

新闻推荐

News Recommendation

张配天-2018202180

目标

给定用户u，新闻池N，从中选取前k个u最有可能点击的新闻，推荐给用户。

- 最大化click through rate ($CTR = \frac{clicked}{viewed}$)

特点

- 1. Large scale:** 用户和新闻数量都很大，新闻更新快
- 2. Sparse:** 很多用户看很少的新闻，使其特征稀疏
- 3. Dynamic:** 用户的兴趣一直发生着变化
- 4. No query:** 人们来看推荐很少会抱有特定的信息诉求，而是
show me something interesting
- 5. No explicit feedback:** 难以得到用户的显示反馈，只有点击数据
- 6. No heterogeneous context:** MIND数据集不提供用户的地理位置、
搜索记录等信息

协同过滤

在user-item矩阵基础上进行计算，忽视item的内容

按照方法，分为：

1. memory-based:

- 根据用户历史记录，计算用户之间两两之间相似度(Jaccard)，以此作为权重，对u的推荐结果就是其和剩余n-1个用户的历史记录加权平均得到的向量中分数大于某一阈值的对应item
- 要算user两两之间的相似度，计算量太大，使用不同方法减少运算量：MinHash+LSH，除了这样的trick外，提出**model-based**

2. model-based:

- 将user和item映射到同一个隐空间中

- **neighbor:**

- 使用kmeans、hierarchical clustering等方法将用户进行聚类，将推荐局限在cluster中

- **Matrix Factorization:**

- 分解user-item矩阵，将用户和item都映射到同一个隐空间中，用户有其特征向量 p_u ，新闻有其特征向量 q_v ，所有用户对所有item的平均值是 μ ， μ 的偏置是 b_u ， v 的偏置是 b_v ，则用户 u 对于新闻 v 的评分为：

$$r = b_{uv} + p_u^T \cdot q_v$$
$$b_{uv} = \mu + b_u + b_v$$

- **基于概率的分解:**

- PLSI，将user和item整合到多个高斯分布上

- **按照对象，分为：**

- **user-oriented:**

- 计算用户之间的相似度，进行上述的聚类等
- 给u推荐与其相似的用户爱看的新闻

- **item-oriented:**

- 计算item之间的相似度，进行上述的聚类
- 给u推荐与其看过的新闻相似的新闻

矩阵分解等办法一旦加入新user、item就得重新训练，而且稀疏情况效率极差，因此提出了
Factorization Machine;

没办法推荐新内容（没人点击），冷启动问题严重，不适用于新闻推荐的场景，因此提出了
基于内容的推荐;

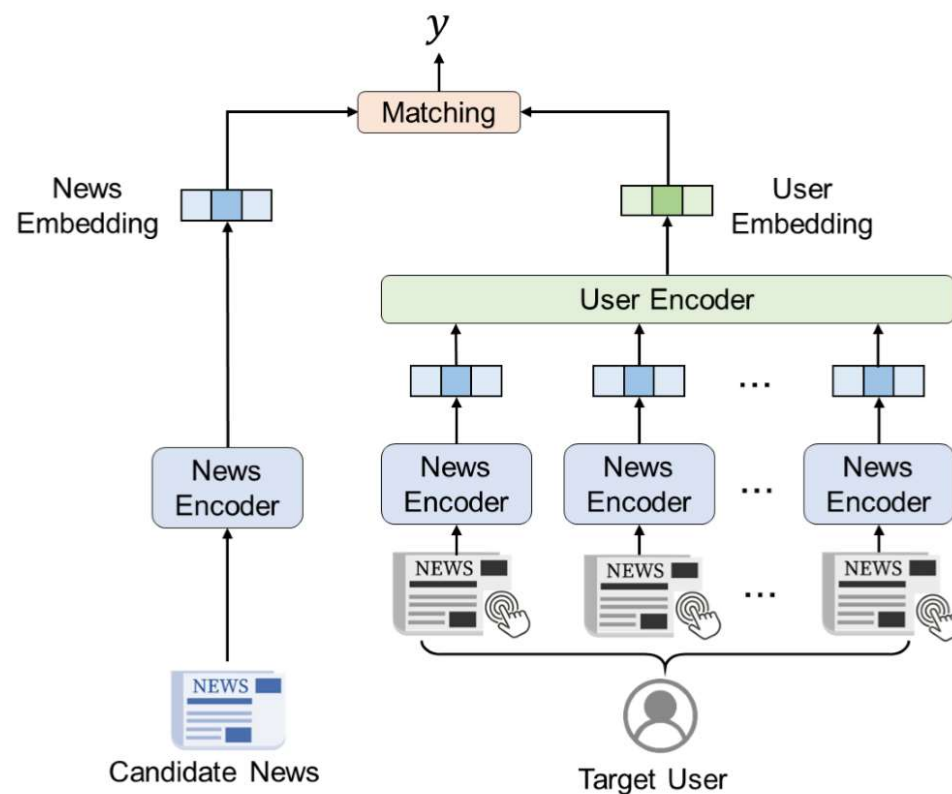
Factorization Machine

将user-item矩阵拉平，并混合一些内容相关成分，形成transaction-attribute矩阵，在此基础上进行计算

Feature vector x																	Target y					
$x^{(1)}$	1	0	0	...	1	0	0	0	...	0.3	0.3	0.3	0	...	13	0	0	0	0	...	5	$y^{(1)}$
$x^{(2)}$	1	0	0	...	0	1	0	0	...	0.3	0.3	0.3	0	...	14	1	0	0	0	...	3	$y^{(2)}$
$x^{(3)}$	1	0	0	...	0	0	1	0	...	0.3	0.3	0.3	0	...	16	0	1	0	0	...	1	$y^{(2)}$
$x^{(4)}$	0	1	0	...	0	0	1	0	...	0	0	0.5	0.5	...	5	0	0	0	0	...	4	$y^{(3)}$
$x^{(5)}$	0	1	0	...	0	0	0	1	...	0	0	0.5	0.5	...	8	0	0	1	0	...	5	$y^{(4)}$
$x^{(6)}$	0	0	1	...	1	0	0	0	...	0.5	0	0.5	0	...	9	0	0	0	0	...	1	$y^{(5)}$
$x^{(7)}$	0	0	1	...	0	0	1	0	...	0.5	0	0.5	0	...	12	1	0	0	0	...	5	$y^{(6)}$
	A	B	C	...	Ti	NH	SW	ST	...	Ti	NH	SW	ST	...	Time	Ti	NH	SW	ST	...		
	User				Movie					Other Movies rated						Last Movie rated						

- 将用户点击item看作是用户（某些特征）与item（某些特征）之间的交互，比如小女孩会喜欢芭比娃娃
- 可以灵活地融合多种的信息，直接作为attribute并入就行
- 时间复杂度 $O(kn)$ ， k 为一条transaction的特征数
- 输入transaction（连续量continuous、独热表达univalent field、multivalent field），输出该条transaction对应用户给对应item的点击概率（评分），同时可以学习得到每一个attribute的表示

基于内容的推荐



表示新闻(news representation) -> 构建用户画像(user profile) -> 初步对比获取候选集(retrieval) -> 详细对比计算item的评分，进行排序(ranking)

News representation:

- 输入新闻，输出新闻的表示向量
- **传统方法**
 - Tf-Idf
 - 矩阵分解 (word-doc矩阵)
- **Neural**
 - Multi-Layer Perceptron, 将n-gram输入MLP得到表示
 - CNN捕捉local text, 词级别、短语级别、句子级别
- **Attention**
 - 普通Attention计算新闻中重要的词
 - Self-Attention计算新闻内部词和词的关联, 去噪
 - Multi-View Attention计算不同级别/领域的新闻表示, 计算重要的级别/领域
- **Graph**
 - 提取新闻中的实体, 用实体在知识图谱中的嵌入丰富其表示
 - 构造Item-Item的co-visitation Graph, GNN得到其表示向量
 - 构造User-Item的二部图, GNN得到其表示向量

User Representation:

- 输入用户点击的新闻的表示，输出用户的表示向量
 - 核心在于如何结合/串连用户点击的新闻
- **Sequential**
 - RNN
 - Session-based
 - 在Session中构建co-visitation graph，学习session的表示，将其作为用户的表达向量
 - Markov序列，用户点击预期回报最大的新闻
- **Not Sequential**
 - 构造随机初始化的Lookup table表示用户；MLP将用户ID映射为用户向量
 - Attention
 - 普通Attention计算对当前用户重要的新闻
 - Self-Attention计算新闻之间的关联
 - Multi-View Attention计算不同子空间/角度的用户表示的重要性
 - Graph
 - 构造User-Item的二部图，用邻居节点的信息丰富自身表达

