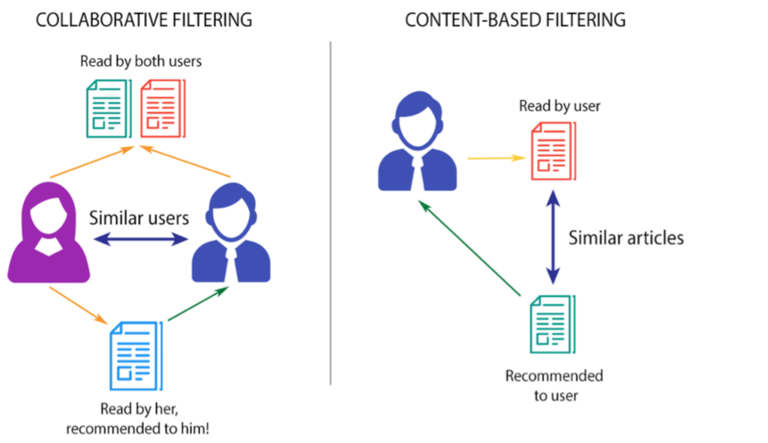
Ensemble of libfm and xlearn for Item Recommendation

Yan Hao (honeyhaoyan@sjtu.edu.cn)
ACM Class, Zhiyuan College, Shanghai Jiao Tong University

Background

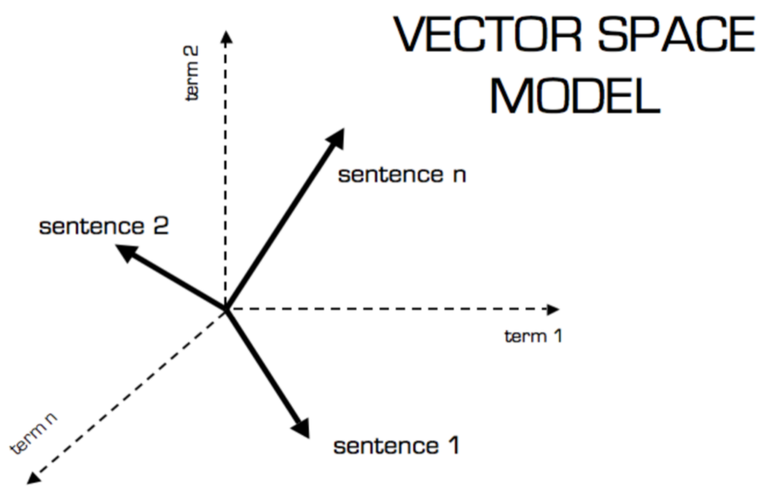
General idea for movie recommendation

approach 1: recommend movies that similar uses are fond of
approach 2: recommend movies that are similar to movies the user has watched and given a positive respond



Vector Space Model

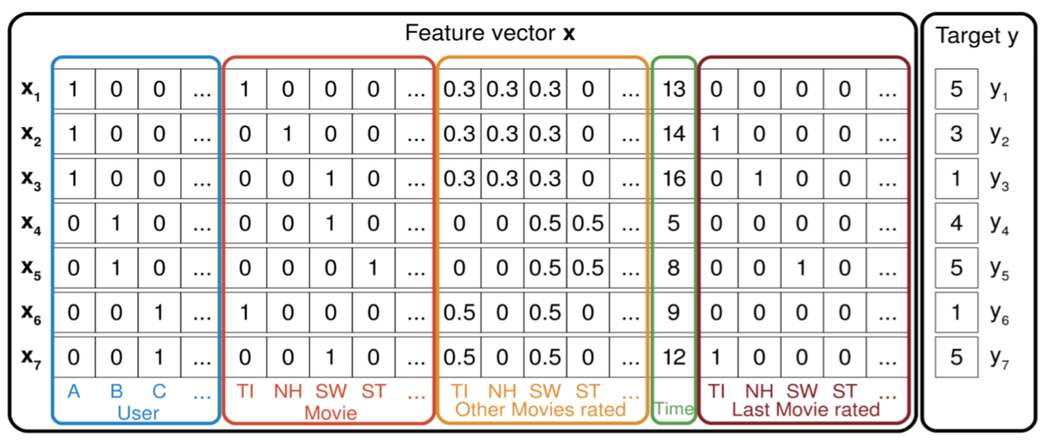
Set up a vector space model and each dimension of it stands for a feature. Training the model using machine learning.



