

A Hybrid Collaborative Filtering Model with Deep Structure for Recommender Systems

会议信息：AAAI-17

来自携程的作者，文章链接：<http://aaai.org/ocs/index.php/AAAI/AAAI17/paper/view/14676>

Abstract

协同过滤算法在实际的推荐系统中有着广泛的应用，却也面临着许多问题。给定的用户-物品的打分矩阵通常是十分稀疏的，效果并不理想。因而，现在开始加入更为丰富的辅助信息来缓解数据稀疏和冷启动问题。目前通过学习用户和商品的隐含表示的模型对打分信息和外部信息的利用还不够充分，本文通过深度学习方法，从打分矩阵和辅助信息中，学习了用户和商品的隐含表示，基于此提出一种混合协同过滤模型。通过在 MovieLens 和 BookCrossing 数据集上的实验验证其有效性。

Introduction

从过去到现在，实际的推荐系统中用的方法可以分为两大类：基于协同过滤和基于内容的过滤方法。后者主要通过用户和物品的描述性信息进行建模与推荐，而前者则在用户-物品的打分矩阵上实现矩阵分解，从而学习到一个低维的用户和物品的分布式向量表示（个人理解：与word2vector词向量有些类似，但是不像词向量那样隐含着词语之间的某种关系，但是相似的用户在这个向量表示的空间中的距离应该比较近，或者他们属于一个cluster）。因此，协同过滤方法要比基于内容过滤的方法具有更优秀的表现。虽然协同过滤主要面临着两大困境：

- 打分矩阵数据稀疏（通常打分数据的密度 $<1\%$ ）
- 冷启动问题，对于历史数据中没有出现过的用户和物品无法建模进行推荐

为了缓解这样的情况，势必要利用一些外部辅助信息作为补充。目前的工作主要是通过正则化方法，把这些额外的信息融合如矩阵分解的学习过程中，然而当这些信息并不十分相关，这样的方法就显得较为苍白 (Agarwal, Chen, and Long 2011)。

随着深度学习，尤其是它在表示学习方面的优异性能，这种方法也被融入推荐系统的协同过滤方法中。一方面，利用限制Boltzmann机进行协同过滤过程，但是没有考虑外部信息，仅仅是通过深度学习模型改进了协同过滤学习用户、物品的表示学习过程。另一方面，就是直接利用深度学习对外部信息进行特征学习，如用CNN处理图片，这一类做法在音乐类的推荐任务中的效果较好，有点类似基于内容的过滤方法。最近又兴起了一种Bayesian stacked denoising auto-encoder (SDAE) 方法，但需要手动设计许多超参。

本文为解决上述问题，提出了一种基于深度学习的混合式协同过滤模型。本文的主要贡献：

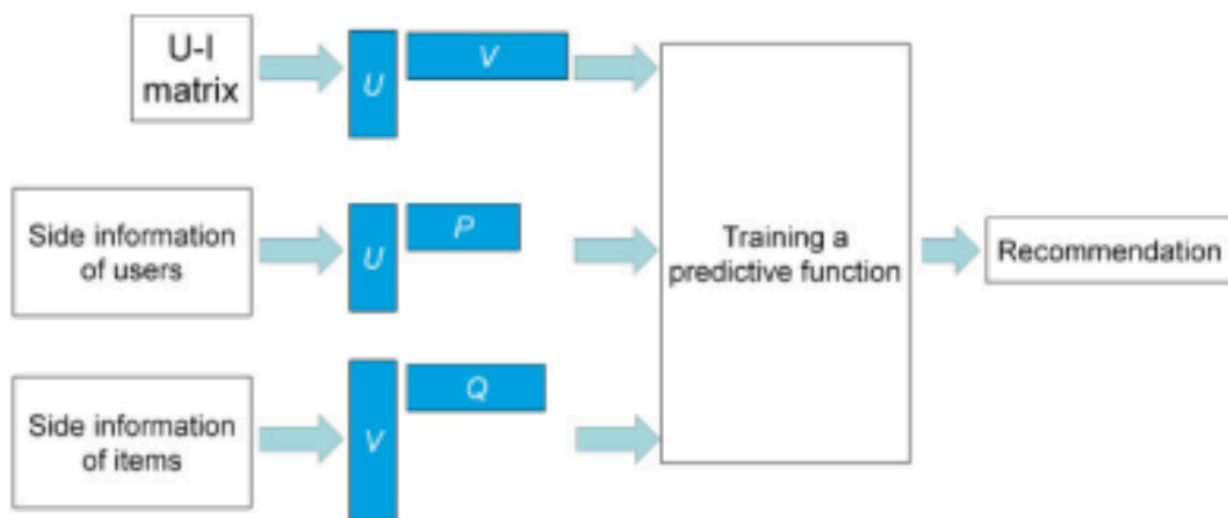
- 在SDAE的基础上提出一种扩展的additional SDAE，更为充分地利用辅助信息中有价值的部分，从而缓解冷启动和数据稀疏的问题；
- 提出一种混合式模型，综合了深度学习模型和矩阵分解模型，能够从外部辅助信息和传统的打分信息中学习到更为有效地特征表示；
- 在三个实际的数据集上验证了该模型的有效性，从RMSE和Recall指标上均超过了目前的state-of-art。