Retrieval:

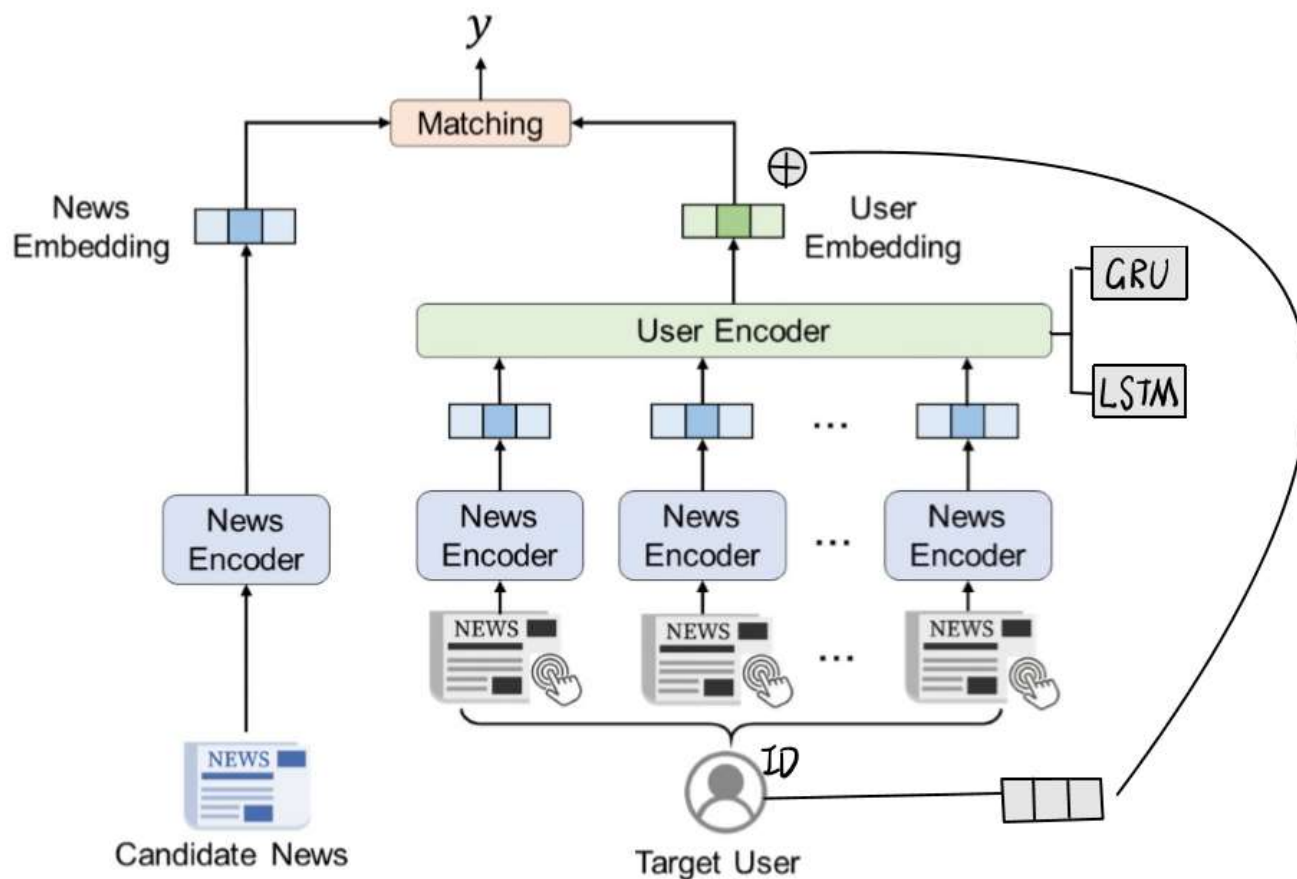
- 生成候选新闻集
 - 给定user profile和item representation在同一个隐空间中，则可以使用KNN进行搜索，初步选取和user相关的item

Ranking:

- 输入用户表示 u 和新闻表示 v ，输出 u 点击 v 的概率，据此进行排序
- **基于期望**
 - Contextual Bandit将期望回报看做是contextual feature的线性求和，Q-Learning
- **基于logistic回归**
 1. 计算user和item的相关度 (relevance)
 - 内积
 - Cos相似度
 - MLP
 2. 根据相关度估计user点击item的概率
 - Sigmoid(relevance)
 - Softmax(relevance), 将负例的信息考虑进来
 3. 最大化用户点击正例概率
 - 二分类, binary cross-entropy
 - 多分类, cross-entropy/MLE
- **基于FM**
 - 将同时包含user和item信息的向量输入FM作为transaction向量
- **基于矢量距离 (metric learning)**
 - 将点击事件也表示为向量，那么user向量+click向量 = item向量

