Recbole algorithm overview (continued)

MCCLK

Dataset

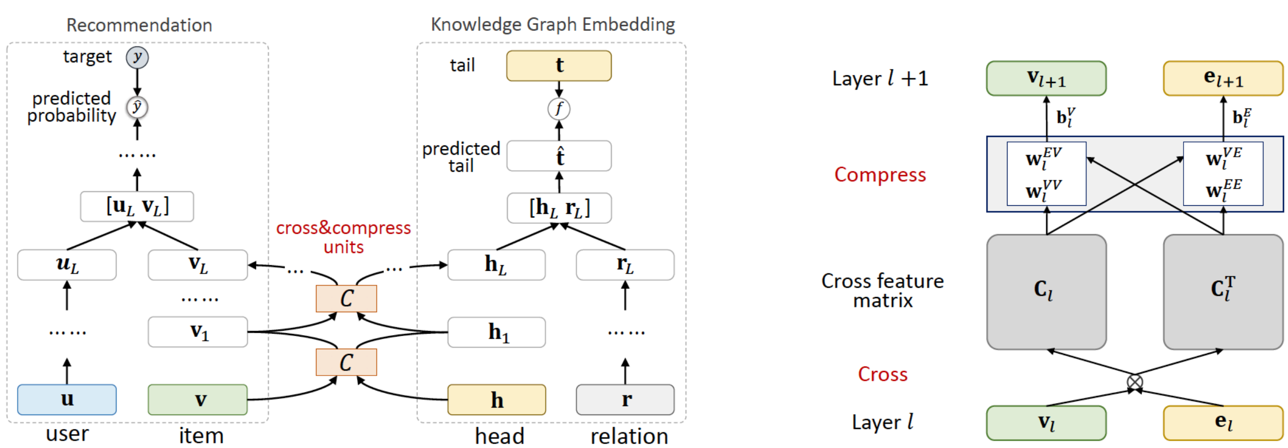
- Book-Crossing
    - book
- MovieLens-1M
- LastFM
- Metric
    - CTR: AUC, F1
    - top-K: Recall@K∈{5, 10, 20, 50, 100}
- 适用数据：sparse, interaction + attribute data

以下是几个比较老的模型：

MKR

Overview

- Multi-Task Feature Learning for Knowledge Graph Enhanced Recommendation



(a) Framework of MKR (b) Cross&compress unit

- 结合KG embedding + 推荐两个任务
    - connected by Cross&compress Unit
        - Cross: item和entity做cross product，生成交叉特征矩阵

        - Compress: 两种cross product加权，输出到下一层
    - Recommender模块，KG模块
        - L层MLP后做Cross&compress

Dataset

- MovieLens-1M
- Book-Crossing
- LastFM
- Bing-News
- Metric
    - CTR: AUC, Accuracy
    - top-K: Precision@K={2, 5, 10, 20, 50}
- 适用数据：sparse, side info（文本，e.g. 电影/图书名称，新闻标题）

KGCN

Overview

- Knowledge Graph Convolutional Networks for Recommender
- 和之前的工作大同小异
    - 基本Pipeline：
        - 聚合neighbor信息（各种aggregator，GNN）
        - 建模高阶关系（multiple-hop）
        - 定义score机制计算概率，e.g. MLP, 内积

Dataset

- MovieLens-20M
- Book-Crossing
- LastFM
- Metric
    - AUC
    - F1
    - Recall@K={1, 2, 5, 10, 20, 50, 100}
- 适用数据：与前面几个模型相同

KGNNLS

Overview

- Knowledge-aware Graph Neural Networks with Label Smoothness Regularization for Recommender Systems
- 将KG转换为user-specific的加权图，用GNN计算每个user对应的item embedding
    - 把KG看作无向图
    - 边权重可训练
- 标签平滑假设：使用标签平滑正则化，确保知识图谱中相邻item有相似的用户相关性label
    - 相当于图上的标签传播
    - 把某一item看作unlabeled
        - 用其他所有实体信息predict出一个label
        - 在与实际label做cross entropy作为正则项

Dataset

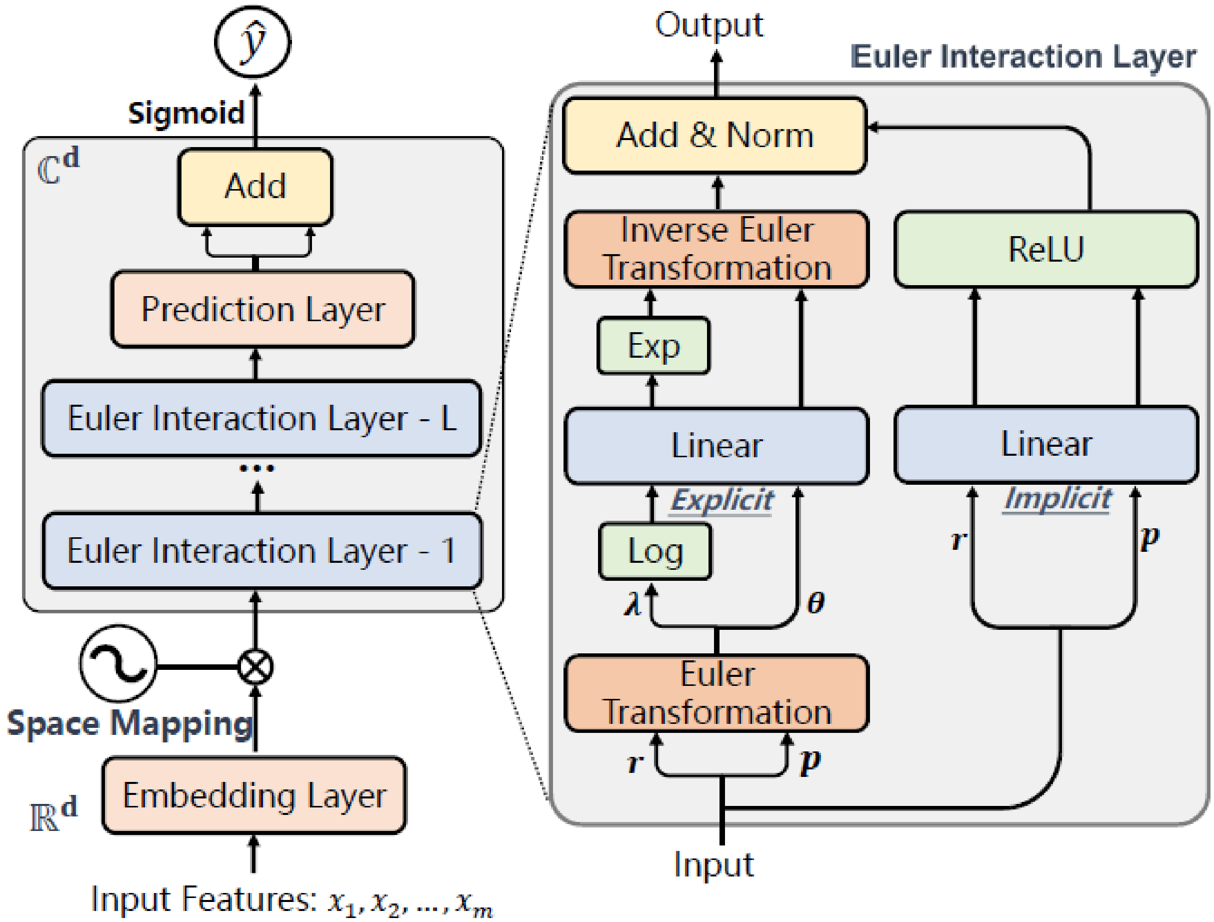
- MovieLens-20M (movie)
- Book-Crossing (book)
- LastFM (music)
- Dianping-Food (restaurant)
    - 大众点评数据
- Metric
    - top-k: Recall@{2, 10, 50, 100}
    - CTR: AUC
- 适用数据：交互数据，实体关系

Context-aware

EulerNet

Overview

- Adaptive Feature Interaction Learning via Euler’s Formula for CTR Prediction



- 用欧拉公式在复向量空间中建模特征交互
  - exponential -> 模和相位linear combination
- 统一捕捉显式和隐式交互，相互增强
- 自适应地学习特征交互阶数

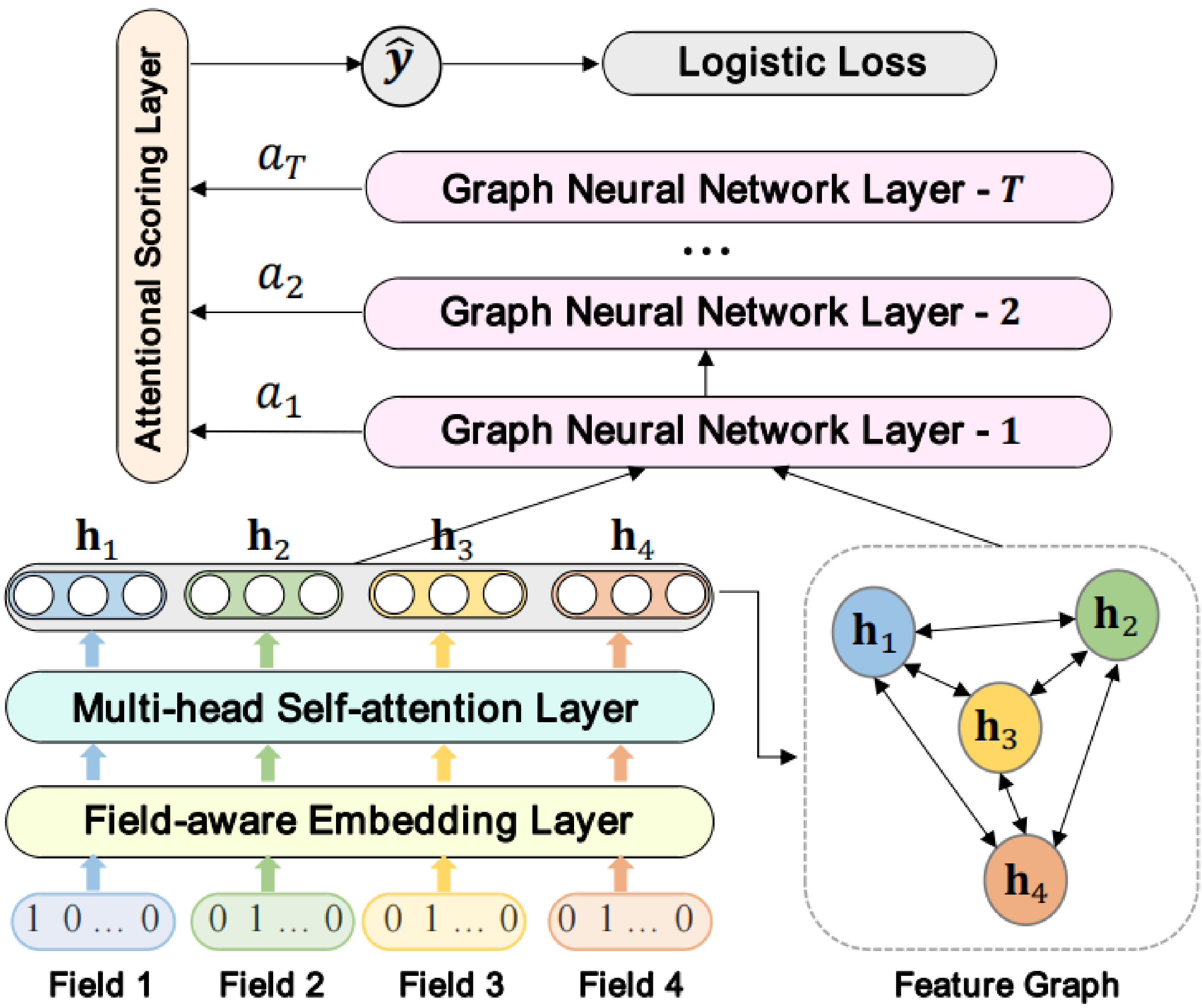
Dataset

- Criteo
  - 7天用户日志，CTR
- Avazu
  - CTR as well
- MovieLens-1M
- Metric
  - AUC
  - Logloss
- 适用数据：稀疏，高维特征数据（e.g. 用户行为日志中提取的多字段特征）

