🤖

Lecture 14: Recommender Systems / Matrix Factorization

Created	@August 10, 2024 7:28 PM

Matrix Factorization

- user rates a very small amount of movies
- you can create a matrix, where rows are users and columns are ratings for each movi
- most of the matrix will be very sparse
- We have to assume that all movies fit into k categories (such as movie genres)
    - We can create a R_v feature vector with the degree to which each movie fits a particular category
    - Then we can create a user u vector with how much they like each category
    - Then you can put the 2 vectors into a rating function to predict how much a user might like a certain movie
        - dot product of the feature and user vectors
- our goal is to argmin L and R
- some recommender systems
    - popularity
    - nearest neighbor
    - item-item

        - plot a co-occurrence matrix that tells you what percent of people who bought this particular item also bought...
        - you have to normalize first because there might be varying amounts in each category
        - take the top k items in the row that had the highest correlation weights and then recommend that to the user
        - for multiple items:
            - take the average of the scores/ratings of each item for all of those categories
            - could weight recent purchases more than older ones
    - feature-based
        - store information for each feature of each movie AND also the user (age, gender, etc)
    - Matrix factorization
        - actual data points will be the original sparse matrix
        - predict data points by learning the movie and user weights (ratings for each genre, author, date published, time of day, etc) and then dot producting the 2 vectors for the particular user or movie that you want to test
        - suffers from cold start problem
        - You can use ridge regularization to fix the problem of having infinite solutions that minimize the MSE on the quality metric
        - can be used for both supervised and unsupervised learning
    - hybrid model