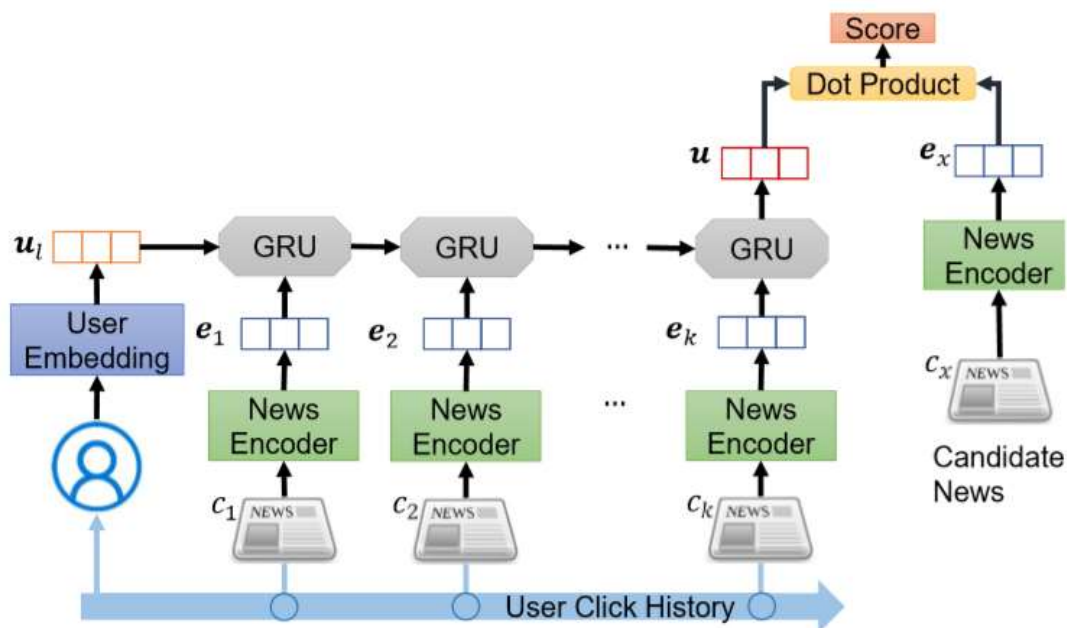
- Long- & short-term, Exploit & Explore

建模用户长短期画像，以及探索用户兴趣

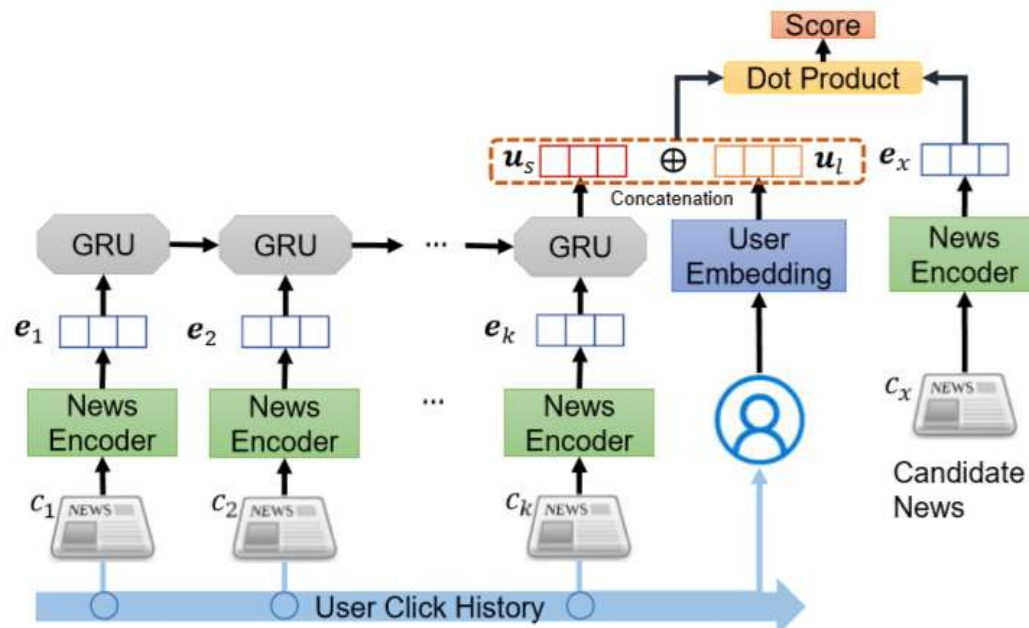


通过
MLP/lookuptable
将用户ID变成向量，
作为用户的长期画
像，也叫作
implicit profile

Neural News Recommendation with Long- and Short-term User Representations



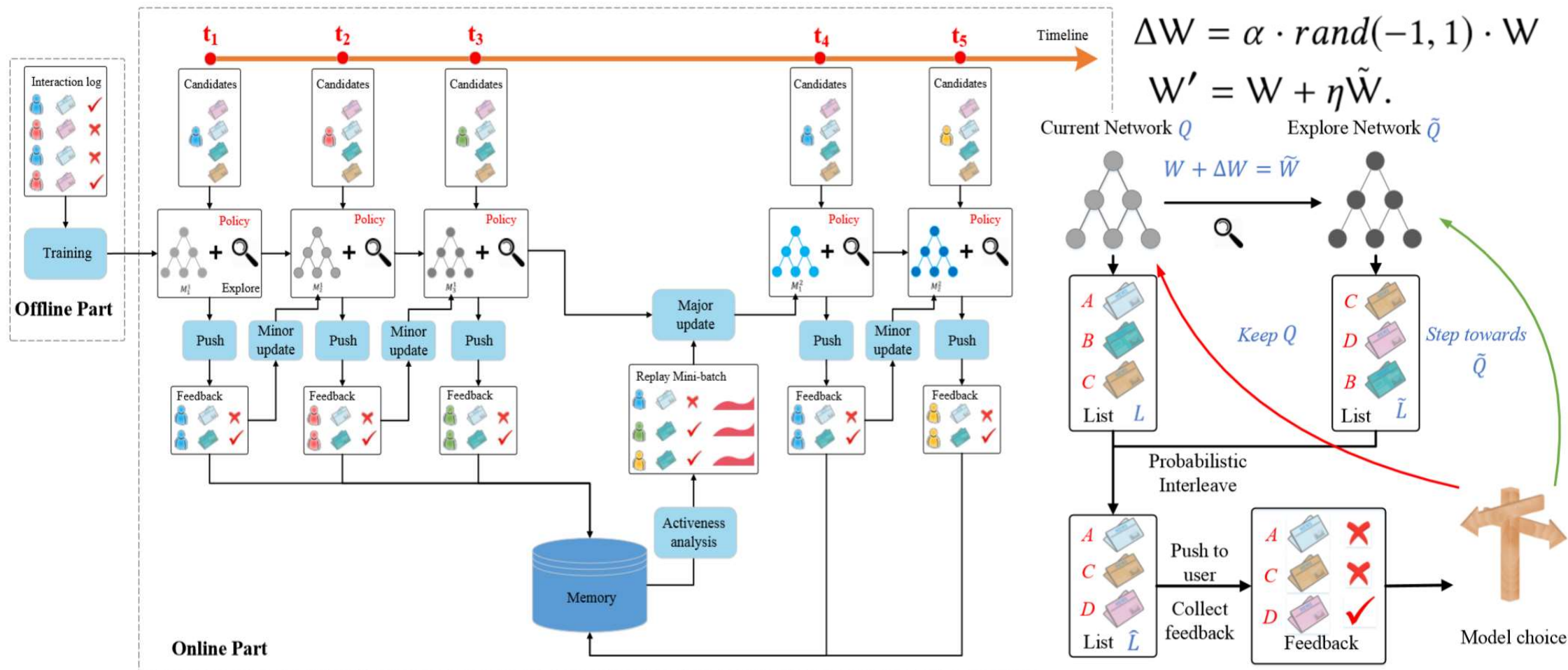
(a) LSTUR-ini.



(b) LSTUR-con.

- **用户有长期兴趣，和短期兴趣，两者应该分开对待；** 将新闻标题中各个单词表示为向量，然后cnn，得到新闻向量
- Short-term user repr:
 - 用GRU串连用户点击的新闻
 - 也可以使用带时间因素的LSTM串连
- Long-term user repr:
 - 使用随机初始化的lookup table，一个用户对应一行，可训；由于不一定每一个用户都有长期画像，使用伯努利分布将其mask；
- 结合long-term和short-term
 - 一种将long-term作为GRU的初始状态，另一种拼接long-term和short-term
 - 拼接两个向量，用MLP映射为实数a，最终分数为 $a \cdot \text{short} + (1 - \text{short}) \cdot \text{long}$

DRN: A Deep Reinforcement Learning Framework for News Recommendation

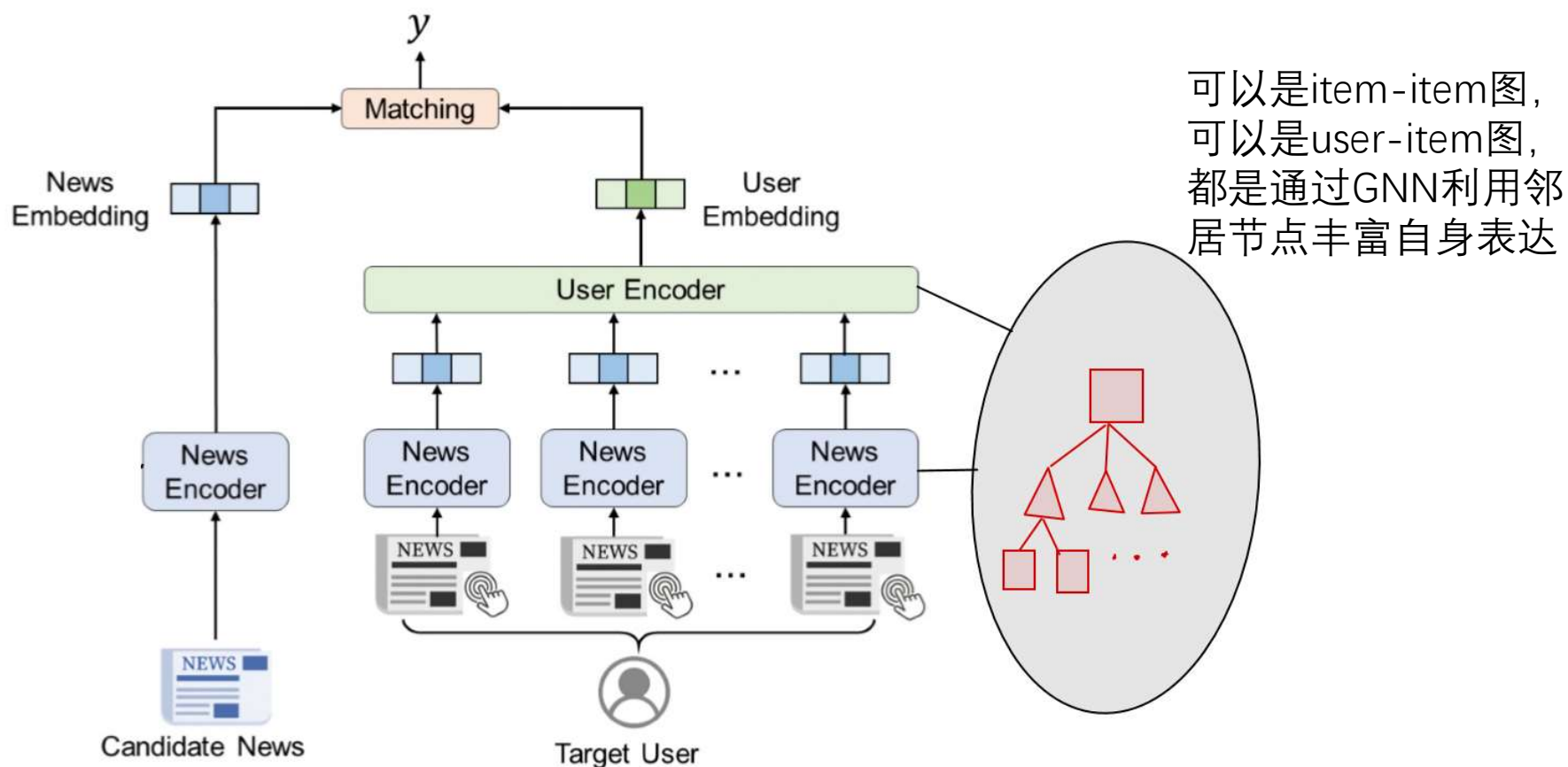


$$y_{s,a,t} = r_{a,t+1} + \gamma Q(s_{a,t+1}, \arg \max_{a'} Q(s_{a,t+1}, a'; W_t); W'_t)$$