Model

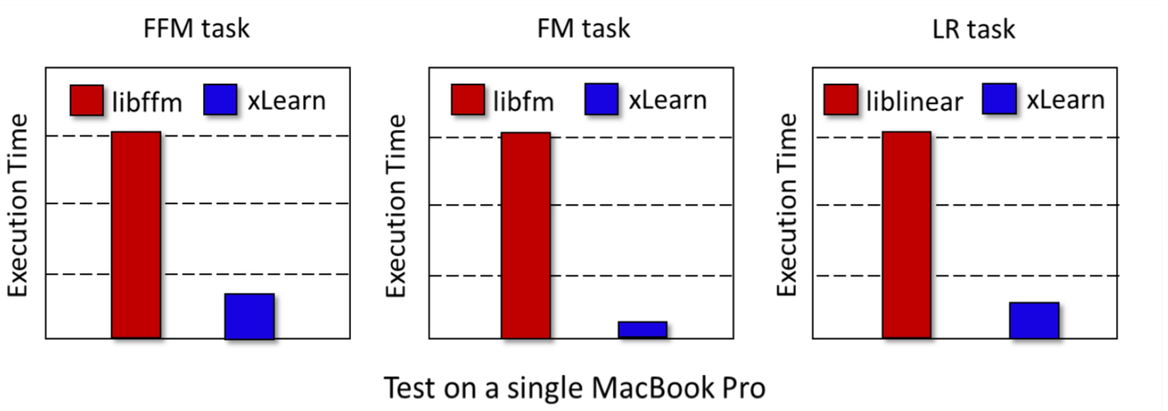
Libfm

Libfm is a software implementation for factorization machines that features stochastic gradient descent (SGD) and alternating least squares (ALS) optimization as well as Bayesian inference using Markov Chain Monte Carlo (**MCMC**).
Here is a example of how libfm works.

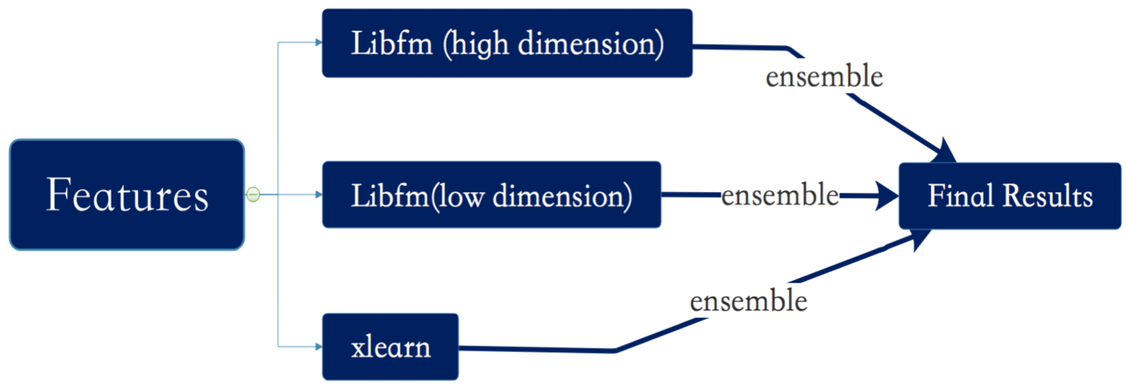


Xlearn

Xlearn is a high performance, easy-to-use, and scalable machine learning package, which can be used to solve large-scale machine learning problems. It is possible alternation of libfm.
The training speed of xlearn is much faster than libfm.



My Approach



Feature Select

- User feature: user ID
- Item feature: the elements in the movie, for instance, whether it is a comedy, fiction movie, documentary, etc.
- Time feature: date & time

Train Model

$$a = \text{the output of low dimension libfm model}$$
$$b = \text{the output of high dimension libfm model}$$
$$c = \text{the output of xlearn model}$$

Final Result

$$\text{output} = \frac{a+b+c}{3}$$

Results

Model	Score	Rank
ensemble of Libfm and xlearn	29.79887	3/35
ensemble of Libfm with different dimensions	30.50439	23/35
Libfm	30.96461	31/35
xlearn	30.81886	29/36

It turns out that :

- Whether we use libfm and xlearn, a sigle model is not enough to fully illustrate the recommendation system.
- The ensemble of libfm models with different parameters works better than a single model, but is still not good enough.
- Calculate the ensemble with libfm and xlearn and get the best result, which rank 3 in the final test.

#	△pub	Team Name
1	—	**Au Revior~**
2	—	**Zhanghao Wu**
3	—	**Haoyan**