FiGNN

Overview

- Modeling Feature Interactions via Graph Neural Networks for CTR Prediction



- 解决问题：现有方法简单拼接字段embedding
- 多字段特征表示为图
  - raw feature先过self-attn
  - 随后接传统GNN pipeline
    - 状态聚合 -> 状态更新（GRU，残差）

Dataset

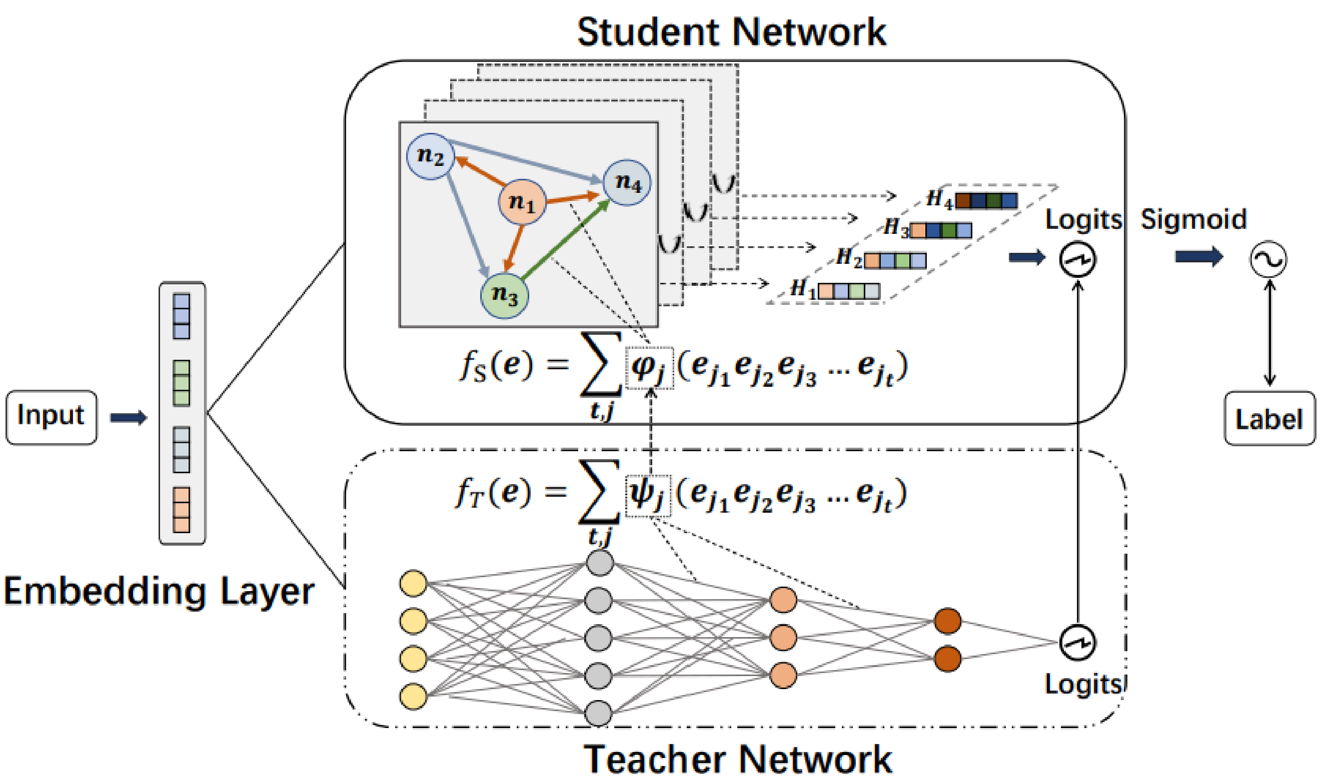
- Criteo
- Avazu
- Metric
  - AUC

- Logloss
- 任务：CTR
- 适用数据：
  - 多字段分类特征数据
  - 稀疏，高维特征
  - 正负样本不平衡

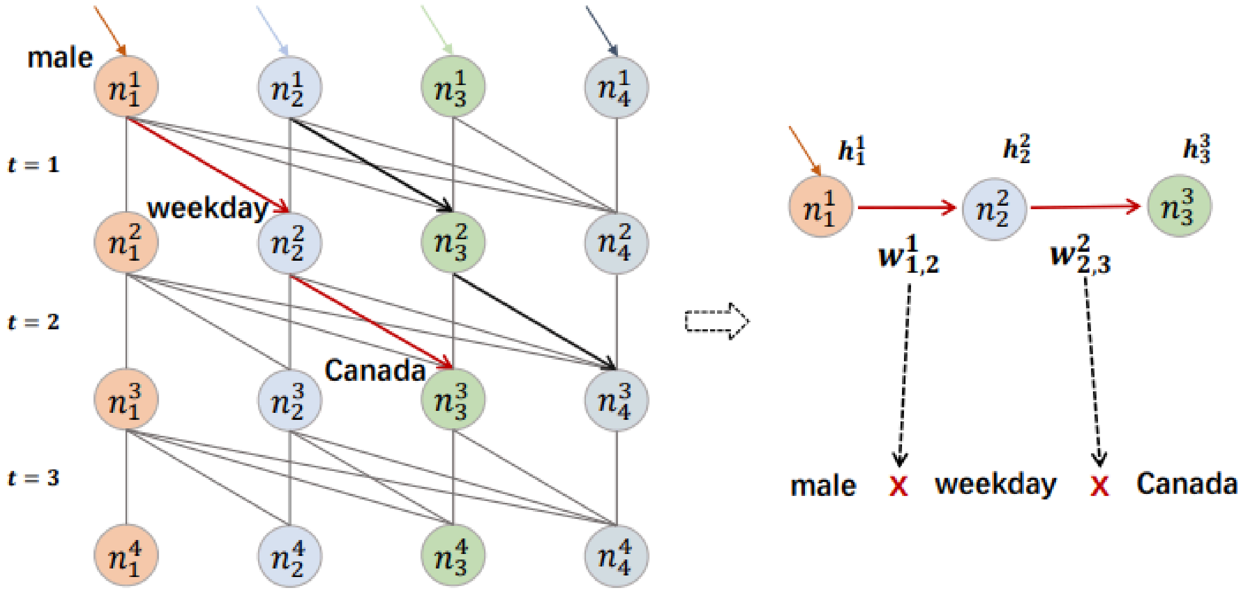
KD_DAGFM

Overview

- Directed Acyclic Graph Factorization Machines for CTR Prediction via Knowledge Distillation.



- “基于知识蒸馏的有向无环图因子分解机”
- 用知识蒸馏从复杂的教师模型向轻量级学生模型传递知识
  - loss为均方误差
- 用dp思想传播hidden state
  - 每个k阶特征交互可以对应到第一层的唯一路径
  - 每个传播层，每个节点聚合所有邻居的状态



- 改进模型：KD-DAGFM+
  - 用于蒸馏显式和隐式特征交互
  - 最后加MLP
  - 教师模型使用xDeepFM，DCNV2，AutoInt+，FiBiNet

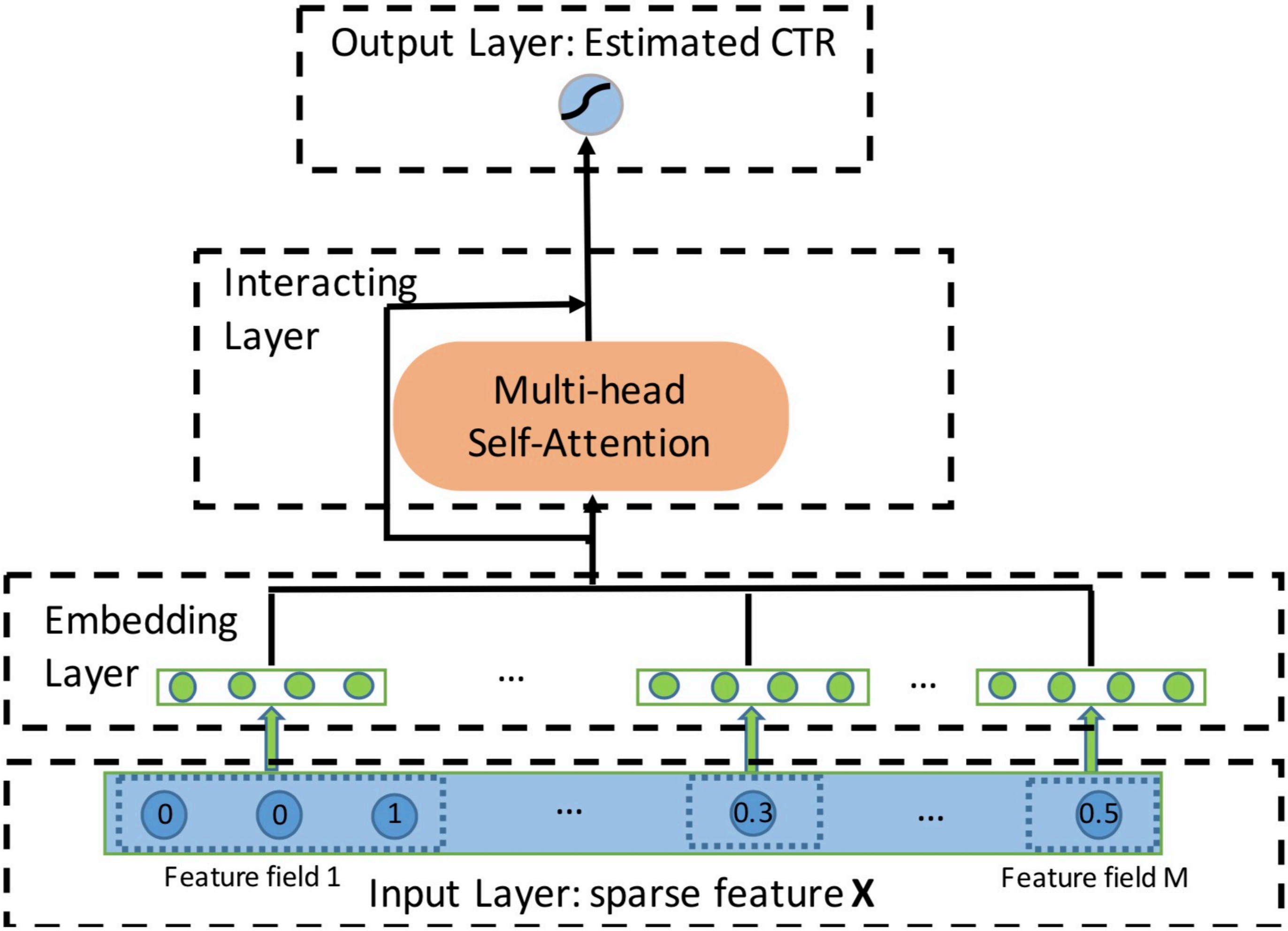
Dataset

- Criteo
  - CTR benchmark
- Avazu
- MovieLens-1M
- WeChat
- Metric
  - AUC
  - Log Loss
- 适用数据：
  - CTR相关
    - u-i交互，用户画像，item特征
    - 大规模工业数据
    - 高维，稀疏，noisy