My Opinions

本文从立意上说，完成了一个多元异构数据融入推荐系统的混合式模型，是目前缓解数据稀疏和冷启动问题的重大解决方案。让我来做这项任务的话，结合深度学习方法，可能会局限于文中提到的相关工作：外部信息仅仅是根据外部信息的特征，使用合适的工具进行向量化（如CNN擅长处理图像、RNN擅长处理文本等其他序列化信息），不会想到这个过程也可以影响矩阵分解的学习过程。

虽然之后想到通过矩阵分解得到的用户、物品向量与各个辅助信息之间通过某种手段建立联系，从而学到具有更深刻表示能力的特征向量，可是效果并不理想。关于从多元异构的辅助信息中抽取特征或者是进行表示学习，也一直在调研，看到最多的方法就是各取所长，利用合适的武器处理合适的任务，产出合适的特征向量，通过concatenation操作，完成特征的扩展与融合，就像Google提出的模型中，wider的做法，但这依赖计算能力和存储空间的做法显然不够科学与合理。

想到有人提出这样的称为 [Collective Matrix Factorization](#) 的模型，就是我想到的典型做法：



这里CMF模型通过分别分解打分矩阵R，用户、物品的辅助信息矩阵，其中用户和物品出现在多个矩阵中，其所分解的隐向量都是一致的、共享的，从而建立联系。

何向南等人提出了Neural Factorization Machine进行了特征融合，可以减少单纯不假思索的concatenation带来的维数灾难，也带来了特征融合的新的方式，看起来更为优雅些。

Preliminaries and Models

Problem Definition

给定打分矩阵， $R \in \mathbb{R}^{m \times n}$ ，用户和物品的辅助信息分别是 $X \in \mathbb{R}^{m \times p}$ 和 $Y \in \mathbb{R}^{n \times q}$

$u_i, v_j \in \mathbb{R}^k$ 分别表示第i个用户和第j个物品的隐含特征向量，k表示维度，那么 $U = u_{1:m}$ ， $V = v_{1:n}$ ，就是我们学习的目标——用户和物品的隐含表示，进而预测未知的用户对物品的打分。