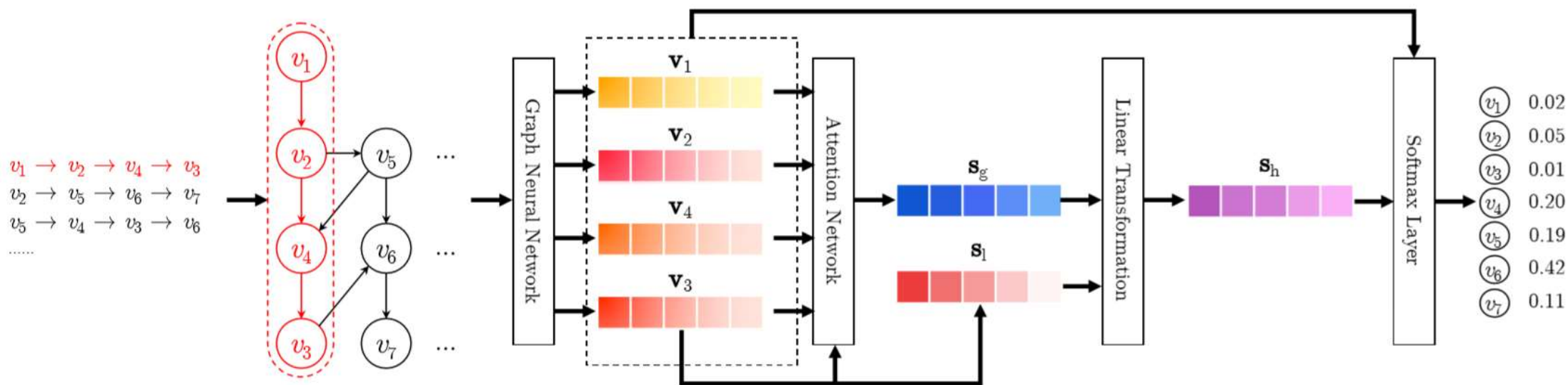
- **强化学习解决用户兴趣的动态变化**, Deep Q-Learning, 线下学习4个feature
- 相比于ucb, greedy等使用更细致的explore方法, Minor update, Major update

- Graph

基于user、item构建图， GNN捕捉high order interaction

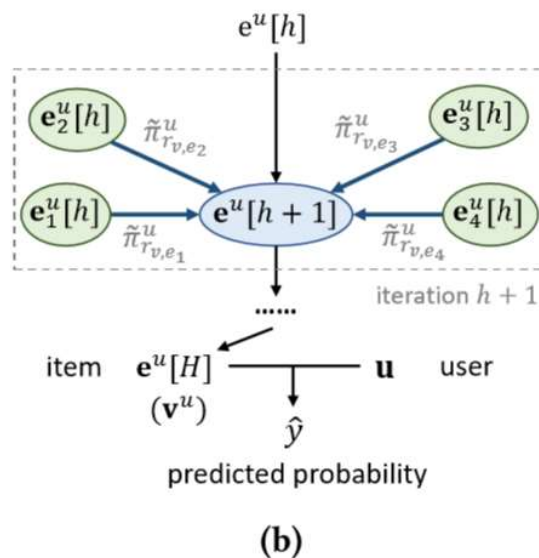
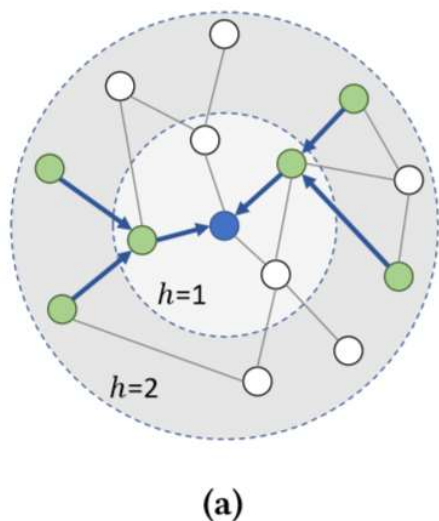


Session-Based Recommendation with Graph Neural Networks



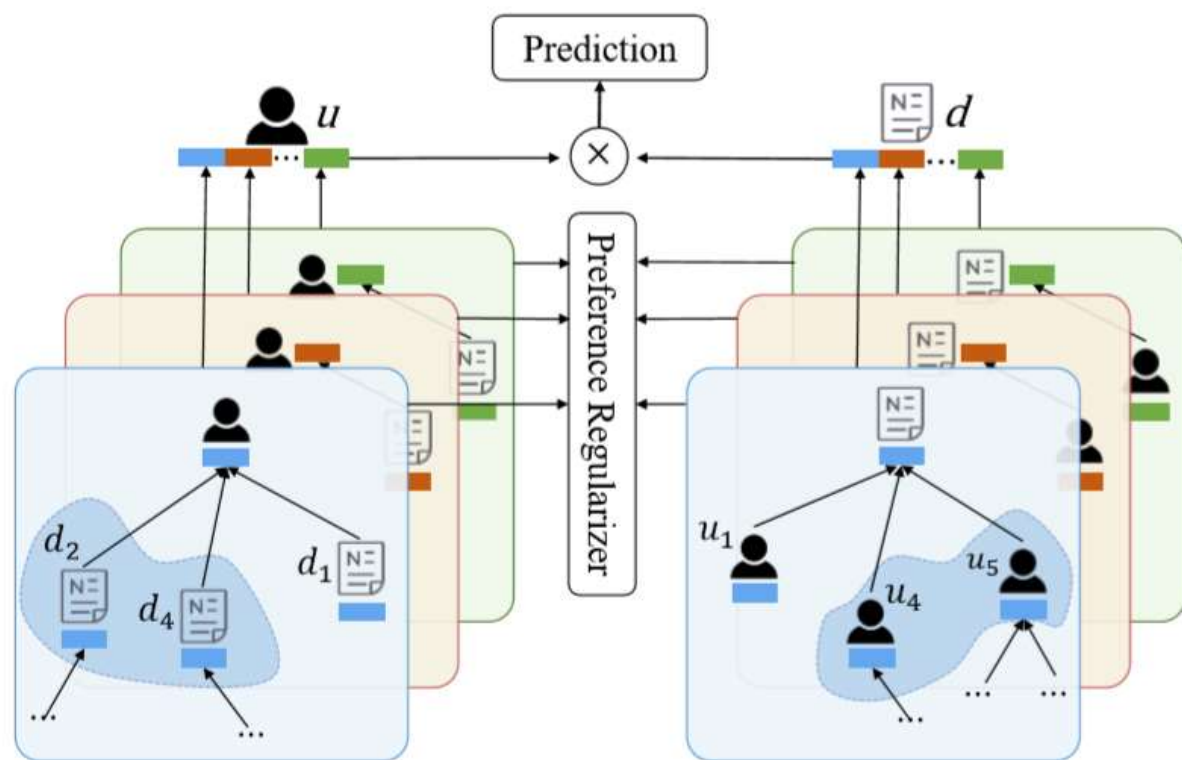
- **建模用户点击序列中item和item之间复杂的交互**
- 针对一个session内的点击序列，以item为节点构造图，边 $v_i \rightarrow v_j$ 代表用户在点击 v_i 后点击了 v_j
 - Session是数据集中划好的
- 各个节点为可训的向量，使用Gated GNN迭代更新每个节点的表示向量
- 将session的表示作为用户的表示，即根据session表示来判断下一个用户会点击的item
 - 使用session中用户最后点击的item的表示作为local embedding
 - 使用session中用户点击的所有item的attention求和作为global embedding

Knowledge Graph Convolutional Networks for Recommender Systems



- 现有的知识图谱表示模型主要适用于链接预测等在图上进行的操作，因此提出用来推荐的图模型
- 使用MovieLens, BookCrossing, Last-FM三种数据集
- 以item为主要的节点，使用Microsoft Satori构造无向图，边为item和item之间一些属性的关联（具有相同评分，具有相同主题）
- 输入用户表达向量 u （可训）, v （可训）
 - 计算 u 对 v 各个边的感兴趣程度，将其作为权重
 - 扩充 v 的表达，参与下一轮迭代，直到收敛
 - 每一轮迭代中可以扩展 H 步