AutoInt

Overview

- Automatic Feature Interaction Learning via Self-Attentive Neural Networks



- Attention is all you need!
- 把数值和分类特征映射到同一低维空间
- 多头自注意力，残差连接

Dataset

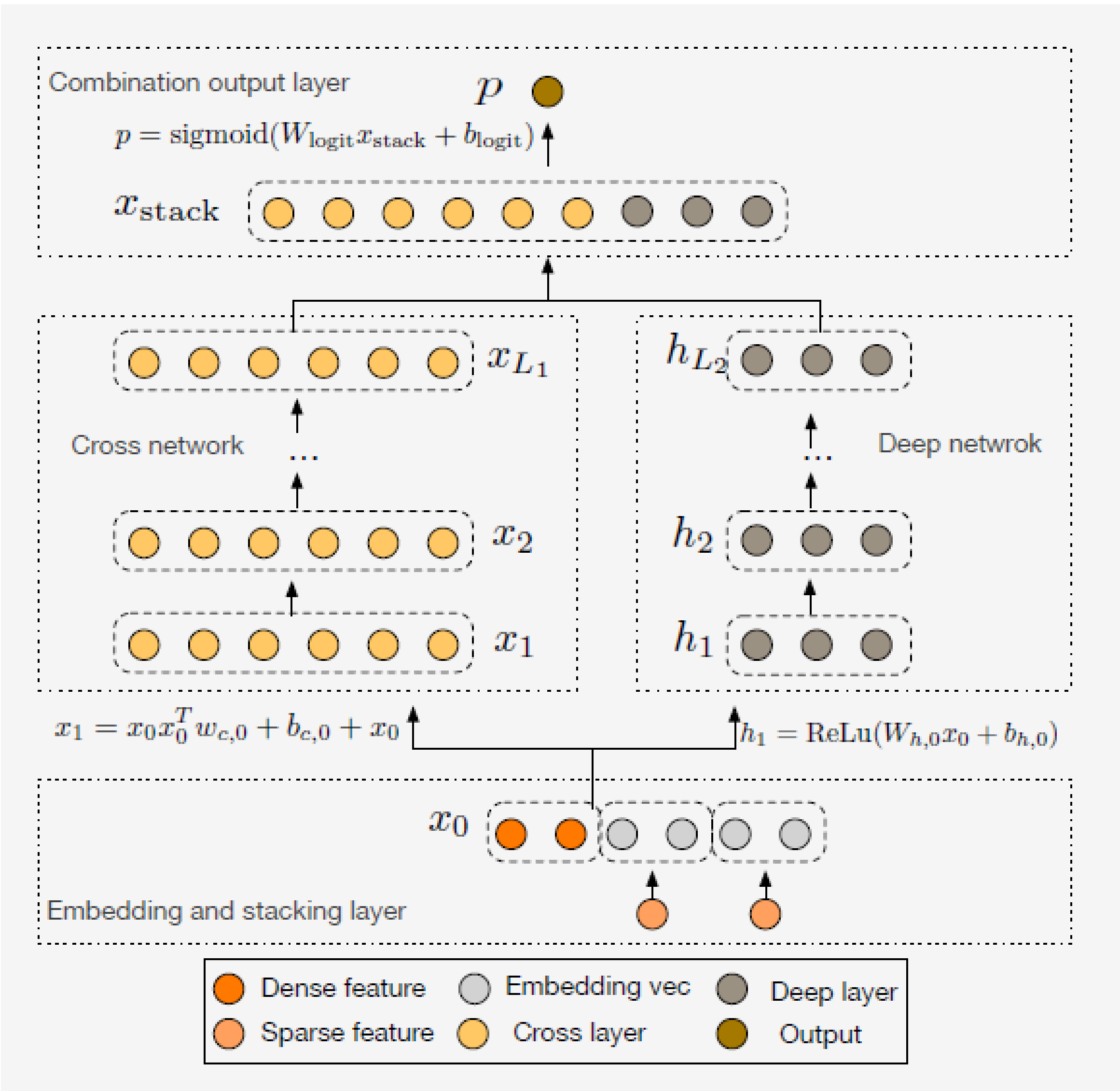
- Criteo
- Avazu
- KDD12
    - KDDCup 2012
- MovieLens-1M
- Metric
    - AUC

- Logloss
- 适用数据：
  - 稀疏，高维
  - 包含数值和分类特征

DCN

Overview

- Deep & Cross Network for Ad Click Predictions



- Deep: MLP
- Cross Network: 每一层显式应用特征交叉，自动计算所有可行的特征组合
- Scalability↑；发现New item

Dataset

- Criteo
- Metric
    - Logloss
- 适用数据：和上述几个模型相似

DCN V2

Overview

- 对DCN的工业级应用改进
- 用低秩结构近似特征交叉，实现更好的性能和延迟的rade-off
- 混合专家架构（MoE），把矩阵分解到多个子空间中，再用gating机制聚合

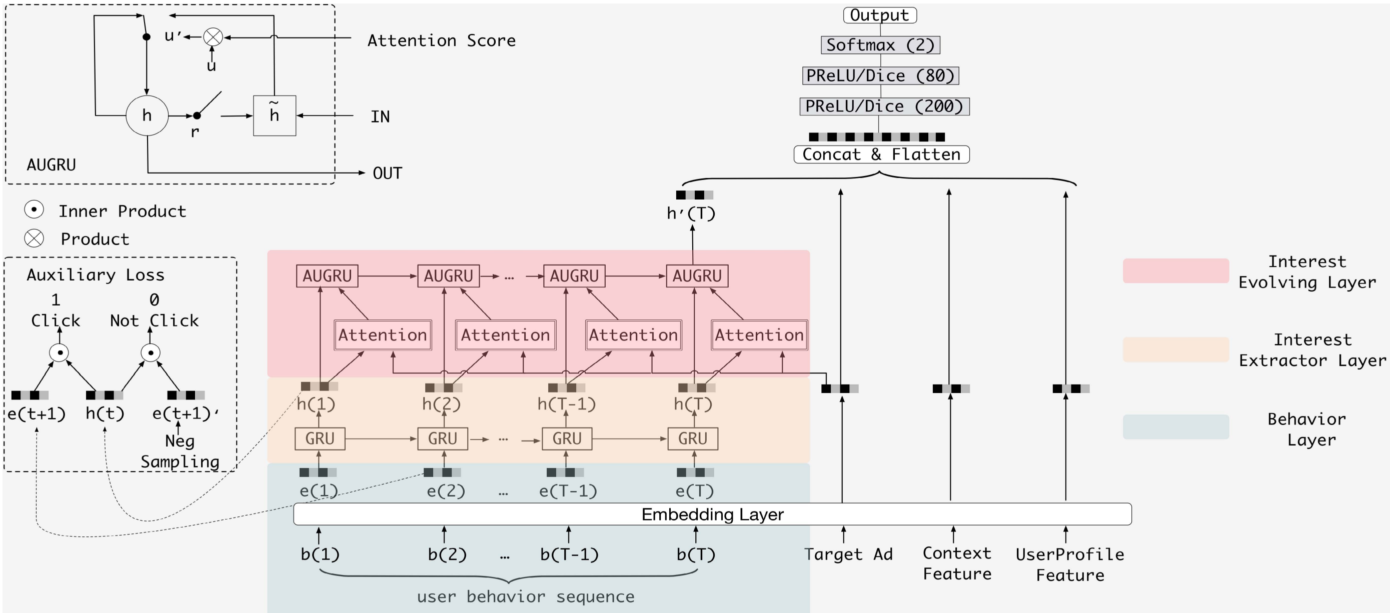
Dataset

- Criteo
- MovieLen-1M
- Metric
    - Logloss
    - AUC
- 适用数据：
    - 大型稀疏特征，web-scale生产数据
    - 低秩结构
        - 权重矩阵具有较大的奇异值差距 / 快速的spectrum decay模式
    - 其他相似

DIEN

Overview

- [Deep Interest Evolution Network for Click-Through Rate Prediction]



- 解决问题：
  - 用户行为背后的隐藏兴趣
  - 兴趣随时间变化
- w/ sequential model
- AUGRU: GRU with attentional update gate
  - 用注意力得分缩放更新门的所有维度，确保与目标item关系较弱的兴趣对隐藏状态的影响较小

$$\tilde{\mathbf{u}}_t' = a_t * \mathbf{u}_t',$$

$$\mathbf{h}_t' = (1 - \tilde{\mathbf{u}}_t') \circ \mathbf{h}_{t-1}' + \tilde{\mathbf{u}}_t' \circ \tilde{\mathbf{h}}_t',$$

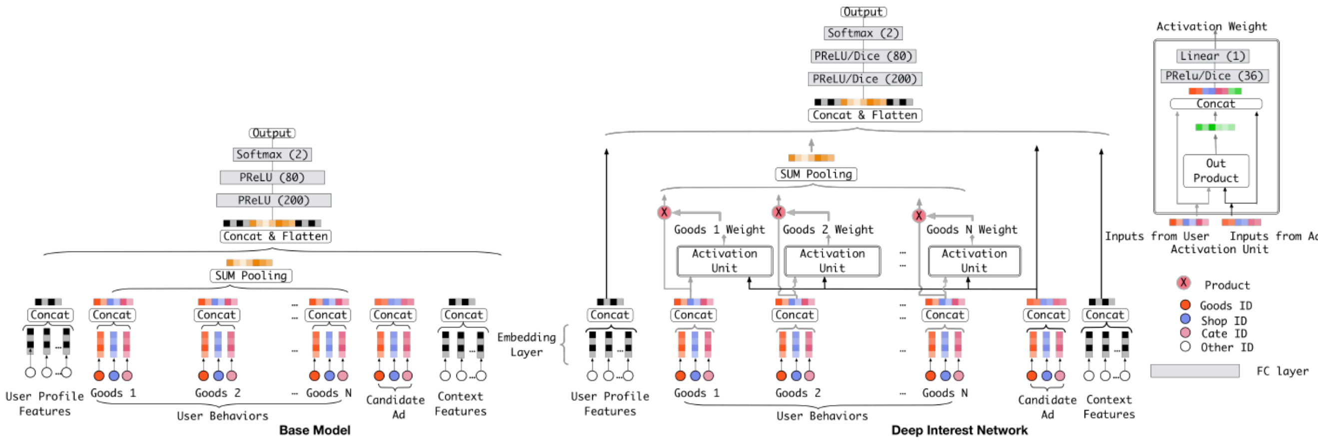
Dataset

- Amazon
  - Books
  - Electronics
- Industrial Dataset
- Metric
  - AUC
- 适用数据：
  - 用户行为数据
  - 多样特征类别
  - 目标商品点击信息

DIN

Overview

- [Deep Interest Network for Click-Through Rate Prediction]



- 固定长度表达能力不强，与candidate ad无关
- 局部激活单元（local activation unit）
    - 自适应地从用户历史行为中学习ad-specific的兴趣表示
- Two techniques:
    - 小批量正则化
        - 只计算每个mini-batch中出现的特征参数的L2范数
    - 数据自适应激活函数
        - Dice
        - PReLU的推广
        - 根据输入数分布自适应调整整流点（均值）

Dataset

- Amazon(Electronics)
- MovieLens
- Alibaba
- Metric
    - AUC
    - RelaImpr (relative improvement)

$$RelaImpr = \left(\frac{\text{AUC(measured model)} - 0.5}{\text{AUC(base model)} - 0.5} - 1 \right) \times 100\%.$$

- 适用数据:

    - pratical: 阿里广告系统
    - 大规模工业级稀疏数据

WideDeep

Previously discussed

DSSM

- 双塔召回
- Previously discussed
- Pipeline:
    - 投影（高->低维）
    - MLP
    - 余弦相似，softmax
- Metric
    - NDCG@{1, 3, 10}

PNN

Overview

- [Product-based neural networks for user response prediction]
- 用乘积层来捕捉字段之间的交互模式
    - Inner product
    - Outer product

Dataset

- Criteo
- iPinYou
- Metric
    - AUC
    - RIG (Relative Information Gain)
    - Log Loss
    - RMSE
- 适用数据：多字段，categorical，高维one-hot特征

FNN

Overview

- Deep Learning over Multi-field Categorical Data
- 最简单的raw MLP
- FM的objective

Dataset

- iPinYou
- Metric
  - AUC
- 适用数据：广告

以下为FM系的模型：

FFM

Overview

- Field-aware Factorization Machines for CTR Prediction

$$\phi_{\mathrm{FFM}}(\boldsymbol{w}, \boldsymbol{x}) = \sum_{j_1=1}^{n} \sum_{j_2=j_1+1}^{n} (\boldsymbol{w}_{j_1, f_2} \cdot \boldsymbol{w}_{j_2, f_1}) x_{j_1} x_{j_2},$$

- 在传统FM基础上，不是直接学习feature间pairwise的权重，而是学字段间权重

Dataset

- Criteo
- Avazu
- Metric
  - logloss
- 适用数据：
  - 只适合分类特征，并可转换为binary特征的数据

FwFM

Overview

- Field-weighted Factorization Machines for Click-Through Rate Prediction in Display Advertising

$$\Phi_{FwFMs}((\boldsymbol{w}, \boldsymbol{v}), \boldsymbol{x}) = w_0 + \sum_{i=1}^{m} x_i w_i + \sum_{i=1}^{m} \sum_{j=i+1}^{m} x_i x_j \langle \boldsymbol{v}_i, \boldsymbol{v}_j \rangle r_{F(i), F(j)}$$

