Abstract

A recommendation engine is an algorithm that is able to predict what a user may or may not like among a list of given items. Those predictions can be used to maximize revenue of a eCommerce service by increasing user conversion rate. The goal of this project is to develop, evaluate and implement a recommendation engine for a particular eCommerce service.

Problem Statement

The general task of Recommeder System is to provide each user with a ranked list of items. We are going to build two recommender systems:

1. a custom algorithm for SVD-based clustering
2. classic collaborative filtering via Factorization Machine

and test them against a baseline of most popular items.

Target metric is top@n - if an item from recommendations was purchased it equals to 1, else 0.

SVD clustering

Rough description of an algorithm for SVD-based clustering.
Let k be a number of clusters.

1. User-item matrix A, where $A_{ij} = 1$ if user i purchased item j at least once is constructed. Each row in this matrix corresponds to a user.
2. Truncated SVD decomposition with k components of A is computed
3. Matrices produced by SVD decomposition are interpreted as **user-to-cluster** matrix and **cluster-to-item** matrix.

Factorization Machines

Factorization Machine is a generalized linear model that uses nested and factorized variable interactions. Implementation of FMs provided by fastFM library is used. FM is fitted to a same user-item matrix as described in SVD clustering, but its parameters are optimized via MCMC sampling.

Testing: local

For local validation rolling cross-validation is used. User purchases are predicted for a next 30 days from a *current date*, then *current date* is increased by some offset. Top-n metric for all recommenders is plotted as a curve on metric-time plot.

Testing: A/B

Production testing is done via A/B test. All active users are splitted in random fashion into 3 groups:

1. receives default recommendations
2. receives recommendations based on SVD clustering
3. receives FM recommendations

Performance of those groups is evaluated via analysis of a daily refreshing dashboard. Indicators that are analyzed include but not limited to total revenue per group, average revenue per unique sessions.