Graph Neural News Recommendation with Unsupervised Preference Disentanglement



- 使用preference regularizer确保隐空间之间Disentanglement:
- 给定第 l 个隐空间的用户表达 $u^{(l)}$, 一层前馈神经网络将 $u^{(l)}$ 映射到 K 维 (一共有 K 个隐空间), softmax计算这个向量属于各个隐空间的概率, 让其分布在当前隐空间 (第 l 个) 的概率最大

- 用户的兴趣是多层面的, 并反应在复杂的user-item交互上
- PCNN得到新闻的表示 h_d
- 随机初始化user表达 h_u , 将 h_d , h_u 投射到 k 个隐空间中
- 每个隐空间构造user-article的二部图
- 以user为例, 计算当前隐空间中其和各个邻居news的相关度, 然后加权集成到自身表达上

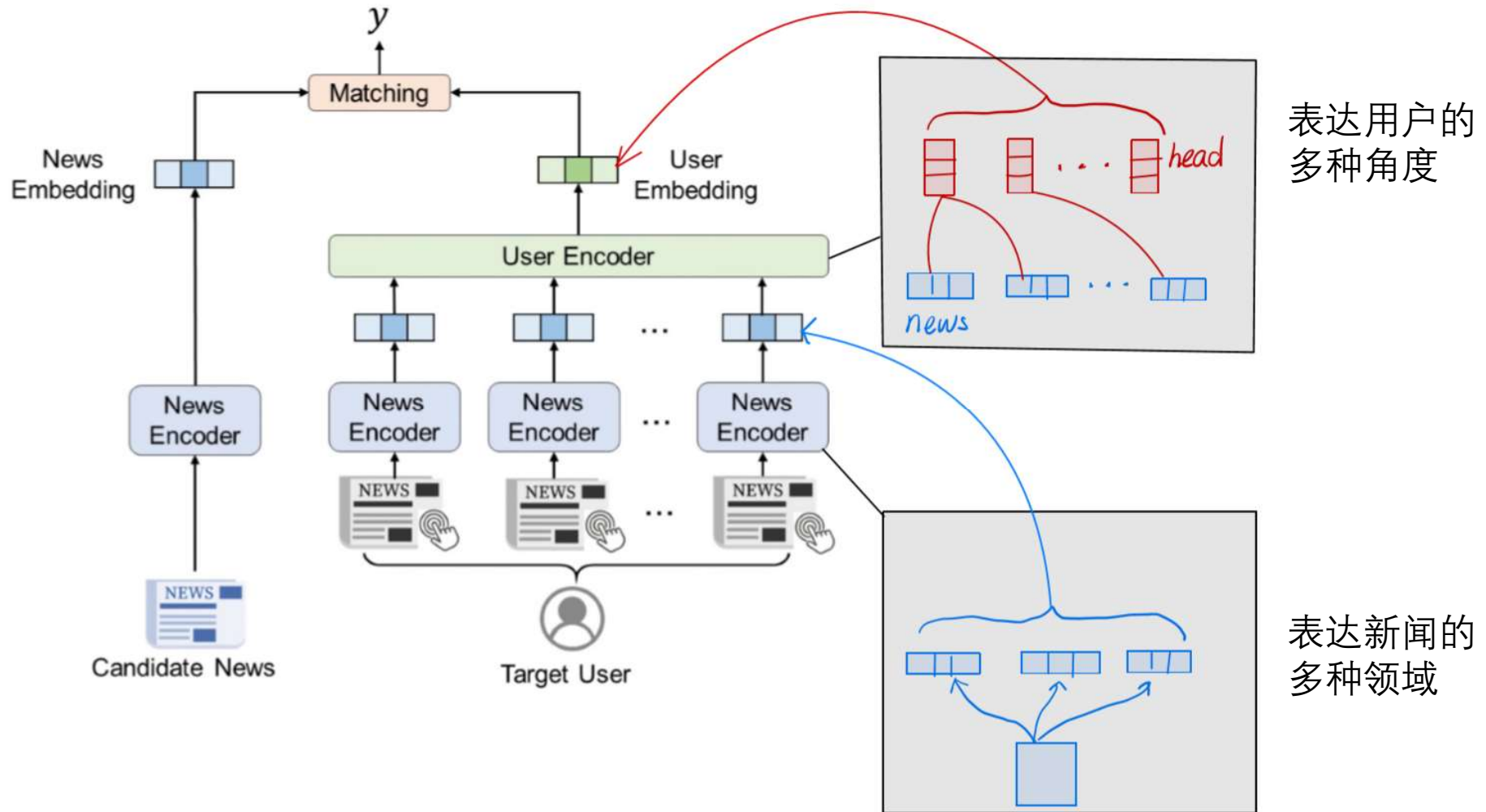
$$r_{d,k}^{(t)} = \frac{\exp(\mathbf{z}_{u,k}^{(t)\top} \mathbf{s}_{d,k})}{\sum_{k'=1}^K \exp(\mathbf{z}_{u,k'}^{(t)\top} \mathbf{s}_{d,k})},$$

$$\mathbf{z}_{u,k}^{(t+1)} = \frac{\mathbf{s}_{u,k} + \sum_{d:(u,d) \in \mathcal{G}} r_{d,k}^{(t)} \mathbf{s}_{d,k}}{\|\mathbf{s}_{u,k} + \sum_{d:(u,d) \in \mathcal{G}} r_{d,k}^{(t)} \mathbf{s}_{d,k}\|_2},$$

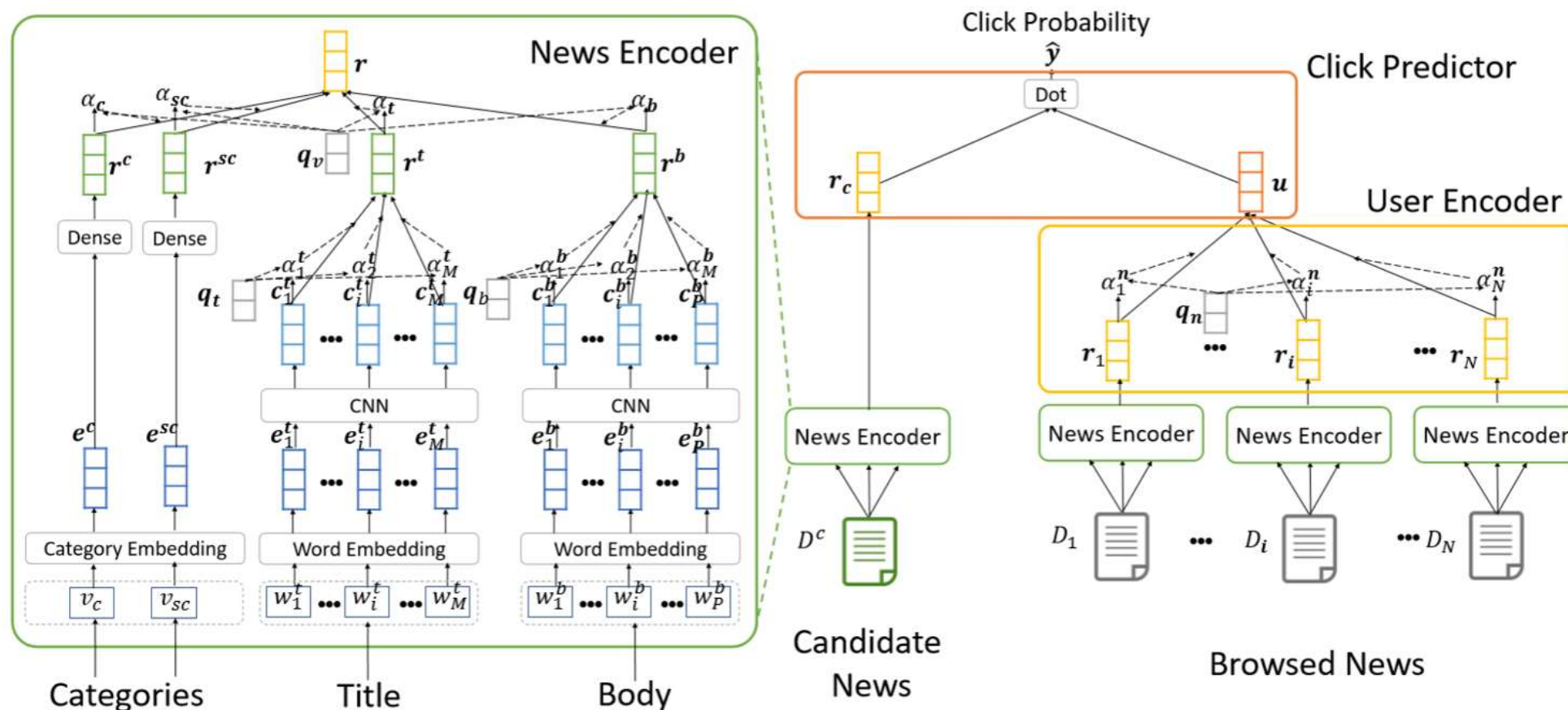
- 迭代 T 次, 直至收敛
- 最后 k 层的表达拼起来得到最终用户表达
- Item同理