- 就是在FFM基础上多加了一个field之间的weight，"field-weighted"

Dataset

- Criteo
- Oath
  - 两周广告点击日志
- Metric
  - AUC
- 适用数据：多字段分类数据

AFM

Overview

- Attentional Factorization Machines: Learning the Weight of Feature Interactions via Attention Networks
- 用attention学习各feature interaction的重要性

Dataset

- Frappe
- MovieLens
- Metric
  - RMSE

DeepFM

Overview

- DeepFM: A Factorization-Machine based Neural Network for CTR Prediction
- DNN + FM（简单加和再过sigmoid）

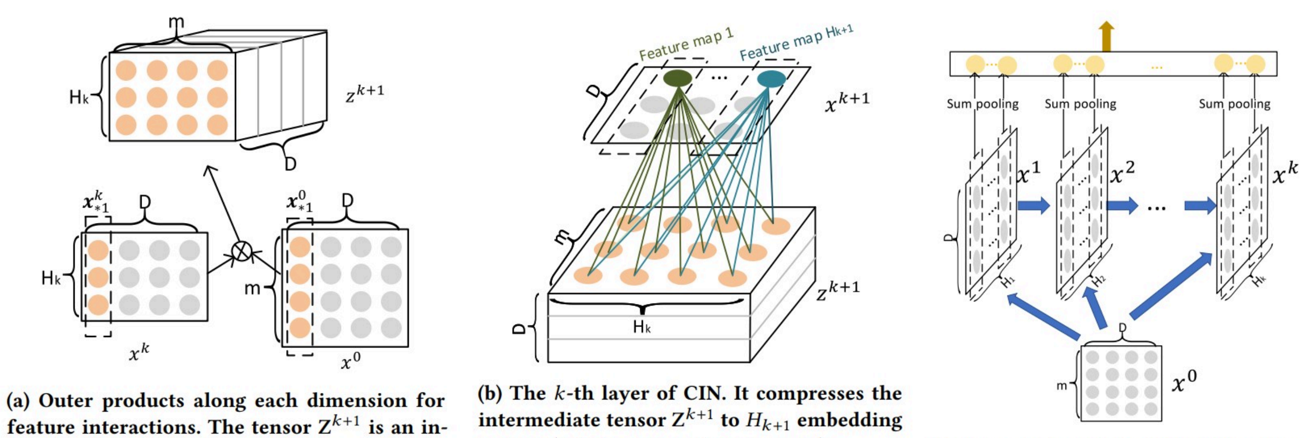
Dataset

- Criteo
- Company*
  - App Store游戏中心，连续7天的用户点击记录
- Metric
  - AUC
  - Logloss
- 适用数据：高维稀疏，CTR

xDeepFM

Overview

- eXtreme Deep Factorization Machine, Combining Explicit and Implicit Feature Interactions for Recommender Systems
- 显式生成特征交互
- CIN:
  - 每一层隐藏层的输出依赖上层和额外输入，类似RNN
  - 滤波器沿embedding维度滑动，生成特征图，类似CNN



(a) Outer products along each dimension for feature interactions. The tensor Z^{k+1} is an intermediate result for further learning.

(b) The k-th layer of CIN. It compresses the intermediate tensor Z^{k+1} to H_{k+1} embedding vectors (aslo known as *feature maps*).

(c) An overview of the CIN architecture.

- CIN + Deep: xDeepFM

    - 想法和DeepFM，Wide&Deep非常类似

Dataset

- Criteo
- Dianping
- Bing News
- Metric
    - AUC
    - Logloss
- 适用数据：
    - numerical & categorical混合
    - 缺乏明确时空上的相关性

NFM

Overview

- Neural Factorization Machines for Sparse Predictive Analytics
- 显式捕捉bi-interaction（二阶交互）->再过多层FFN

Dataset

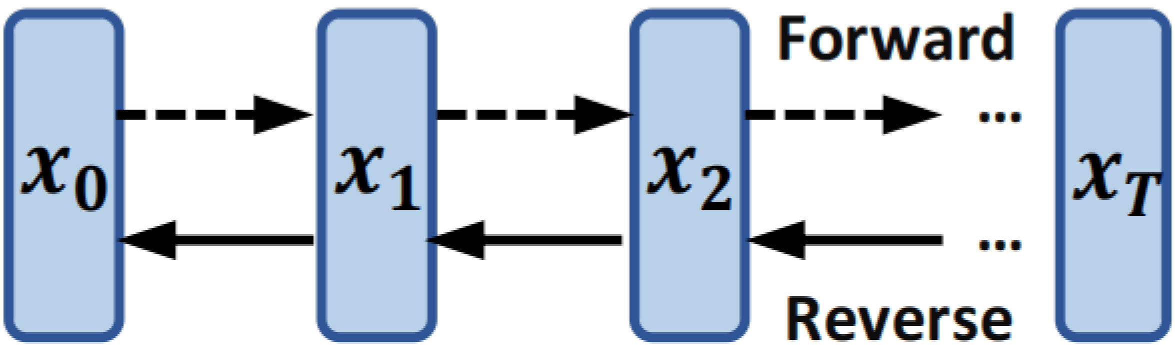
- Frappe
    - app用户日志
- MovieLens
- Metric
    - RMSE
- 适用数据：稀疏特征，categorical

Genearal Recommendation

DiffRec

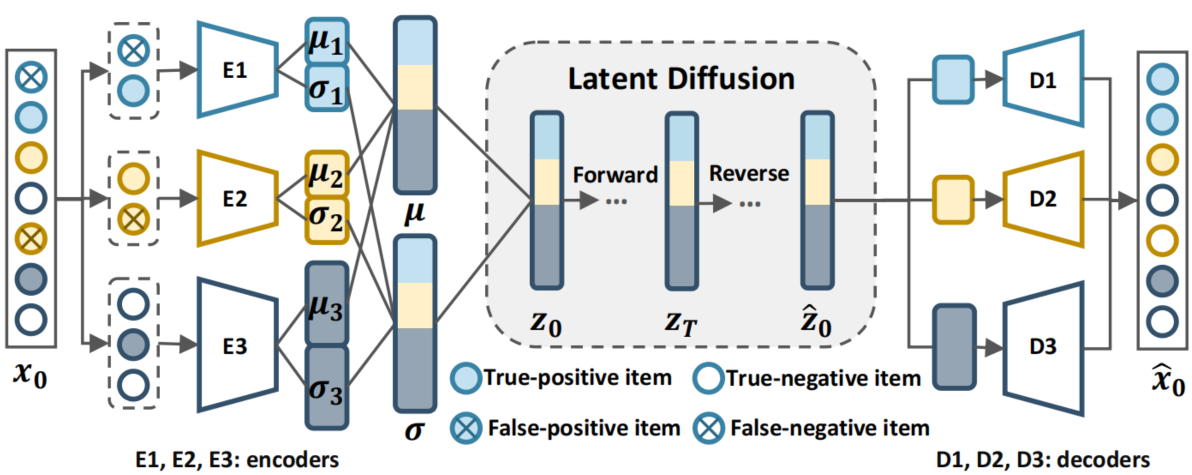
Overview

- Diffusion Recommender Model



(b) Illustration of DiffRec.

- 逐步用高斯噪声腐蚀用户交互历史，然后迭代恢复原始交互
  - 前向过程显著降低噪声尺度，保留用户个性化信息
- Inference: 用重构后的$\hat{x}_0$分布计算probability，推荐排名靠前的商品
- 两个扩展：
  - L-DiffRec
    - 用k-means对物品进行聚类，压缩维度



  - T-DiffRec
    - time-aware reweighting

        - 时序靠后的交互赋予更大权重
- 有点类似data augmentation?

Dataset

- Amazon-book
- Yelp
- ML-1M
- Metric
    - Recall@k
    - NDCG@k
    - k=10, 20
- 适用数据:
    - 高噪声用户交互
    - large-scale
    - 用户偏好随时间变化

NCL

Overview

- Improving Graph Collaborative Filtering with Neighborhood-enriched Contrastive Learning
- A+B: Graph CF + CL
    - 将结构和语义邻居纳入图协同过滤，做对比学习
- GNN:
    - 归一化聚合，BPR loss（pairwise）
- 和邻居做CL
    - Structural：所有邻居平等对待
    - Semantic：仅考虑聚类之后在同prototype之下的邻居

Dataset

- MovieLens1M
- Yelp
- Amazon Books
- Gowalla
- Alibaba-iFashion
- Metric

- Recall@N
- NDCG@N
- N = 10, 20, 50
- 适用数据：隐式反馈，sparse，nothing specific

SimpleX

Overview

- A Simple and Strong Baseline for Collaborative Filtering
- 关注点：loss function & 负采样比例
  - **简单**鲁棒的baseline
- CCL
  - cosine contrastive loss

$$\mathcal{L}_{CCL}(u,i) = (1 - \hat{y}_{ui}) + \frac{w}{|\mathcal{N}|} \sum_{j \in \mathcal{N}} \max(0, \ \hat{y}_{uj} - m)$$

  - m: margin; 正样本 + neg sample
- 基本pipeline：
  - 所有交互的items做aggregation
    - avg
    - attn
  - 聚合的item embed和user embed加权加和
  - 和candidate item余弦相似，最后计算出CCL

Dataset

- Amazon-Books
- Yelp18
- Gowalla
- further test:
  - Amazon-CDs, Amazon-Movies, Amazon-Beauty
  - CiteUlikeA, Movielens-1M
- Metric
  - Recall@20
  - NDCG@20
  - wide range of comparison
- 适用数据：

- large-scale
- 计算资源有限场景

NCEPLRec

Overview

- Noise Contrastive Estimation for One-Class Collaborative Filtering
- 解决问题：总是推荐popular items，缺乏个性化推荐
- NCE: Noise Contrastive Estimation
  - 平衡observed & unobserved
- NS-AutoRec
  - 通过负采样训练嵌入层
  - 仅考虑重构error，NS作正则项
- NCE-AutoRec
  - unobserved interaction的预测期望用负采样期望近似

$$\underset{\tilde{\mathbf{r}}_i}{\operatorname{argmax}} \sum_j r_{i,j} \left[\log p(\tilde{r}_{i,j} = 1) + E_{p(j')}[\log p(\tilde{r}_{i,j'} = 0)] \right],$$

Dataset

- Goodbooks
- MovieLens-20M
- Netflix
- Yahoo
- Metric
  - R-Precision, NDCG, MAP@K, Precison@K, Recall@K and F1-score@K
- 适用数据：
  - large-scale, real-life
  - implicit feedback
  - 用户偏好多样化

ADMMSLIM

Overview

- Sparse Recommendations for Many Users
- by Netflix
- 优化原始Slim目标函数，训练时间与用户数无关，可扩展到大规模用户群体

$$
\begin{aligned}
\min_{B} \quad & \frac{1}{2} \cdot \|X - XB\|_F^2 + \frac{\lambda_2}{2} \cdot \|B\|_F^2 + \lambda_1 \cdot \|B\|_1 \\
\text{s.t.} \quad & \operatorname{diag}(B) = 0 \\
& B_{i,j} \geq 0 \quad \forall i, j \in \mathcal{I}
\end{aligned}
$$

- B涉及多个函数和约束，重新定义为等效优化问题，前两项$f(B)$，最后一项$g(C)$, s.t. $B = C$
- 这个constraint用拉格朗日乘数约束，然后这个优化用类似ALS的方法交替更新B和C（有闭式解）

$$
L_\rho(B, C, \Gamma) = f(B) + g(C) + \langle \Gamma, B - C \rangle_F + \frac{\rho}{2} \cdot \|B - C\|_F^2
$$

- 可调整各约束和正则项修改目标函数，灵活，提供ablation
- 优点：转化为优化问题，收敛快，scalability

Dataset

- ML-20M
- Netflix Prize
- Million Song Data(MSD)
- Metric
  - Recall@20
  - Recall@50
  - NDCG@100
- 适用数据：
  - 大量用户
  - 冷启动
  - catalog size > # user

