

基于深度学习的组合推荐算法在电影系统中的应用

顾亚鑫, 武国宁

中国石油大学(北京)理学院, 北京

收稿日期: 2025年4月23日; 录用日期: 2025年5月22日; 发布日期: 2025年5月28日

摘要

如何根据不同用户需求将有效信息第一时间呈现给用户是推荐系统的一个重要研究方向。其中, 基于深度学习的推荐系统不仅能解决信息冗余还能提升个性化推荐效果, 引起了广泛的关注。由于单一的推荐算法往往并不能理想化地满足客户的需求, 本文介绍了一种基于深度学习的组合推荐算法, 对基于内容的推荐算法(CB)、矩阵分解推荐算法(MF)和Wide&Deep模型推荐算法进行优化, 分别通过岭回归、增加隐式反馈信息和贝叶斯优化的优化方法, 并通过转换型的组合方式构成组合推荐算法。实验结果表明, 组合算法在Movielens数据集上的推荐效果优于单一模型, 评估指标均较于单一模型有基本提升, 能够实现精准的内容投放。

关键词

深度学习, 组合推荐算法, Wide&Deep模型, 个性化推荐

Application of Combined Recommendation Algorithm Based on Deep Learning in Movie System

Yaxin Gu, Guoning Wu

College of Science, China University of Petroleum Beijing, Beijing

Received: Apr. 23rd, 2025; accepted: May 22nd, 2025; published: May 28th, 2025

Abstract

How to present effective information to users in a timely manner according to their different needs is an important research direction in recommendation systems. Among them, recommendation

文章引用: 顾亚鑫, 武国宁. 基于深度学习的组合推荐算法在电影系统中的应用[J]. 计算机科学与应用, 2025, 15(5): 703-717. DOI: 10.12677/csa.2025.155142

systems based on deep learning not only solve information redundancy but also improve personalized recommendation performance, which has attracted widespread attention. Due to the fact that a single recommendation algorithm often cannot ideally meet the needs of customers, this article introduces a combination recommendation algorithm based on deep learning, which optimizes content-based recommendation algorithm (CB), matrix factorization recommendation algorithm (MF), and Wide&Deep model recommendation algorithm. The optimization methods include ridge regression, adding implicit feedback information, and Bayesian optimization, and the combination recommendation algorithm is composed of a transformational combination. The experimental results show that the recommendation effect of the combined algorithm on MovieLens dataset is better than that of the single model, and the evaluation indicators are basically improved compared with that of the single model, which can achieve accurate content delivery.

Keywords

Deep Learning, Combined Recommendation Algorithm, Wide&Deep Model, Personalized Recommendation

Copyright © 2025 by author(s) and Hans Publishers Inc.

This work is licensed under the Creative Commons Attribution International License (CC BY 4.0).

<http://creativecommons.org/licenses/by/4.0/>



Open Access

1. 引言

推荐系统可以被看作是一个搜索排名系统, 其中输入是一组用户和上下文信息, 输出是一个排序的物品列表。协同过滤[1][2]是最早和最有影响力的推荐算法之一, 其利用用户与物品之间的交互信息为用户产生推荐, 在各种协同过滤技术中, 矩阵分解(MF) [3] [4]是最流行的一种, 它将用户和物品投影到共享的潜在空间中, 使用潜在特征向量来表示用户或物品。但协同过滤技术存在数据稀疏和冷启动问题, 此外它采用浅层模型无法学习到用户和物品的深层次特征。为了解决这些限制, 基于内容的推荐方法[5]出现了, 利用用户已选择的物品来寻找其他类似属性的物品进行推荐, 主要特征分为用户端信息、物品端信息[6] [7]以及交互信息[8]。但是这种方法需要有效的特征提取, 传统的浅层模型依赖于人工设计特征, 其有效性及可扩展性非常有限, 这制约了基于内容的推荐方法的性能。

近年来, 深度学习基于物品自身可提取到的“可见”内容进行相似度或偏好建模, 无法捕捉更深层次。在图像处理、自然语言理解和语音识别等领域取得突破性进展, 为推荐系统的研究带来了新的机遇[9]-[13]。深度学习的最大优势是能够通过一种通用的端到端的过程学习到数据的特征, 自动获取到数据的高层次表示, 而不依赖于人工设计特征。并且, 它能够建模用户-物品之间的复杂交互[14], 捕捉顺序、时间和上下文等多种信息, 从而在大规模数据中表现出卓越的学习与推断能力。而这恰好可以改善基于内容的推荐方法的性能验证依赖于有效的数据特征提取的缺点。Elkanky 等人[15]在深度结构化语义模型(DSSM) [16]的基础上进行了扩展, 提出多视角深度神经网络(Multi-View DNN), 通过同时对用户和物品两种信息实体进行语义匹配, 有效解决了传统基于内容的推荐系统中难以获取用户特征的问题。Zheng 等人[17]同样借鉴了 DSSM 的框架, 将评论信息融入推荐系统, 以缓解由于用户行为数据稀疏而带来的模型训练困难, 并提出了深度协作神经网络(DeepCoNN)。此外, Cheng 等人[10]在兼顾记忆(memorization)与泛化(generalization)能力的思路下, 将宽度(Wide)与深度(Deep)两种模型架构相结合, 提出了深广学习(Wide&Deep Learning)模型, 从而在多个推荐场景中都取得了良好的效果。类似地, He 等人[18]基于协同过滤方法提出了 NCF 模型(Neural network based Collaborative Filtering), 使用多层感知器(MLP)来取代 MF

