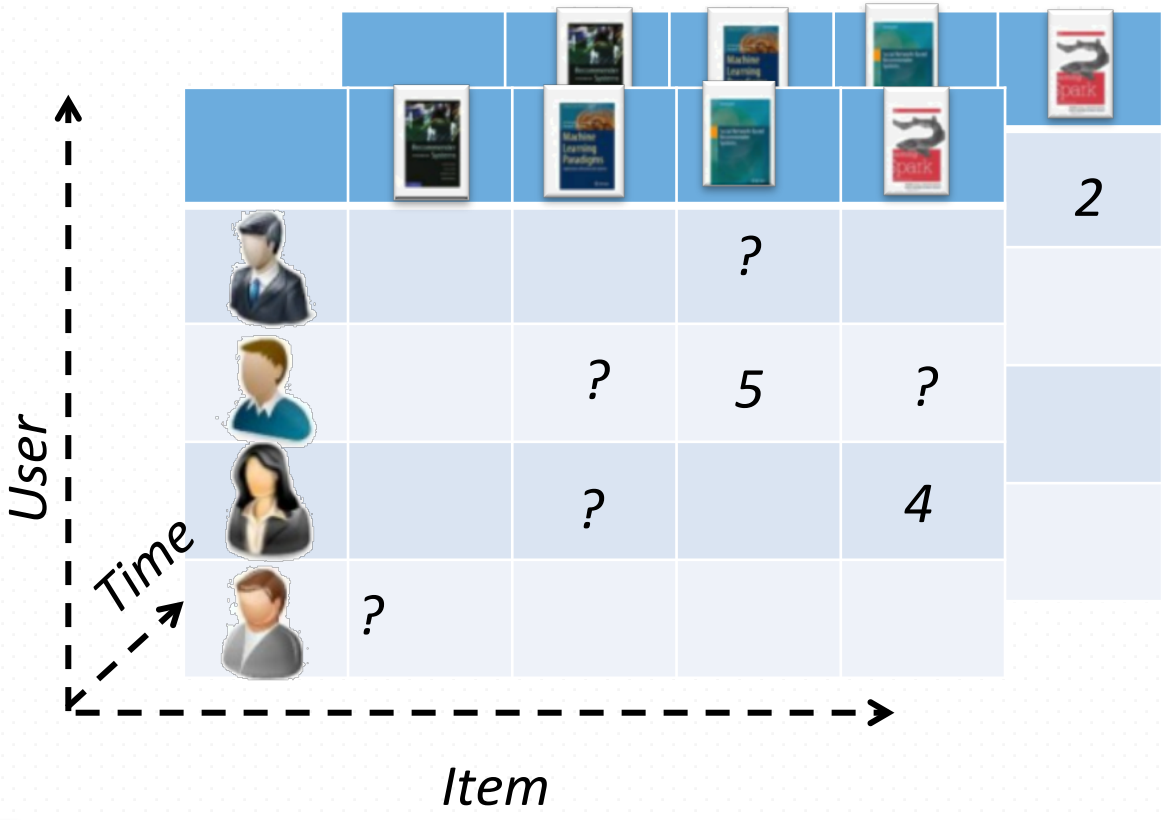
Neural Tensor Factorization for Temporal Interaction Learning

Xian Wu, Baoxu Shi, Yuxiao Dong, Chao Huang, Nitesh Chawla

Department of Computer Science and Engineering
University of Notre Dame

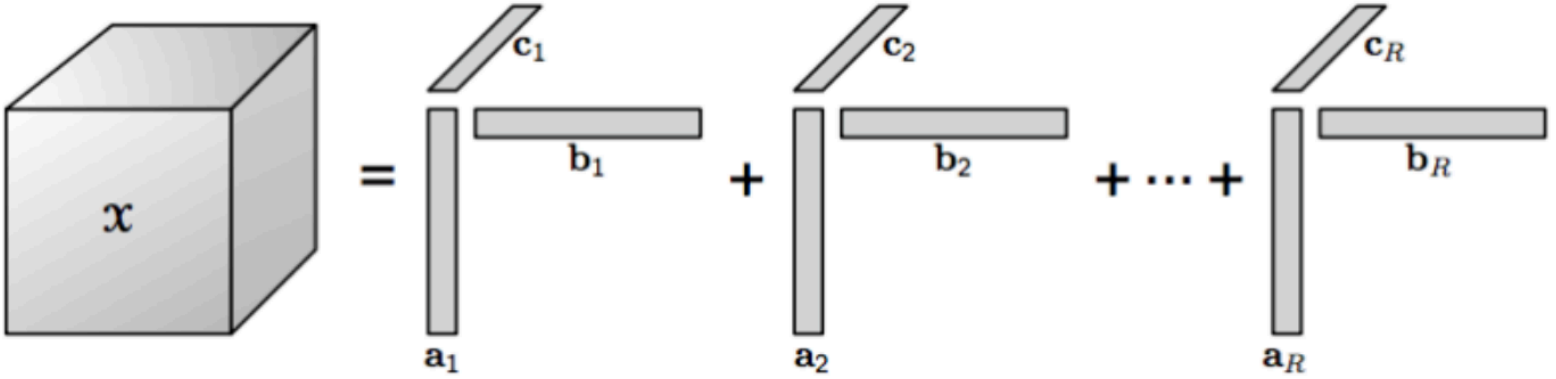
WSDM 2019

Background&Problem Formulation



Objective:
learn a predictive model that can infer the unknown values with the observed ones