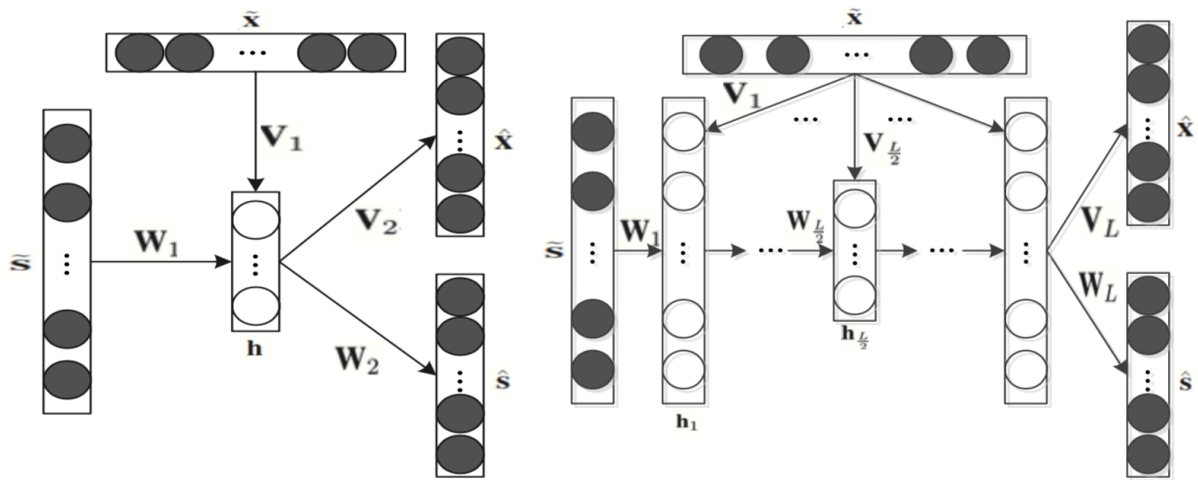
用矩阵分解的角度来说，就是求解

$$\arg \min_{U, V} \mathcal{L}(R, UV^T) + \lambda(\|U\|_F^2 + \|V\|_F^2)$$

的过程。

Additional Stacked Denoising Autoencoder

1. 自动编码器 (Autoencoder, AE) , 是一种神经网络模型, 包含一个编码器和解码器。编码器进行编码, 对于一个向量输入 s , 进行函数 $g(\cdot)$ 的作用, 解码器进行解码, 把编码器的结果重构为输入 s , 即保证了 $f(g(s)) \approx s$
2. 进一步, 在自动编码器中的输入中融入一些噪音, 可以提高模型的泛化能力, 使得学习到的结果更为鲁棒, 这就是DAE
3. Stacked DAE, SDAE是一个多层的AE, 前一层自编码器作为后一层自编码器的输入, 进一步压缩了特征维数, 有些类似DSSM, 可以更好地拟合数据, 开始向多层神经网络发展
4. Additional SDAE, aSDAE则是融合了外部信息的SDAE, 如下图所示, 这张图只是说明了用户打分融合用户外部信息, 物品角度也是类似的:



(a) Additional Denoising Autoencoder(aDAE) (b) Additional Stacked Denoising Autoencoder(aSDAE)

个人的理解, 有些类似Seq2Seq的模型, 但是是一种简化的方式。(b)图中左侧实心点的 \tilde{S} 是用户的打分信息, 上面实心的 \tilde{x} 是用户辅助信息, 右侧实心的 \hat{x} 和 \hat{S} 是通过AE解码后的结果, 是要与输入尽可能一致的。

具体过程: 一方面, 用户打分的数据通过一个多层自编码器, 得到一个中间表示, 图中的 $h_{\frac{L}{2}}$, 随后通过解码器部分, 最终的目的是解码出一个与输入类似的表示: 公式表示为:

$$h_l = g(W_l h_l + V_l \tilde{x} + b_l), \text{ where } l \in \{1, 2, \dots, L\}$$