SGL

Overview

- Self-supervised Graph Learning for Recommendation
- 解决问题:
  - 监督信号稀疏
  - 数据分布偏向高度节点
  - prone to noise
- 数据增强方法:
  - 节点dropout
  - 边dropout
  - random walk
- CL: 最大化同一节点不同视图间的一致性
- MTL: 总loss = cross-entropy + CL loss + regularizatoin

Dataset

- Yelp2018
- Amazon-Book
- Alibaba-iFashion
- Metric
  - Recall@20
  - NDCG@20
- 适用数据:
  - 数据分布biased
  - observed interaction高噪

SLIMElastic

Overview

- Sparse Linear Methods for Top-N Recommender Systems
- Refer to ADMMSLIM

Dataset

- 购买交易记录

    - ccard
    - ctlg2
    - ctlg3
    - ecmrc
- 评分数据
    - BX
    - ML10M
    - Netflix
    - Yahoo
- Metric
    - HR
    - ARHR
        - 对每个用户命中以位置倒数加权：$\frac{1}{\#users}\sum_{i\in hit}\frac{1}{p_i}$
- 适用数据：purchase, rating

EASE

Overview

- Embarrassingly Shallow Autoencoders for Sparse Data
- 在SLIM基础上，drop掉非负约束和L1正则
- 同样提供闭式解（形式简单），方法和ADMMSLIM略有不同

$$\hat{B}_{i,j} = \begin{cases} 0 & \text{if } \ i = j \\ -\frac{\hat{P}_{ij}}{\hat{P}_{jj}} & \text{otherwise.} \end{cases}$$

$$\hat{P} \triangleq (X^\top X + \lambda I)^{-1}$$

Dataset

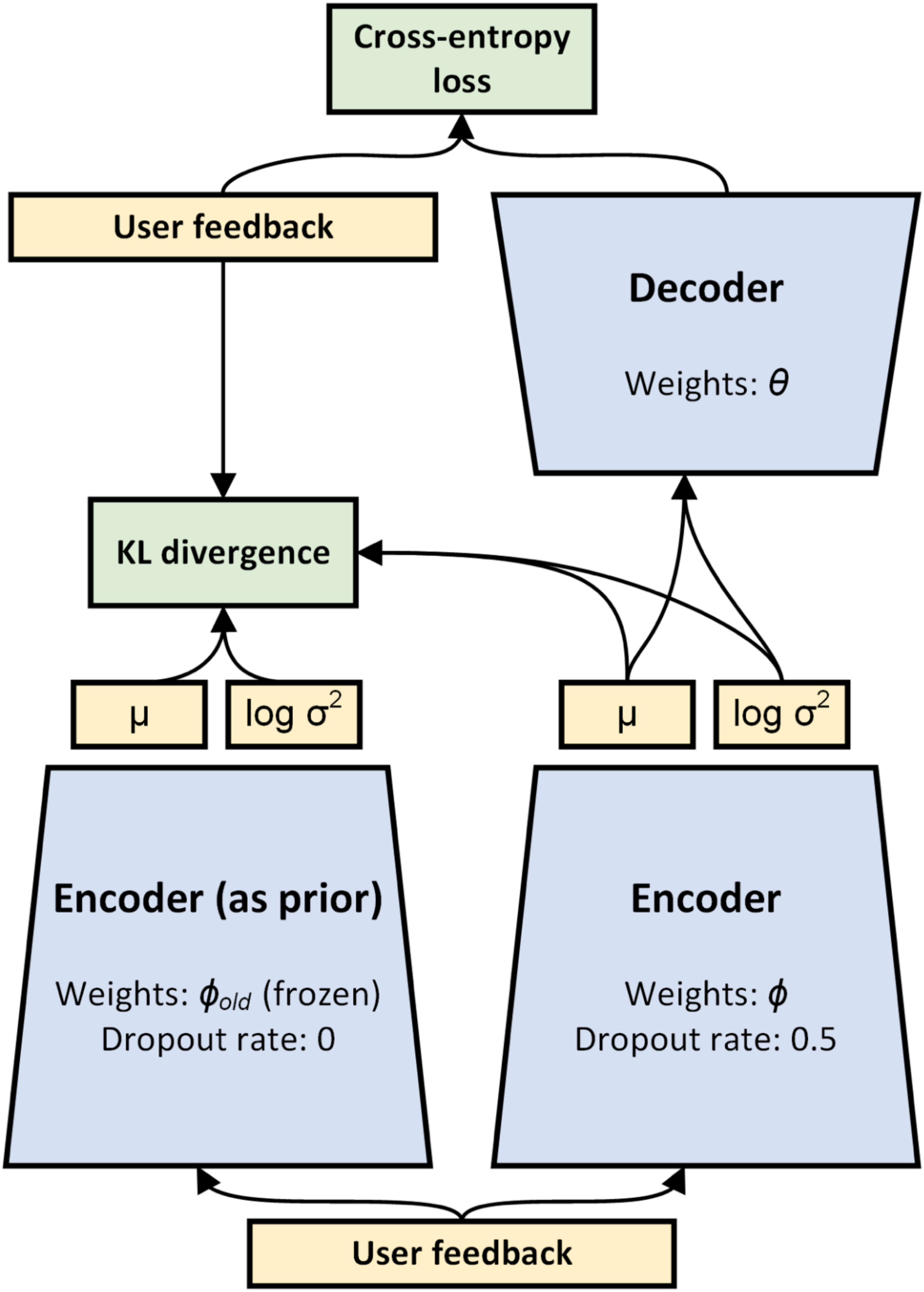
- ML-20M
- Netfix Prize
- Million Song Data (MSD)
- Metric
    - Recall@{20, 50}
    - NDCG@100

- 适用数据：
  - 个性化相关性高，更好推荐长尾item
  - sparse, implicit, etc.

RecVAE

Overview

- A New Variational Autoencoder for Top-N Recommendations with Implicit Feedback
- Auto Encoder Paradigm


Cross-entropy loss
User feedback
Decoder
Weights: θ
KL divergence
μ
log σ^2
μ
log σ^2
Encoder (as prior)
Weights: ϕ_{old} (frozen)
Dropout rate: 0
Encoder
Weights: ϕ
Dropout rate: 0.5
User feedback

- 把传统VAE的高斯分布改为多项式分布

$$\mathcal{L} = \mathbb{E}_{q_\phi(z|\tilde{x})}\mathbb{E}_{q(\tilde{x}|x)}\Big[\log p_\theta(x|z) - \beta'(x)\mathrm{KL}\left(q_\phi(z|\tilde{x})\|p(z|\phi_{old}, x)\right)\Big],$$

- 把原始的ELBO改为去噪VAE的形式（体现在期望上）
- 复合先验：标准高斯与上一epoch后验近似（$q_\phi(z \mid x)$）的weighted sum

$$p(z|\phi_{old}, x) = \alpha\mathcal{N}(z|0, \mathrm{I}) + (1-\alpha)q_{\phi_{old}}(z|x),$$

- 超参β不再是β-VAE中的常数，而是与当前用户interaction数量成正比（体现在$\beta'(x)$）
- 类似ALS，user和item embedding交替进行

Dataset

- MovieLens-20M
- Netflix Prize
- Million Songs
- Metric
  - Recall@{20, 50}
  - NDCG@100
- 适用数据：
  - 电商，内容推荐等隐式反馈场景
  - 处理稀疏，噪声

RaCT

Overview

- Towards Amortized Ranking-Critical Training for Collaborative Filtering
- 借鉴RL中Actor-Critic（玩家-评委）的想法
  - Critic：近似排名指标
  - Actor：针对指标优化
- 训练方法：
  - 用MLE预训练actor网络，standard in VAE
    - $\mathbb{E}_{q_\phi(z|x)} \log p_\theta(x \mid z)$
  - 预训练critic网络，最小化MSE（w/ groud truth）

- 类似GAN，actor as G, critic as D

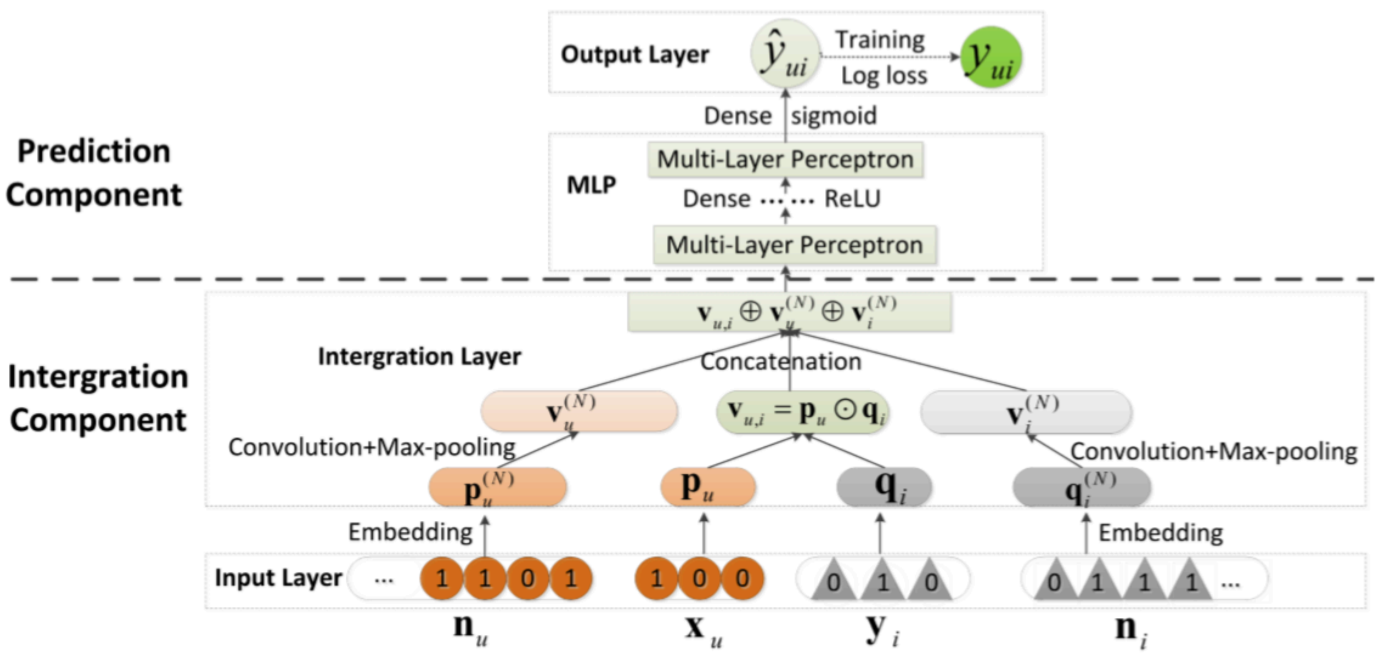
Dataset

- ML-20M
- Netflix
- MSD
- Metric
  - R@20
  - R@50
  - NDCG@100
- 适用数据：implicit, large-scale, sparse

NNCF

Overview

- A Neural Collaborative Filtering Model with Interaction-based Neighborhood



- 增强局部信息利用
- 编码领域信息：
  - 在邻居隐向量上卷积再Max-pooling
- 全局和局部隐向量concat在一起，再过MLP
- 想法非常trivial

