

LightGCN

轻量级卷积神经网络

潘微科

深圳大学计算机与软件学院

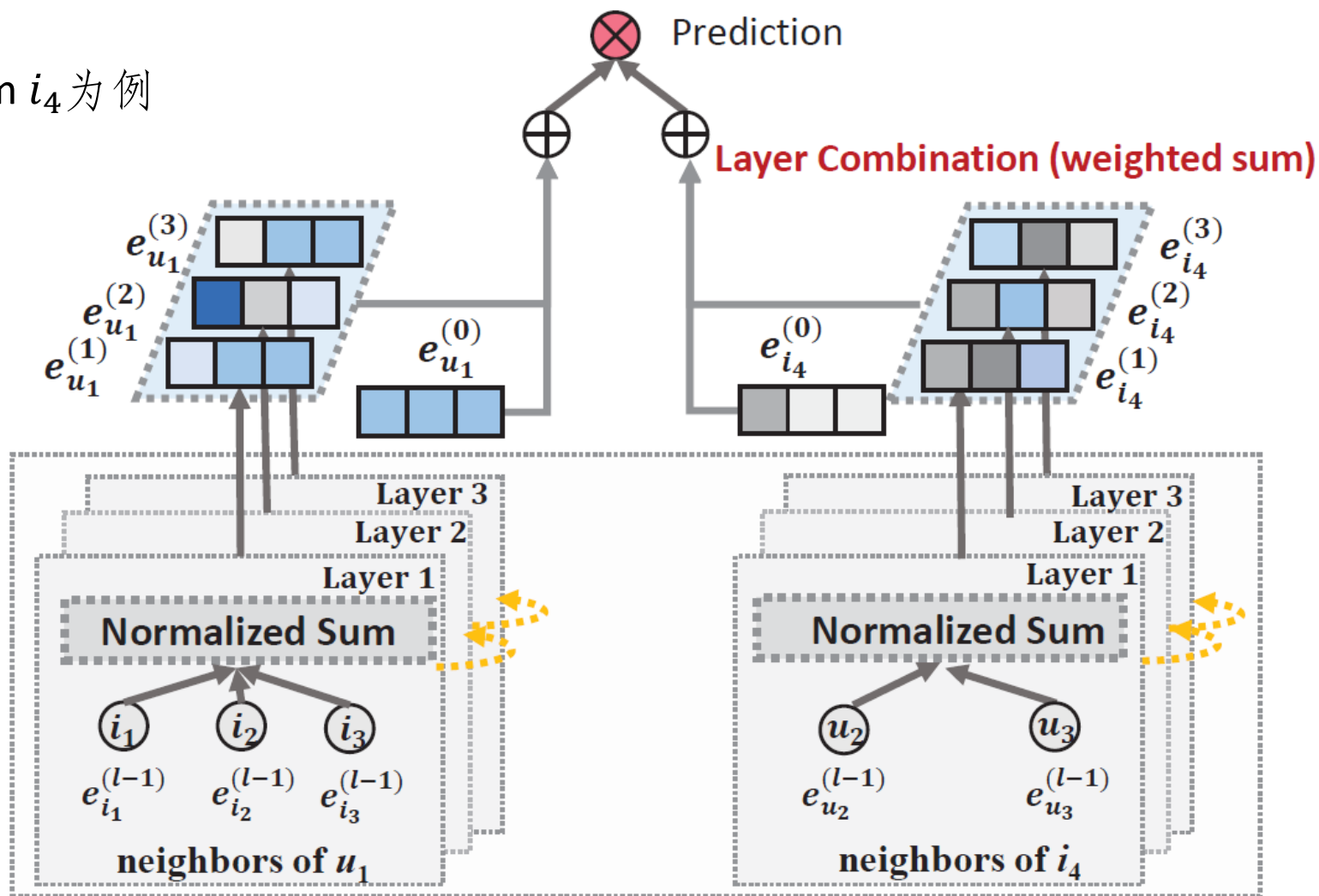
Xiangnan He, Kuan Deng, Xiang Wang, Yan Li, Yongdong Zhang, Meng Wang. **LightGCN: Simplifying and powering graph convolution network for recommendation.** SIGIR 2020.