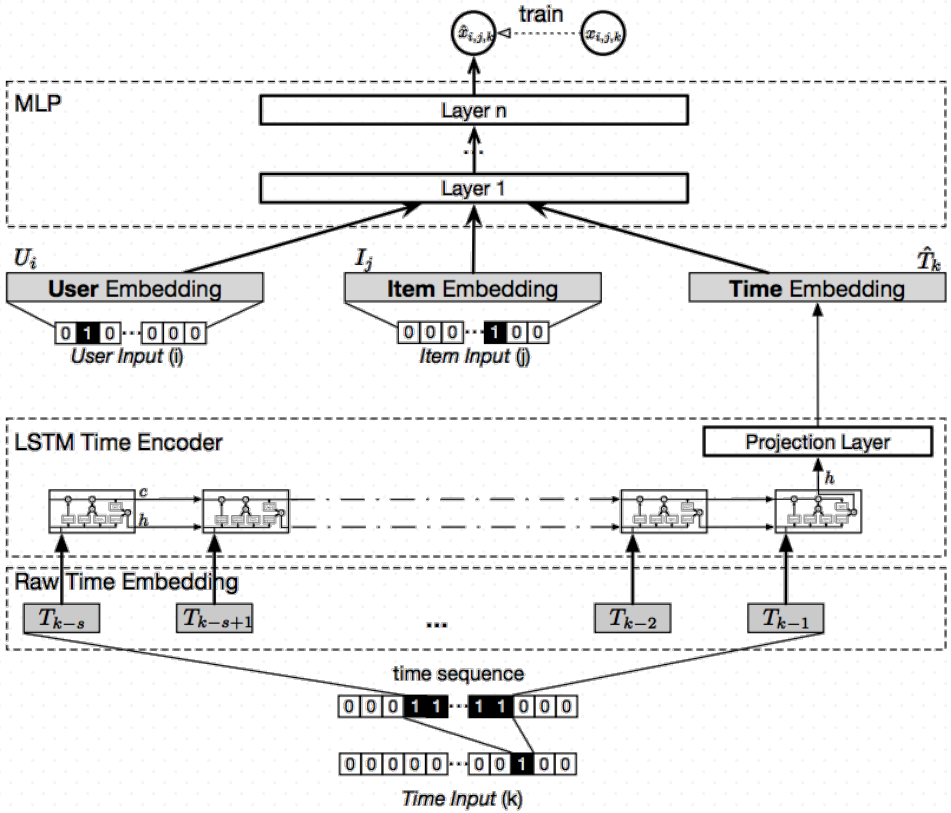
Limitations



Limitation 1: Independence assumption in temporal dimension

Limitation 2: Linear interaction across dimensions

The Proposed Framework-Neural Tensor Factorization



To capture the complex temporal dynamics, we utilize LSTM to encode the evolving interactions.

To model the non-linearity of multi-dimensional interactions, we use MLP on top of the first component

Experimental Settings

Netflix Rating Data

Beselines:

1.Latent Factor based Models: Probabilistic Matrix Factorization (PMF), Bayesian Probabilistic Matrix Factorization (BPMF) , Bayesian Probabilistic Tensor Factorization (BPTF)

2.Neural network based Models: Temporal Deep Semantic Structured Model (TDSSM), Recurrent Recommendation Networks (RRN), Neural Collaborative Filtering (NCF)

Metrics: Root Mean Square Error(RMSE) and Mean Absolute Error (MAE)

Experiments

Month - 2004	Jan		Mar		May	
Metrics	RMSE	MAE	RMSE	MAE	RMSE	MAE
PMF	0.9385	0.7331	0.9274	0.7263	0.9243	0.7171
BPMF	0.9879	0.7686	0.9829	0.7659	0.9741	0.7541
TDSSM	1.0031	0.8001	1.0386	0.8488	0.9897	0.7886
RRN	1.0062	0.7936	0.9901	0.7798	0.9721	0.7584
NCF	0.9498	0.7517	0.9364	0.7357	0.9421	0.7408
NTF*dot*	0.9869	0.7763	0.9736	0.7702	0.9600	0.7523
NTF(*ReLU*)	0.9192	0.7204	0.9111	0.7169	0.9127	0.7131
NTF(*sigmoid*)	0.9158	0.7178	0.9113	0.7148	0.9141	0.7110
NTF(*tanh*)	0.9178	0.7187	0.9128	0.7185	0.9135	0.7111

density degree: Jan-3.67%, Mar-3.86%, May-4.11%

Observation:

the sparser the data, the larger performance gain we can achieve

Conclusion

We developed a Neural network based Tensor Factorization(NTF) to model temporal interactions.

Extensive experiments show that NTF significantly outperforms baseline methods.

Thanks