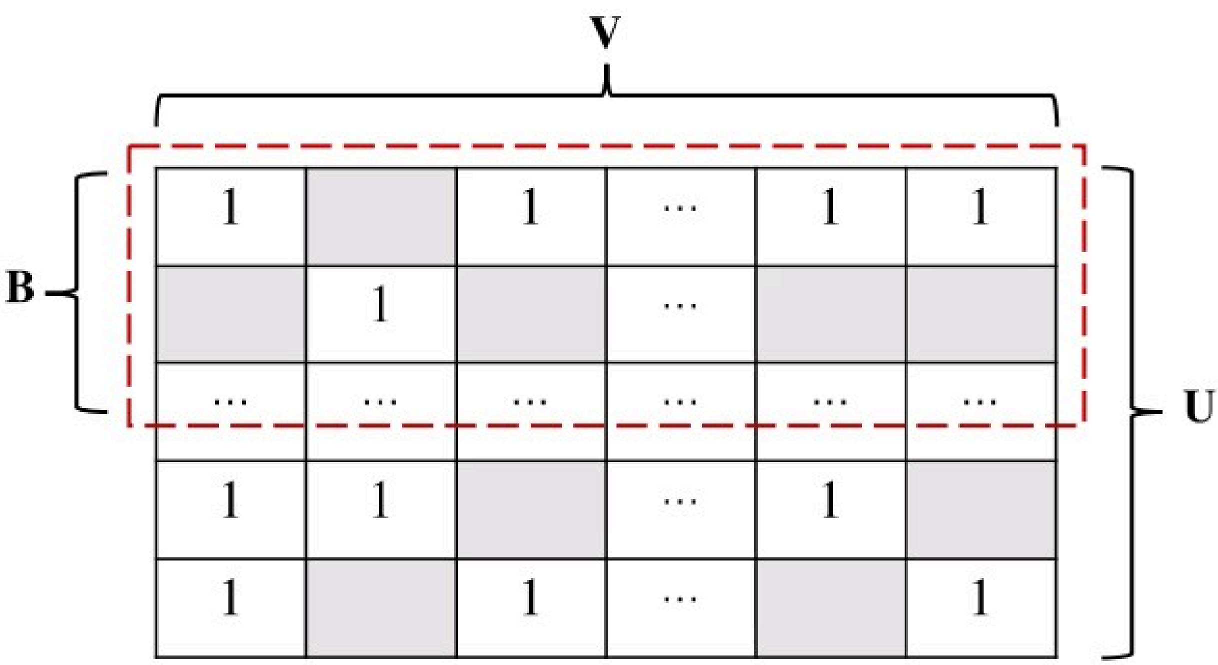
Dataset

- Delicious
    - 社交书签网络服务
- MovieLens
- Rossmann
    - 药店销售记录，kaggle competition
- 适用数据：implicit, sparse, same with traditional CF

ENMF

Overview

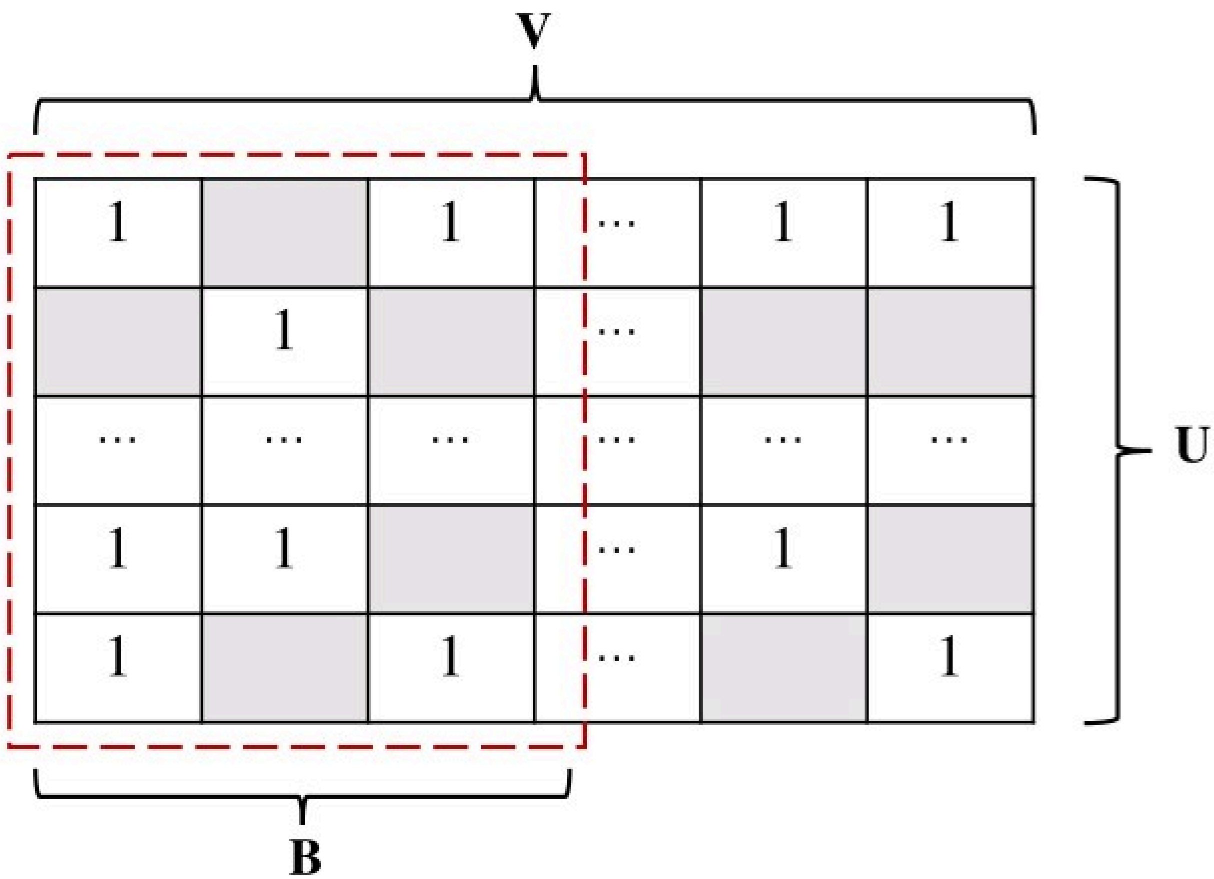
- Efficient Neural Matrix Factorization without Sampling for Recommendation
- 负采样不够鲁棒，不易优化
- 从整个训练数据学习神经模型，w/o sampling
- 提出三个优化techniques：
    - User-based optimization method (ENMF-U)
        - weighted regression loss
        - user batch（横向）



        - 把loss分为正数据损失$\mathcal{L}_1^P(\Theta)$ & 所有数据损失$\mathcal{L}_1^A(\Theta)$
        - 优化：nested sum的重排

$$\mathcal{L}_1^A(\Theta) = \sum_{i=1}^{d}\sum_{j=1}^{d}\left(\left(h_i h_j\right)\left(\sum_{u\in\mathbf{B}} p_{u,i}p_{u,j}\right)\left(\sum_{v\in\mathbf{V}} c_v^- q_{v,i}q_{v,j}\right)\right)$$

- Item-based optimization method (ENMF-I)
  - item batch（纵向）



- Alternating-based optimization method (ENMF-A)
  - motivated by ALS，交替优化user和item

Dataset

- Ciao
  - 购物评分
- Epinions
- Movielens
- Metric
  - HR
  - NDCG
- 适用数据：large-scale, implicit

<u>以下为VAE/DAE系的模型：</u>

CDAE

Overview

- Collaborative Denoising Auto-Encoders for Top-N Recommender Systems

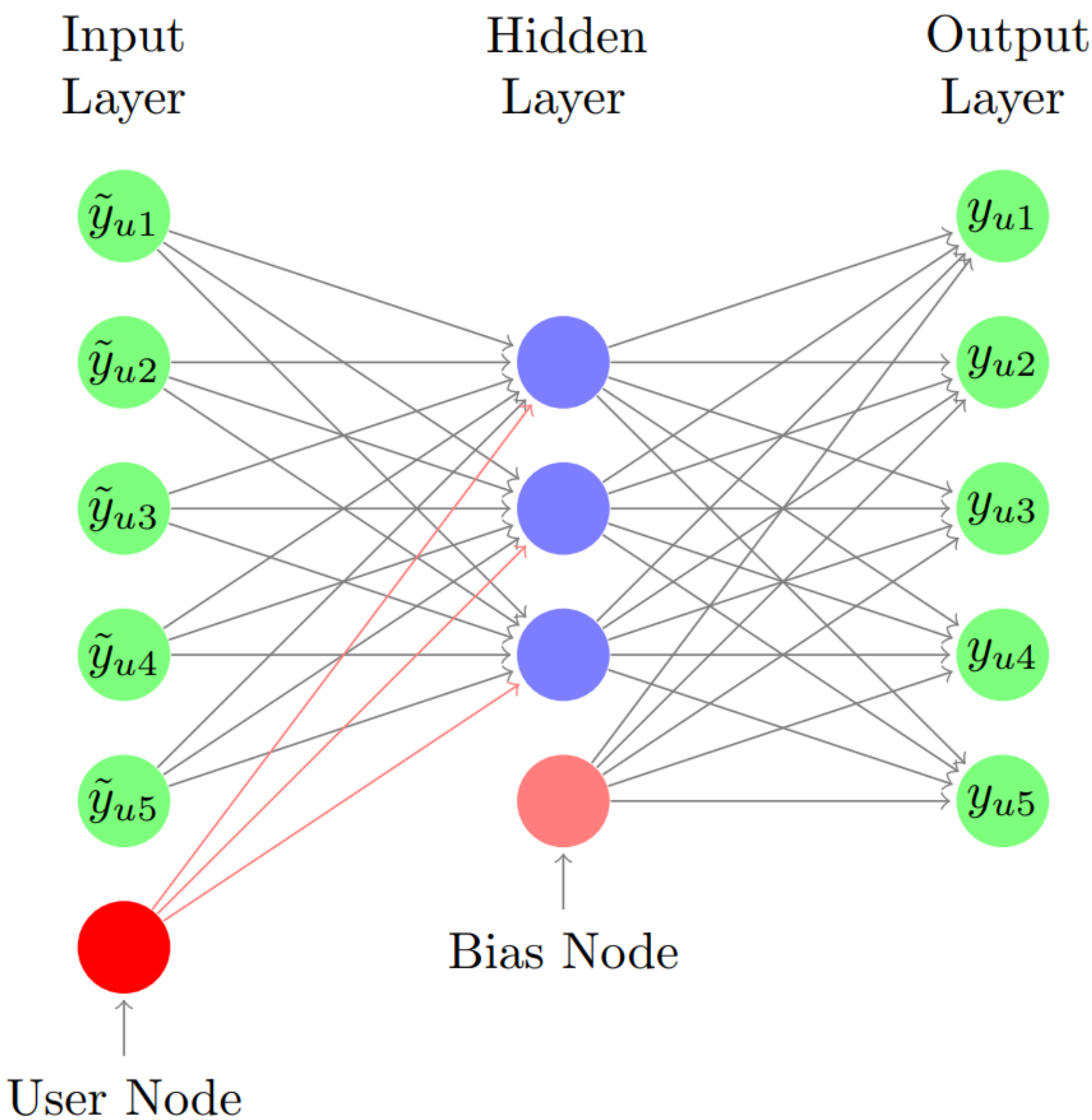


Figure 1: A sample CDAE illustration for a user u. The links between nodes are associated with different weights. The links with *red* color are user specific. Other weights are shared across all the users.