介绍

- 研究的问题：单类协同过滤（OCCF）
- 研究的动机：
 - 将图卷积神经网络（GCN）应用于OCCF问题
 - 对神经图协同过滤（NGCF, SIGIR 2019）算法进行简化和优化（效果提升）
- 模型的名称：LightGCN

模型示意图

- 以 user u_1 和 item i_4 为例



构图 (graph construction)

- 用户-物品二部图: user-item bipartite interaction graph

卷积操作 (convolution operation)

- 卷积操作包含：近邻聚合、表征更新
- 即示意图中的“Normalized Sum”

$$\mathbf{e}_u^{(k+1)} = \sum_{i \in \mathcal{N}_u} \frac{1}{\sqrt{|\mathcal{N}_u|} \sqrt{|\mathcal{N}_i|}} \mathbf{e}_i^{(k)}$$

$$\mathbf{e}_i^{(k+1)} = \sum_{u \in \mathcal{N}_i} \frac{1}{\sqrt{|\mathcal{N}_i|} \sqrt{|\mathcal{N}_u|}} \mathbf{e}_u^{(k)}$$

- 注：在LightGCN中，**可训练参数**是第0层的user embedding和item embedding，因为其他层的embedding都可以由它们得到。
- 注：从模型参数数量的角度来看，与矩阵分解的参数数量相当。

表征读出

- 作为 user u 和 item i 的 final embedding
- 即示意图中的 “Layer Combination (weighted sum)”

$$\mathbf{e}_u = \sum_{k=0}^K \alpha_k \mathbf{e}_u^{(k)}$$

$$\mathbf{e}_i = \sum_{k=0}^K \alpha_k \mathbf{e}_i^{(k)}$$

- 注：权重用 $1/(K+1)$ 通常能取得较好的结果。

预测公式 (prediction rule)

- user u 对item i 的偏好

$$\hat{y}_{ui} = \mathbf{e}_u^T \mathbf{e}_i$$

优化函数 (optimization function)

- BPR loss + 正则化项

$$L_{BPR} = - \sum_{u=1}^M \sum_{i \in \mathcal{N}_u} \sum_{j \notin \mathcal{N}_u} \ln \sigma(\hat{y}_{ui} - \hat{y}_{uj}) + \lambda \|\mathbf{E}^{(0)}\|^2$$

- 注：优化的时候使用Adam优化器，同时使用mini-batch的方式。

相关工作

- NGCF (Neural Graph Collaborative Filtering)
- Xiang Wang, Xiangnan He, Meng Wang, Fuli Feng, Tat-Seng Chua. **Neural graph collaborative filtering**. SIGIR 2019.

- 简化了其中的 **feature transformation** (特征转换) 和 **nonlinear activation** (激活函数)

$$\mathbf{e}_u^{(k+1)} = \sigma\left(\underline{\mathbf{W}}_1 \mathbf{e}_u^{(k)} + \sum_{i \in \mathcal{N}_u} \frac{1}{\sqrt{|\mathcal{N}_u| |\mathcal{N}_i|}} (\underline{\mathbf{W}}_1 \mathbf{e}_i^{(k)} + \underline{\mathbf{W}}_2 (\mathbf{e}_i^{(k)} \odot \mathbf{e}_u^{(k)}))\right)$$

$$\mathbf{e}_i^{(k+1)} = \sigma\left(\mathbf{W}_1 \mathbf{e}_i^{(k)} + \sum_{u \in \mathcal{N}_i} \frac{1}{\sqrt{|\mathcal{N}_u| |\mathcal{N}_i|}} (\mathbf{W}_1 \mathbf{e}_u^{(k)} + \mathbf{W}_2 (\mathbf{e}_u^{(k)} \odot \mathbf{e}_i^{(k)}))\right)$$

- 注: \mathbf{W}_1 和 \mathbf{W}_2 是要学习的参数, 用于做 feature transformation。
- 注: 在“表征读出”阶段, NGCF 将 $K+1$ 个表征拼接起来, 而非线性加权相加 (LightGCN)。

实验结果

- 相比NGCF, LightGCN平均提升了约**16%**。

谢谢!