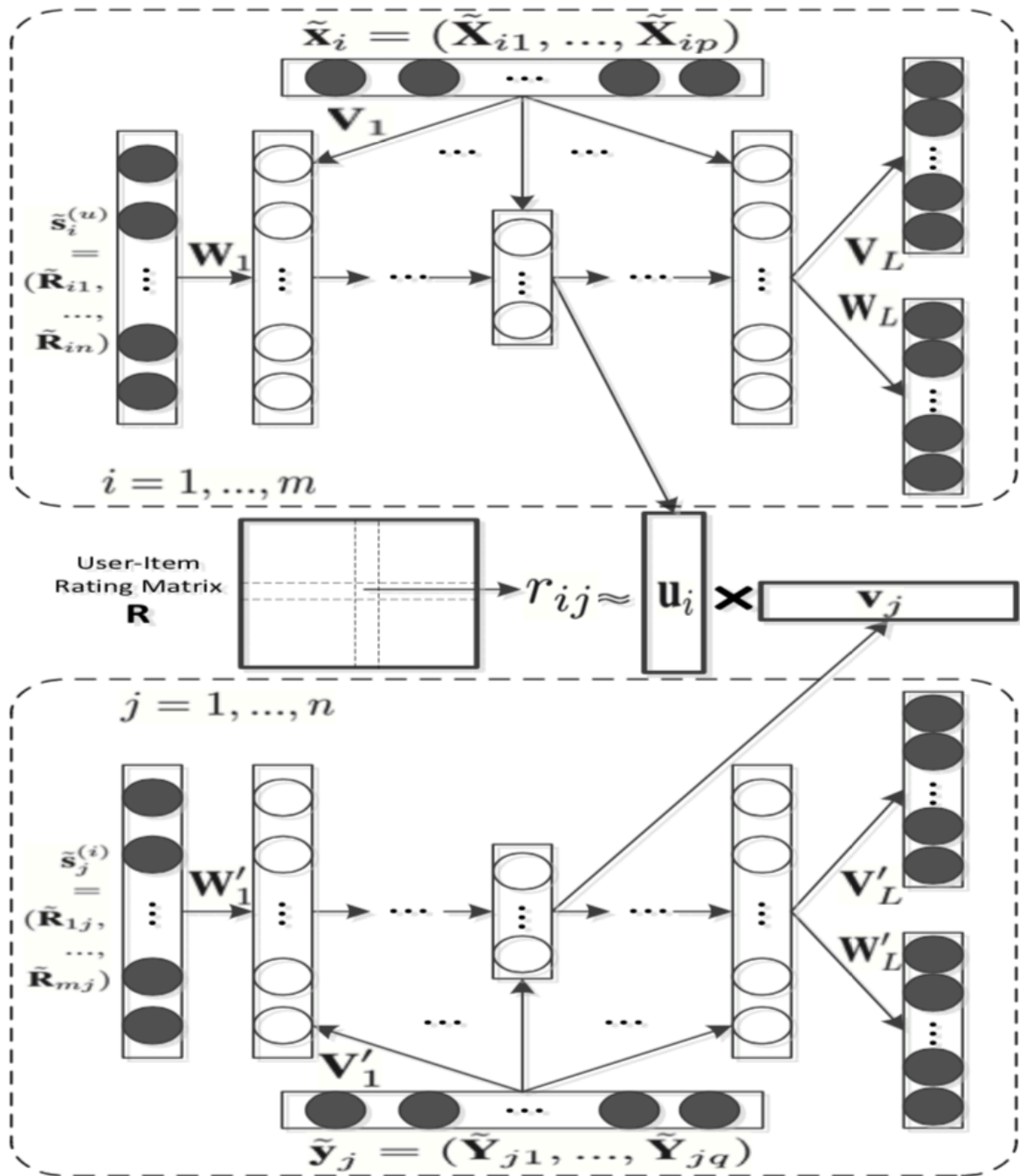
$$\hat{s} = f(W_L h_L + b_{\hat{s}})$$

$$\hat{x} = f(V_L h_L + b_{\hat{x}})$$

我们需要学习的就是 W, V, b 这一系列的参数。注意到, 这里生成的每一层 h 具有的辅助信息输入 \tilde{x} 都是不变的。

Hybrid CF Model

有了上面的aSDAE, 提出了一种混合的协同过滤模型, 即融入了打分信息作为损失函数:



简而言之，aSDAE的左侧 S 输入来自原本的打分矩阵的每一行/列，中间的 $h_{\frac{L}{2}}$ 作为矩阵分解 u_i, v_j 表示，用来计算打分： $\hat{r}_{ij} = u_i^T v_j$

最终的目标，最小化的损失包括三部分：

1. 打分损失： $r_{ij} - u_i^T v_j$
2. aSDAE中自编码解码用户和物品的损失，包括重构 S 和辅助信息 X, Y
3. 参数正则损失

通过梯度更新下降更新参数，优化训练模型

Experiments

Settings

数据集来自 MovieLens 1K/1M 和 Book-Crossing，外部辅助信息应该是自己从外部获取的？（我根本不知道MovieLens里面的id都具体指那一个用户或物品，标准数据集并没有这些）。用户的辅助信息包括年龄、性别、职业、邮政编码等，编码成长度为1943的二进制向量；物品的辅助信息物品名称、发布日期、电影所属类别，编码为长度是1822的二进制向量。

评价指标是RMSE和Recall@K

Results and Analysis

结果是比传统的方法更好。具体看他的分析：

1. 相比经典方法，如PMF，不融合外部辅助信息，显然利用了更过的信息创造了更多的价值，学习到的用户、物品表示更为丰满了
2. CMF方法，它融合的方式过于苍白且简单，其他的深度学习模型，也是简单的特征concatenation等，学习到的用户和物品的隐藏向量表示存在缺陷。
3. （这个是我自己想的，原文没有解释）信息之间不是简单的加在一起，像人一样信息之间的相关，蕴含着推理（如否定、递进、无关的关系）这样的联系或者是过程。

Conclusion

本文的结论十分简单，提出了这样的基于深度学习的混合式协同过滤模型的推荐系统，能获得更好的性能。一些常见的深度模型（AE，CNN，RNN 等）都可以应用于推荐系统中，但是针对不同领域的推荐，可能需要选择合适的模型。他们的未来计划是，探索其他的深度模型（CNN，RNN）来取代提出的aSDAE。

My Last Words

本文由于出自工业界，是一个介绍模型方法，列结果的套路。理论分析似乎有些欠缺，很难理解背后的原因，似乎仅仅通过实验证明了模型的有效性，究竟为什么奏效分析不足，所以我自己想了一点原因，让人怀疑能被AAAI录用的结果（开玩笑，膜拜携程大佬）。

深度学习越来越火，并且烧到了推荐系统领域，并且打败了很多传统的方法。让我看到了一种新的融合方式，比我能想到的与collective matrix factorization，即便是collective neural network matrix factorization都要厉害。想想我自己从review作为side information（即辅助信息），或者是社交关系，也都是可以用aSDAE来做的，但是还是太过幼稚，见识浅陋，比不上这种融合模式。