- Multi-View

多种角度/领域/级别

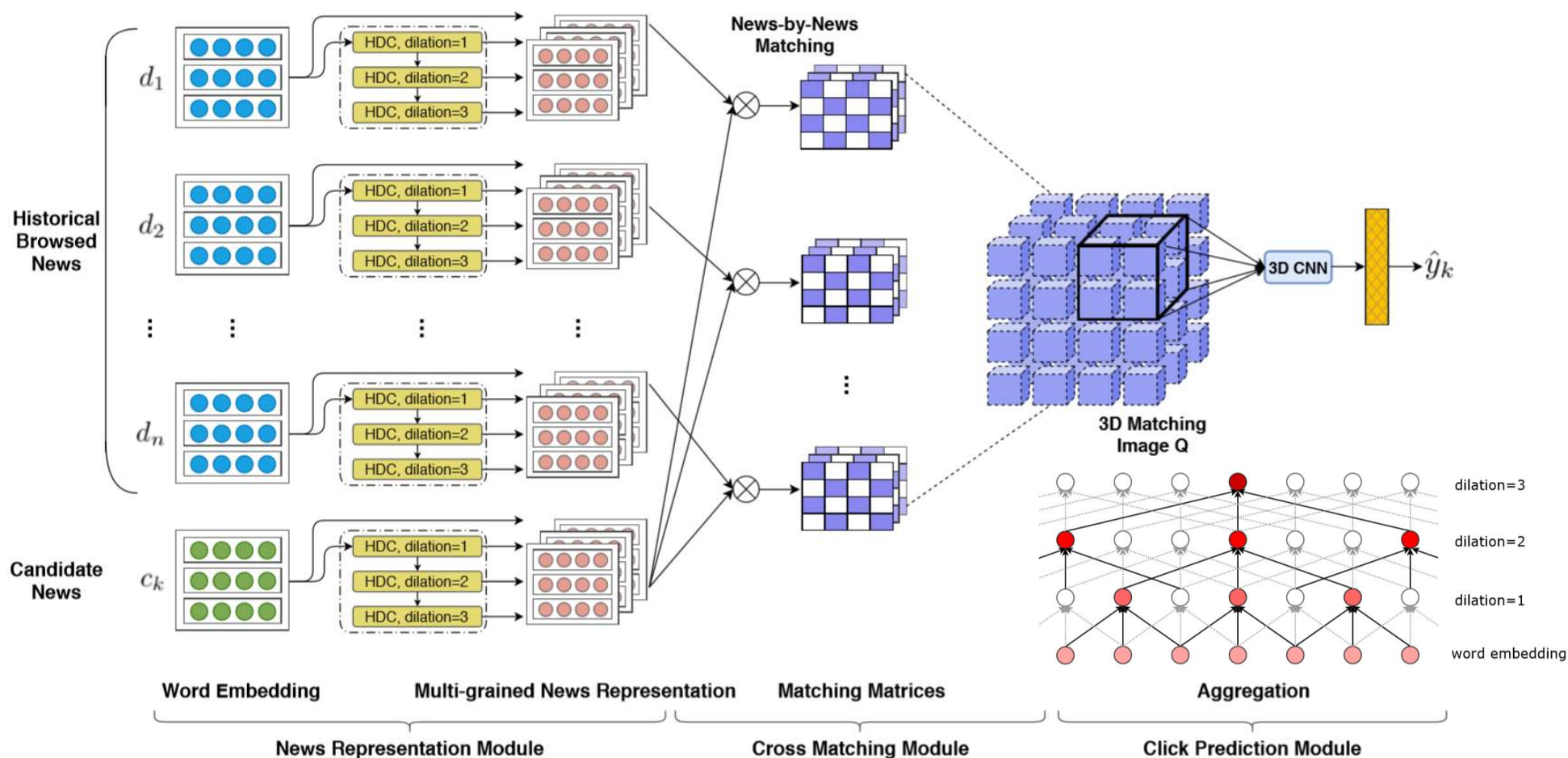


Neural News Recommendation with Attentive Multi-View Learning



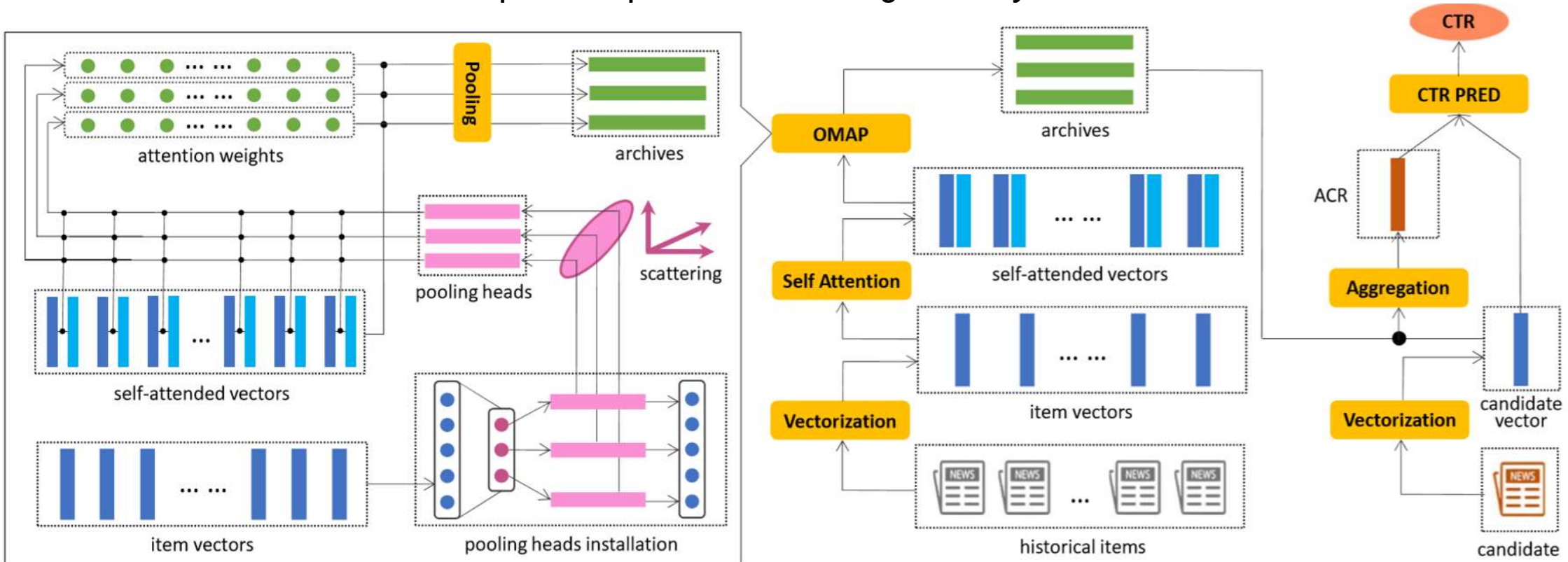
- 新闻的标题、正文、分类有其不同的含义和重要性
- 在word-embedding的时候分开考虑title中的word、body中的word，将其视作不同的view，multi-view attention
- 这里的query都是训练得到的，其实可以使用用户id的dense vector作为query，personalized attention

Fine-grained Interest Matching for Neural News Recommendation



- **为了捕捉新闻中多个层级的文本信息**，使用带膨胀率的HDCN处理新闻文本，每层表达叠在一起变成三维
- 分别与candidate news对比，计算每个层级上的两者相关度
- Candidate和所有history news的相关度矩阵叠在一起形成最终矩阵Q，使用3D-CNN得到最终的向量包含不同层级的用户和新闻的信息，MLP得分

Hi-Fi Ark: Deep User Representation via High-Fidelity ArchiveNetwork



- 很多方法推荐时需要访问用户所有的浏览历史，慢，为了快速准确地推荐，将用户的画像压缩到多个archive中
- 输入用户点击过的item，首先使用CNN将其转化为向量，使用self-attention去除噪声，之后选取该向量空间的K个正交基作为pooling head，刻画用户k个不同方面的兴趣
- 以每一个pooling head为query，对用户历史记录中每个item的向量做attention集成，形成k个archive，证明得到如果pooling head保持正交性，则这些archive近乎无损地保留了用户的偏好
- 推荐时以candidate的向量为query做attention后集成archive，得到用户向量