- 自编码器框架，从被污染的输入学习
    - 单隐藏层
    - input -> latent -> reconstruction

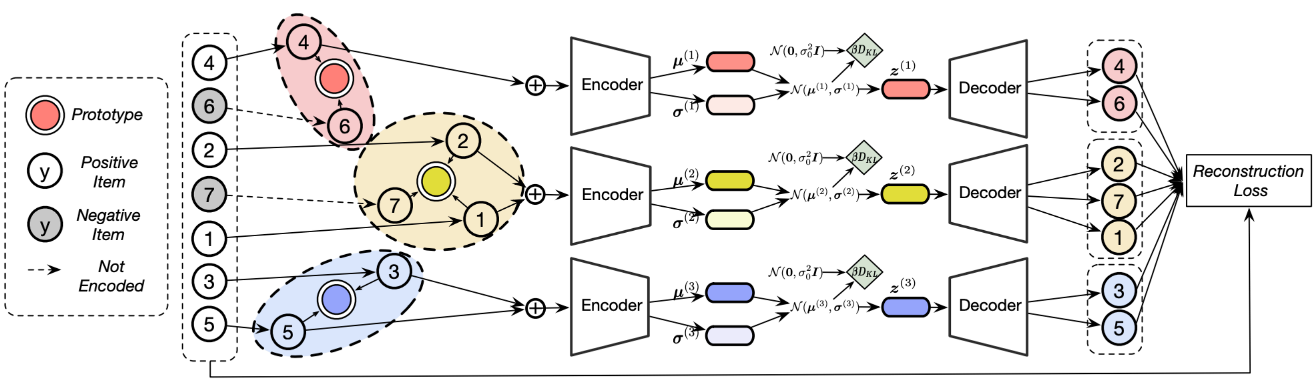
Dataset

- MovieLens-10M
- Netflix
- Yelp
- Metric
  - MAP@{1, 5, 10}
- 适用数据：implicit, also corrupted data (nature of **denoising**)

MacridVAE

Overview

- Learning Disentangled Representations for Recommendation



- 宏观分离（macro disentanglement） & 微观分离（micro disentanglement）
  - Macro: 捕获对高层次概念的偏好（e.g. 类别）
    - latent $\mathbf{z}_u \in \mathbb{R}^{Kd}$, K:高层次概念数量
    - 同时推理独热向量$\mathbf{C}$，聚类item
    - 用VAE paradigm做推理
  - Micro: 独立的低层次因素（e.g. 大小，颜色）
    - 鼓励维度间的独立性
    - 做法：ELBO中KL散度写成互信息形式去噪，并采用β-VAE to force independence (set β high)

Dataset

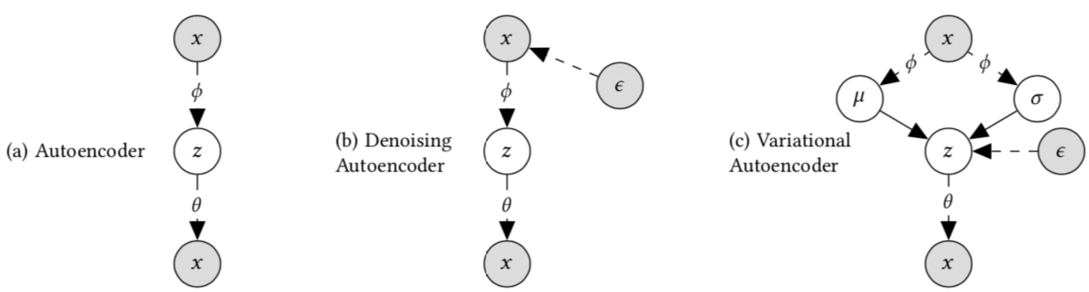
- Netflix Prize
- Movielens
  - ML-100k

  - ML-1M
  - ML-20M
- AliShop-7C
  - 7C: 7 categories
- Metric
  - NDCG@100
  - Recall@20
  - Recall@50
- 适用数据：user behavior, u-i interaction

MultiVAE & MultiDAE

Overview

- Variational Autoencoders for Collaborative Filtering
- VAE4Rec的鼻祖
- 用多项式似然$\mathbf{x}_u \mid \mathbf{z}_u \sim Mult(N_u, MLP(\mathbf{z}_u))$
- Taxonomy：
  - DAE: delta变分分布（仅在$g_\phi(\mathbf{x}_u)$处有概率密度）
  - VAE: 参数化高斯近似



Dataset

- ML-20M
- Netflix
- MSD
- Metric
  - Recall@20
  - Recall@50
  - NDCG@100

- 适用数据：implicit, large, sparse, high-dim

LINE

Overview

- Large-scale Information Network Embedding
- designed to：大型信息网络嵌入低维空间
- 优化目标函数同时保留全局&局部结构
    - 一阶近邻（local） & 二阶近邻（global）
        - 一阶：两节点联合概率，sigmoid（observed）
        - 二阶：建模由上下文节点生成另一节点的条件概率，softmax（unobserved）

$$p_2(v_j|v_i) = \frac{\exp(\vec{u}_j'^T \cdot \vec{u}_i)}{\sum_{k=1}^{|V|} \exp(\vec{u}_k'^T \cdot \vec{u}_i)},$$

- 负采样方法：unigram分布（$P_n(v) \propto d_v^{\frac{3}{4}}$，out degree）
- 图学习方法可以应用到推荐系统（observed/unobserved模式），把图看作interaction关系即可

Dataset

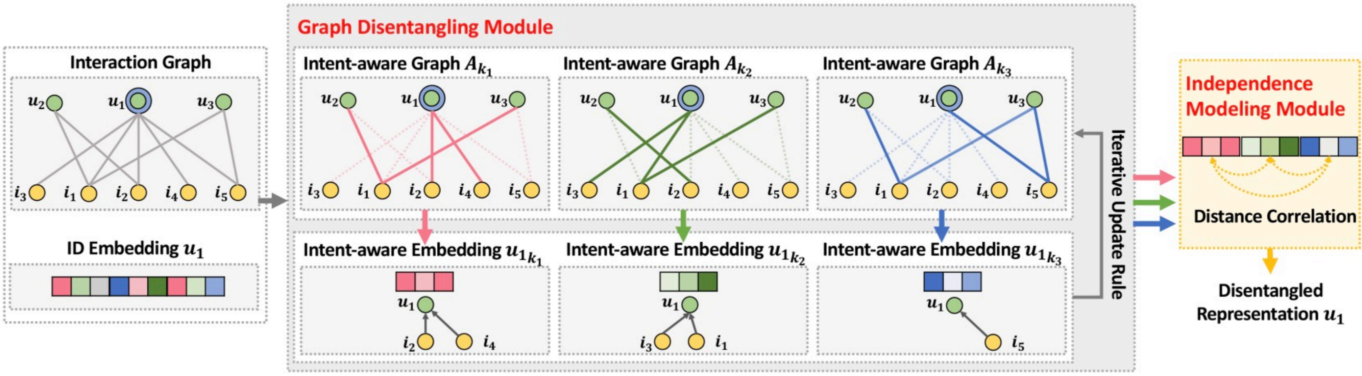
- 论文中使用数据为各种网络数据，e.g. 语言网络（维基百科），社交网络（flicker, youtube），引用网络（DBLP）
- 适用数据：interaction, implicit即可

以下为一些graph-based模型：

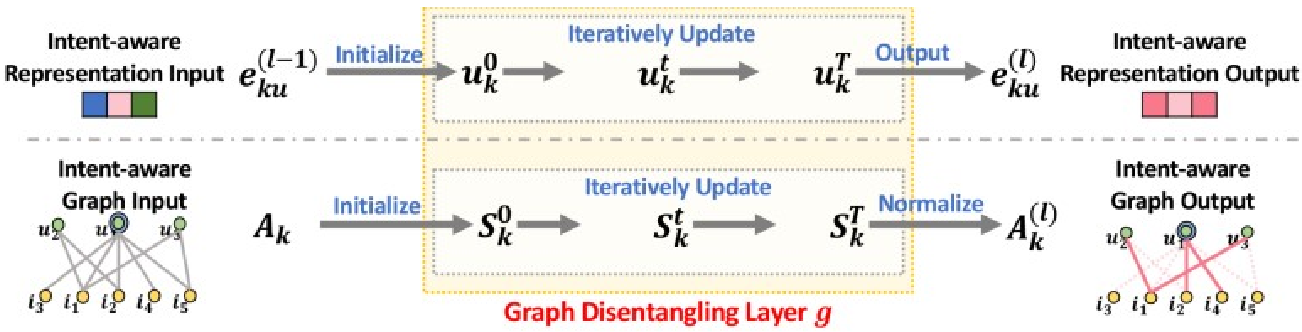
DGCF

Overview

- Disentangled Graph Collaborative Filtering



- 解决问题：uniform modeling无法有效捕捉用户意图的多样性
- Could refer to MacridVAE
- 基本Pipeline:
  - K个latent intent，每个包含一组user/item分块的表示（$\mathbf{u}_k$，$\mathbf{i}_k$）
  - 维护两组矩阵，交替更新



    - k intent下user表示
      - aggregate所有该user相邻的item，weighted by (u,i) score $S_k(u,i)$
    - k intent下的全局图表示
      - update $S_k(u,i)$，S += (u^t, i^0) attn score（i不更新，nonlinear activation）
  - 正则项：鼓励latent factor间独立性，使用distance correlation

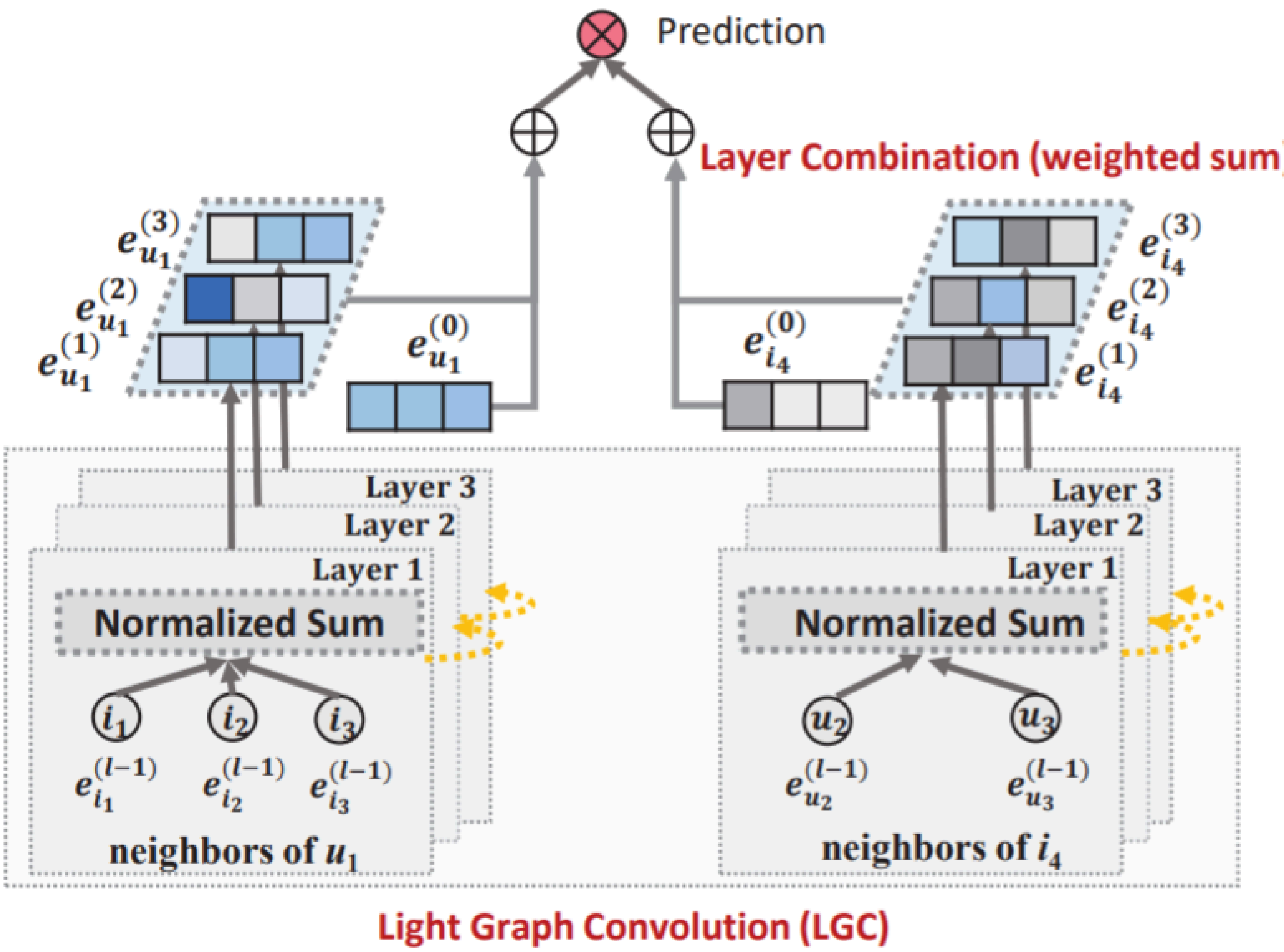
Dataset

- Gowalla
  - location-based check in
- Yelp2018$*$
- Amazon-Book
- Metric
  - recall@20
  - ndcg@20
- 适用数据：多样化用户意图，其他相似

LightGCN

Overview

- Simplifying and Powering Graph Convolution Network for Recommendation



- Light：仅包含GCN中领域聚合
    - AGG: normalized sum（u & i）
    - 各层组合：weighted sum
    - inner product做预测
- BPR loss

