and Matrices Description and Case Study Objectives**

The task was to implement the collaborative filtering learning algorithm and apply it to a dataset of movie ratings. This dataset consists of ratings on a scale of 1 to 5. The dataset has nu = 943 users, and nm = 1682 movies.

The matrix Y (a movies x users matrix) stores the ratings y(i,j) (ranging from 1 to 5).

The matrix R is an binary-valued indicator matrix, where R(i, j) = 1 if user j gave a rating to movie i, and R(i,j) = 0 otherwise.

The number of movie features is n = 100, and therefore, x(i) ∈ℝ 100 and **θ** (j) ∈ℝ 100

Correspondingly, X is a nm x 100 matrix and Theta is a nu x 100 matrix.

The i-th row of X corresponds to the feature vector x(i) for the i-th movie and the j-th row of Theta corresponds to one parameter vector θ(j) for the j-th user. Both x(i) and θ(j) are n-dimensional vectors.

The objective of collaborative filtering is to predict movie ratings for the movies that users have not yet rated, that is, the entries with R(i, j) = 0. This will allow us to recommend the movies with the highest predicted ratings to the user.

The collaborative filtering algorithm in the setting of movie recommendations considers a set of n-dimensional parameter vectors x(1) ,….., x(nm) and θ(1),…..,θ(nu)where the model predicts the rating for movie i by user j as y(i,j) = (θ (j))T x(i). Given a dataset that consists of a set of ratings produced by some users on some movies, you wish to learn the parameter vectors

x(1) ,….., x(nm) and θ(1),…..,θ(nu) that produce the best fit (minimizes the squared error).