Collaborative Metric Learning with Memory Network for Multi-Relational Recommender Systems

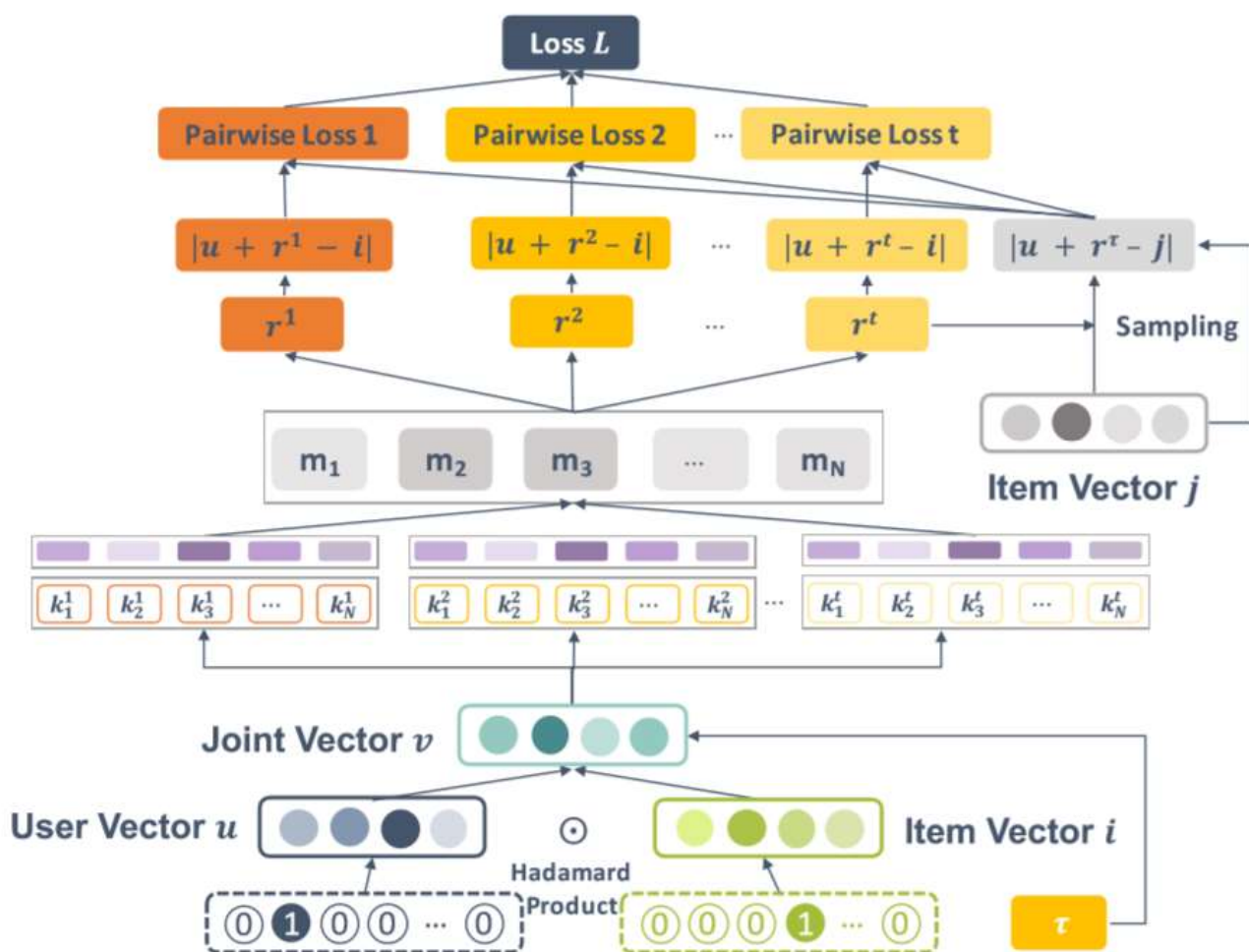


Figure 2: The architecture of MRMN.

• 考虑多种用户反馈，点击、跳过、收藏、加入购物车、购买等

• 将user ID和item ID投射到d维dense vector

• 两者Hadamard product得到d维联合向量 v

• 针对每一种反馈，为每一个用户构造一个memory matrix，每一列都是一个d维key，这个key是学习得到的，可以理解为当前反馈类型的feature

• 构造一个全局共享的memory matrix M ，每一列是构成关系表示的单元

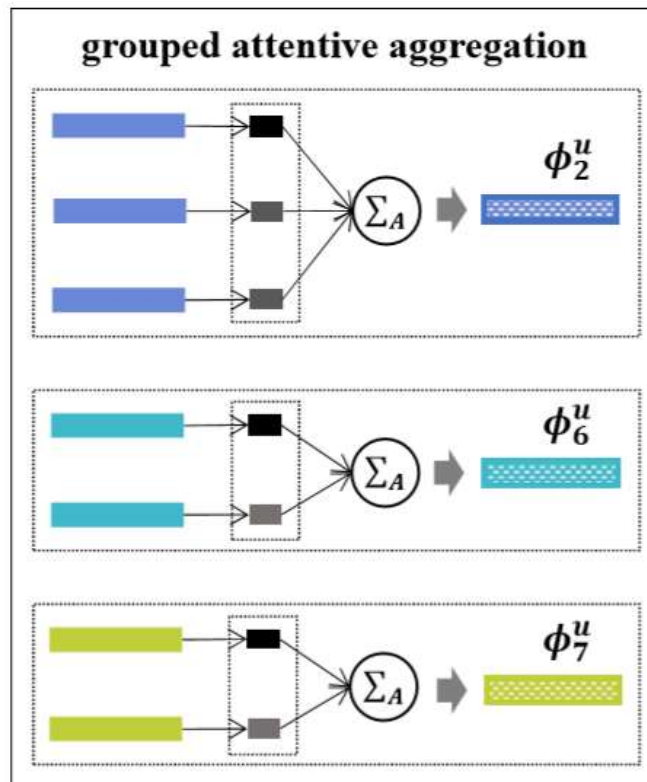
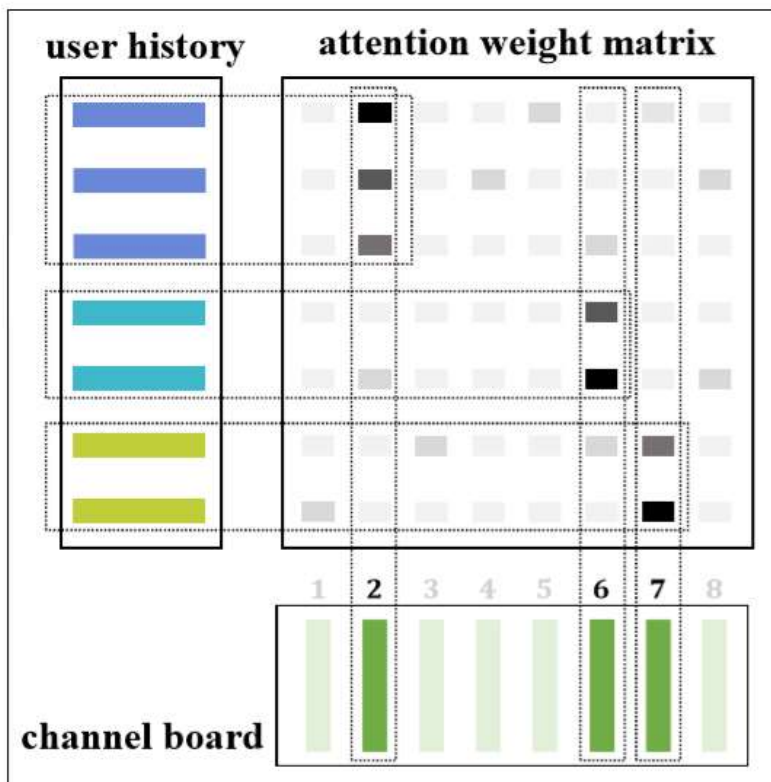
• 根据联合向量 v 和用户各个反馈的memory matrix的权重集成这些单元，形成各个关系的表示

• 矢量距离计算损失，让所有正例的得分高于负例

$$\mathcal{L} = \sum_{\tau} \sum_{(u, i, \tau) \in \eta} \sum_{(u, j, \tau) \notin \eta} \phi(s(u, i, \tau) + \lambda^{\tau} - s(u, j, \tau))$$