Dataset

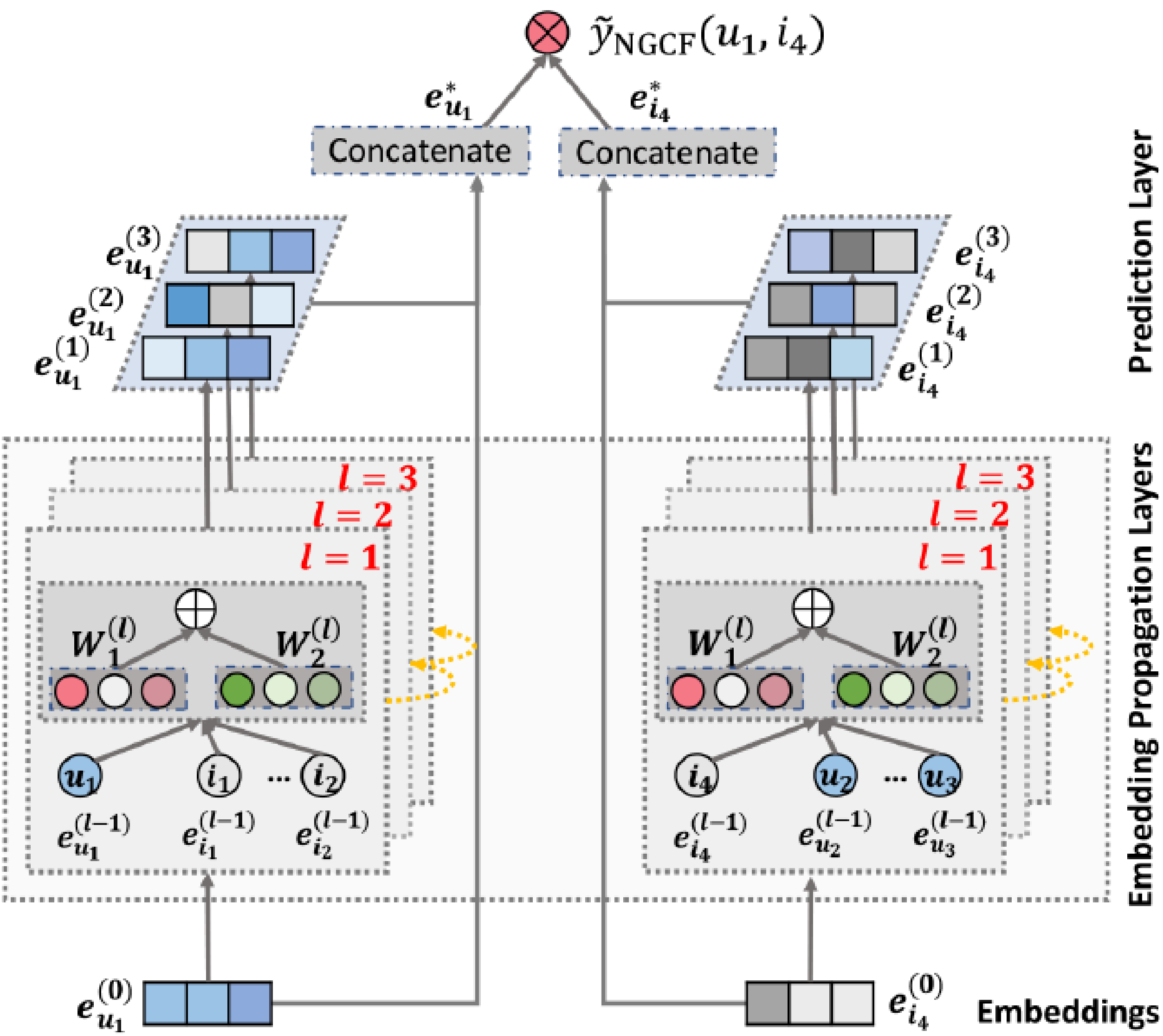
- Gowalla
- Yelp2018
- Amazon-Book
- Metric
    - recall@20

- ndcg@20
- 适用数据：轻量，长期依赖，其他同

NGCF

Overview

- Neural Graph Collaborative Filtering



- 和LightGCN相比：
  - AGG时，包括self-connection（user2user自己的信息传递）
  - activation: LeakyReLU
- LightGCN claim这两处在消融中useless

Dataset

Same as LightGCN

GCMC

Overview

- Graph Convolutional Matrix Completion
- Bipartite Graph(user -- item)
    - 转化为link prediction任务
- 仅考虑1-hop
- 其他类似
    - AGG: normalized sum
    - Decoder: softmax
    - both w/ trainable matrix added
- loss: NLL (negative log likelihood)

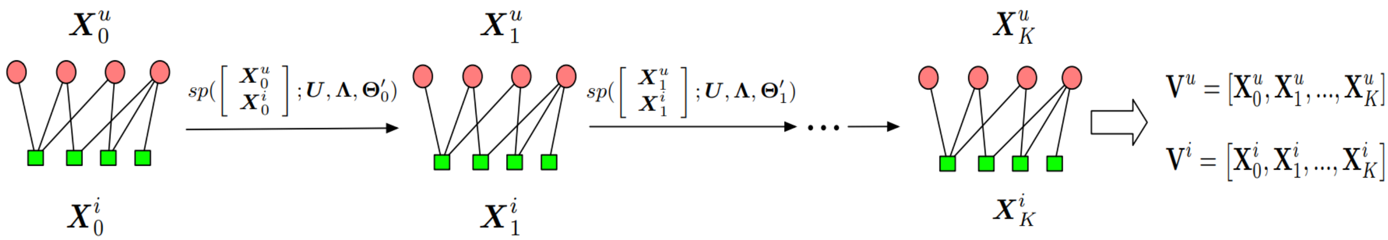
Dataset

- MovieLens(100K, 1M, 10M),
- Flixster
    - also movie rating
- Douban
- YahooMusic
- Metric
    - RMSE
- 适用数据：解决冷启动，任何可以转化为二分图

SpectralCF

Overview

- Spectral collaborative filtering



- 频谱卷积：动态放大/衰减各频率分量，捕捉连通性信息
- 图傅里叶变换
  - 拉普拉斯矩阵，eigenvector组成U
  - x（图信号）--U^T--> $\hat{x}$（频域信号）--U--> 新图信号
- 卷积核g_θ applied before U，调节频率分量
  - 优化：卷积核用多项式近似，降低复杂度
  - 节点表示拓展到C维，层数扩展到K层，各层结果concat

$$\begin{bmatrix} X_{new}^u \\ X_{new}^i \end{bmatrix} = \sigma\left((UU^\top + U\Lambda U^\top)\begin{bmatrix} X^u \\ X^i \end{bmatrix}\Theta'\right),$$

- loss: BPR

Dataset

- MovieLens-1M
- HetRec
  - extension of ML, also movie ratings, transformed into implicit
- Amazon Instant Video
- Metric
  - Recall@M
  - MAP@M
  - M={20, 40, 60, 80, 100}
- 适用数据：implicit, cold start

以下是一些item-based算法：

NAIS

Overview

- [Neural Attentive Item Similarity Model for Recommendation]
- item-to-item

- softmax分母平滑化处理

$$\hat{y}_{ui} = \mathbf{p}_i^T \left(\sum_{j \in \mathcal{R}_u^+ \setminus \{i\}} a_{ij} \mathbf{q}_j \right),$$

$$a_{ij} = \frac{\exp(f(\mathbf{p}_i, \mathbf{q}_j))}{\left[\sum_{j \in \mathcal{R}_u^+ \setminus \{i\}} \exp(f(\mathbf{p}_i, \mathbf{q}_j))\right]^\beta},$$

- loss: cross-entropy

Dataset

- MovieLens-1M
- Pinterest
- Metric
  - HR@10
  - NDCG@10
- 适用数据：item-based CF，不同长度的用户历史

FISM

Overview

- Factored Item Similarity Models for Top-N Recommender Systems
- estimated score $\hat{r}_{ui}$: same as NAIS
- FISMrmse
  - 目标函数：squared error
- FISMauc：BPR loss

Dataset

- ML100K
- Netflix
- Yahoo Music
- Metric
  - HR

- ARHR
- 适用数据：sparse, implicit, scalable

DMF

Overview

- MLP分别编码user/item
- loss: normalized cross-entropy
    - divided by max(rating)，同时考虑explicit / implicit

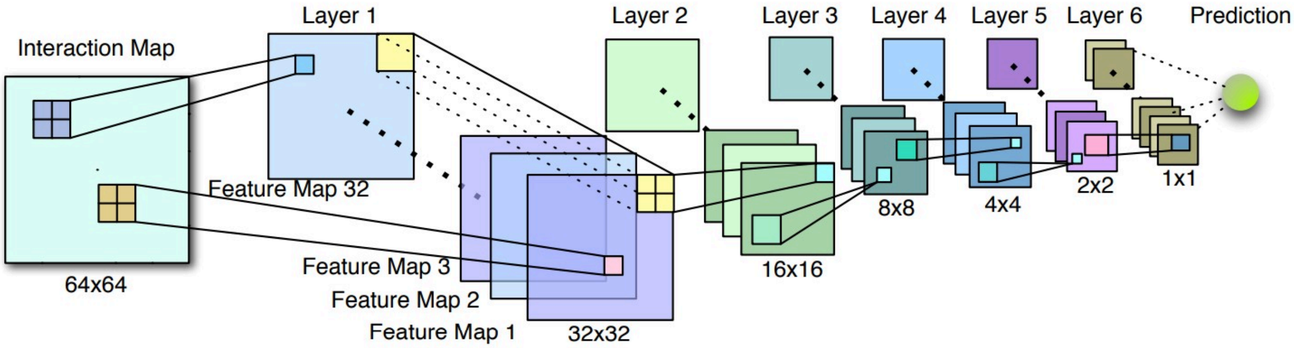
Dataset

- MovieLens 100K, 1M
- Amazon Music, Movie
- Metric
    - NDCG@10
    - HR@10
- 适用数据：explicit & implicit unified

ConvNCF

Overview

- Outer Product-based Neural Collaborative Filtering
- one-hot feature -> embedding (w/ embed_size = K)
- 使用外积构造interaction map
    - $\mathbf{E} = \mathbf{p}_u \otimes \mathbf{q}_i = \mathbf{p}_u \mathbf{q}_i^T \in \mathbb{R}^{K \times K}$
    - 编码维度间二阶correlation
- CNN抓取interaction map中的重要信号
    - 2×2局部信息，输出为所有维度correlation



- loss: BPR
- prediction: $\hat{y}_{ui}$为最后一层输出reweighted

Dataset

- Yelp
  - Business rating
- Gowalla
  - check-in dataset
- Metric
  - NDCG@10
  - HR@10
- 适用数据：sparse, implicit

NeuMF

Overview

- Neural Collaborative Filtering
- GMF: $\phi = \mathbf{p}_u \odot \mathbf{q}_i$ (linearity)
- MLP: non-linearity
- 两者concat后过激活函数
- loss: cross-entropy

Dataset

- MovieLens 1M
- Pinterest
- Metric
  - NDCG@10
  - HR@10

- 适用数据：same as ConvNCF

BPR

Overview

- BPR-Opt通用优化架构
  - 可适用于：MF, adaptive kNN
  - MF: $\hat{X} = WH^T, \Theta = W, H$
  - kNN: $\hat{x}_{ui} = \sum_{k \in I_u^+ \setminus i} s_{ik}, \Theta = S$为cosine相似度矩阵
- General objective: $\sum \log \sigma(\hat{x}_{uij}) - \lambda \|\Theta\|^2$
  - $i \in I_u^+, j \in I_u^-$
  - $\hat{x}_{uij} = \hat{x}_{ui} - \hat{x}_{uj}$, pairwise
- 训练：随机选择三元组SGD

Dataset

- Rossmann
- Netflix
- Metric
  - AUC
- 适用数据：any implicit

ItemKNN

Overview

- Item-based top-N recommendation algorithms
- Similarity:
  - cosine
  - 条件概率：$\frac{\sum_{R_{q,j}>0} R_{q,j}}{Freq(i)Freq(j)^\alpha}$，每行normalized
  - 可推广到set of items

Dataset

- ctlg1,2,3
  - 目录零售采购记录

- ecmrc
    - 电商网站
- ccard
    - 商场信用卡消费记录
- Movies
    - em
        - EachMovie
    - ml
- skill
    - 简历中IT技术
- Metric
    - HR@10
    - ARHR@10
- 适用数据：implicit