• Lambda用来表示反馈类型的偏置，比如购买的lambda就比点击的lambda高

Octopus: Comprehensive and Elastic User Representation for the Generation of Recommendation Candidates



- 现有的multi-head方法都仅考虑了固定head的个数，导致了under/over channelization，本文聚焦于生成候选集
- 构造很多个channel，每个channel两个head，group head初始化为所有item repr(DSSM)空间的正交基，能够覆盖用户的所有兴趣，在训练过程中保持正交；aggregation head用于生成当前channel下的用户画像；
- 以item repr为query，group head为key，计算和用户浏览历史中每一个item最相关的channel，激活相应channel
- 将激活相同channel的用户历史点击item分为一组，使用当前channel的aggregation head作为query，item repr作为key和value，生成当前channel的用户画像

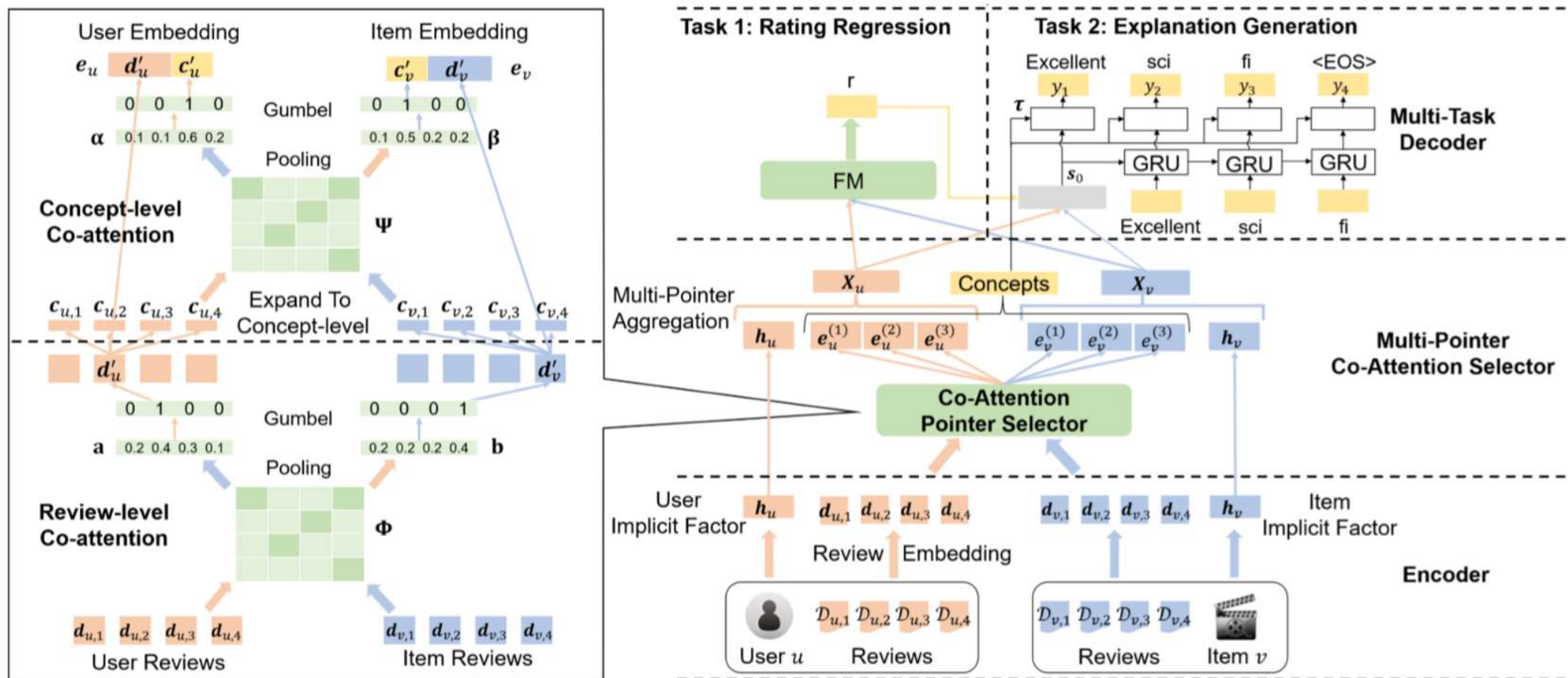
- 每一个channel对应的用户画像都有k-nearest-neighbor，为不同channel分配不同的名额 (k)
- 使用MLP将每一个user repr映射为归一化的权重 (分布)
- 使用和待聚类item最相似的user repr作为lable，训练多分类器
- a-proportional计算每个channel的名额

$$\beta_l = \frac{(\gamma_l)^\alpha}{\sum_{\gamma_* \in \{\gamma_*\}^u} (\gamma_*)^\alpha}, \text{ where } \alpha > 0.$$

- Interaction

更细粒度上建模user和item之间的交互

Co-Attentive Multi-Task Learning for Explainable Recommendation



- 为用户生成推荐的解释有助于提升用户的信任程度，提高CTR，还有别的multi-task应用
- 从多个review 中选出重要的一个，从这一个中选出重要的多个concept (entity)
- 这种co-attention都是基于review， 因为用户和item都有其review（由词组成）， 所以建模词和词之间的交互

- [1] Google News Personalization Scalable Online Collaborative Filtering
- [2] A contextual-bandit approach to personalized news article recommendation
- [3] Content-Based News Recommendation
- [4] Personalized News Recommendation A Review and an Experimental Investigation
- [5] Personalized news recommendation based on click behavior
- [6] SCENE A Scalable Two-Stage Personalized News Recommendation System
- [7] Content-based Collaborative Filtering for News Topic Recommendation
- [8] Embedding-based news recommendation for millions of users
- [9] Matrix factorization techniques for recommender systems
- [10] Factorization Machines
- [11] Wide & Deep Learning for Recommender Systems
- [12] Maximum Likelihood, Logistic Regression, and Stochastic Gradient Training
- [13] Learning deep structured semantic models for web search using clickthrough data
- [14] Combining Content-Based and Collaborative Filtering in an Online Newspaper
- [15] Neural Networks: Tricks of the Trade (Second edition)
- [16] Deepfm a factorization machine based neural network for ctr prediction
- [17] News session based recommendation fusing deep neural networks
- [18] Towards better representation learning for personalized news recommendation a multi-channel deep fusion approach
- [19] Session-based recommendations with recurrent neural networks
- [20] Neural News Recommendation with Attentive Multi-View Learning
- [21] Neural News Recommendation with Long- and Short-term User Representation
- [22] Neural News Recommendation with Multi-Head Self-Attention
- [23] Npa Neural news recommendation with personalized attention
- [24] DKN Deep Knowledge-Aware Network
- [25] DRN A Deep Reinforcement Learning Framework for News Recommendation
- [26] Adaptive User Modeling with Long and Short-Term Preference for Personalized Recommendation
- [27] Co-Attentive Multi-Task Learning for Explainable Recommendation
- [28] Explainable Recommendation through Attentive Multi-View Learning
- [29] Fine-grained Interest Matching for Neural News Recommendation
- [30] Geography-Aware Sequential Location Recommendation
- [31] Hi-Fi Ark Deep User Representation via High-Fidelity Archive Network
- [32] Knowledge Graph Convolutional Networks for Recommender Systems
- [33] Multi-Task Feature Learning for Knowledge Graph Enhanced Recommendation
- [34] Neural News Recommendation with Heterogeneous User Behavior
- [35] Neural News Recommendation with Topic-Aware News Representation
- [36] Personalized Multimedia Item and Key Frame Recommendation
- [37] Reviews Meet Graphs Enhancing User and Item Representations for Recommendation with Hierarchical Attentive Graph Neural Network
- [38] Session-Based Recommendation with Graph Neural Networks
- [39] Graph Neural News Recommendation with Unsupervised Preference Disentanglement
- [40] LightRec A Memory and Search-Efficient Recommender System
- [41] Towards Explainable Conversational Recommendation
- [42] Knowledge-Aware Document Representation for News Recommendations
- [43] Combining Collaborative Filtering and Search Engine into Hybrid News Recommendation
- [44] Collaborative Metric Learning with Memory Network for Multi-Relational Recommender Systems
- [45] Sequential Recommendation with User Memory Networks
- [46] Leveraging Demonstrations for Reinforcement Recommendation Reasoning over Knowledge Graphs
- [47] EntityDuet Neural Ranking Understanding the Role of Knowledge Graph Semantics
- [48] Octopus Comprehensive and Elastic User Representation for the Generation of Recommendation