和用户交互功能。然而, 深度推荐模型本质上是数据饥渴的, 为了利用深度架构, 需要大量的训练数据。但是, 大多数用户通常只与大量可用物品中的一小部分进行交互[19]。因此, 数据稀疏性问题阻碍深度推荐模型充分发挥其潜力[20], 这也说明了深度学习在推荐系统中还需要进一步改进和优化。

在面对相对复杂的问题时, 单一的推荐算法往往并不能理想化地满足客户的需求, 因此本文围绕如何利用深度学习技术和传统推荐算法相结合的问题展开: 首先, 介绍 MF 算法、CB 算法以及 Wide&Deep 模型推荐算法的基本原理与特点; 接着, 依托在奇异值分解(SVD)基础上增加隐式反馈并结合岭回归(Ridge)与贝叶斯优化算法, 对以上三种推荐算法进行深度分析与性能优化。此后, 通过对数据特征的精细分类和超参数设置, 将不同算法进行“转换式”整合, 形成了组合推荐算法, 并在 ml-latest-small 数据集上进行验证。实验结果表明, 本文提出的组合方法在评分准确性方面明显优于单一模型, 实现了对不同用户的精准内容投放, 为解决信息冗余与个性化推荐需求提供了新的思路与方法。

2. 相关工作

2.1. 协同过滤推荐算法

协同过滤算法(CF)的数据通常表示为一个用户 - 物品评分矩阵, 矩阵中的元素表示用户对物品的评分。算法的关键在于构建用户或物品的相似度矩阵, 图 1 展示了其基本流程。

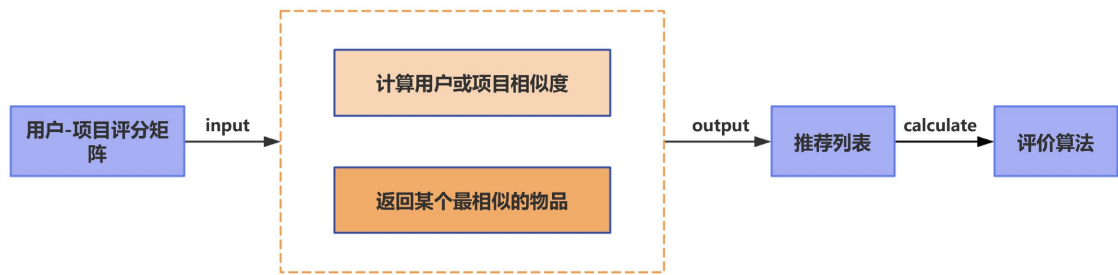


Figure 1. Basic workflow of the collaborative filtering recommendation algorithm

图 1. 协同过滤推荐算法基本的流程

CF 算法的三种类型包括基于用户的协同过滤(UBCF)、基于物品的协同过滤算法(IBCF)、基于模型的 CF 算法[1]。在基于模型的协同过滤算法中核心在于模型的构建, 包括矩阵分解(MF)、聚类模型、基于决策树的模型、贝叶斯网络等[21]。SVD 是经典的 MF 方法, 本文采用 SVD++方法是 SVD 方法的改进, 通过学习用户和物品的隐含特征来优化推荐效果。

原有的 MF 方法的基本思路是通过对稀疏矩阵进行模型分解, 预测得到缺失的空白数据, 将二维用户 - 物品评分矩阵补全完整。即假设 $U = \{u_1, u_2, \dots, u_m\}$ 代表有 m 位用户, $I = \{i_1, i_2, \dots, i_n\}$ 代表有 n 个物品, 所以该评分矩阵可以通过一个 $m \times n$ 的矩阵进行表示, $R = [r_{1,1}, \dots, r_{1,n}; \dots; r_{m,1}, \dots, r_{m,n}]$, 可以表示为:

$$R_{i,j} = \begin{bmatrix} r_{1,1} & r_{1,2} & \cdots & r_{1,n} \\ r_{2,1} & r_{2,2} & \cdots & r_{2,n} \\ \vdots & \vdots & \ddots & \vdots \\ r_{m,1} & r_{m,2} & \cdots & r_{m,n} \end{bmatrix} \quad (1)$$

式中 $R_{i,j}$ 代表第 i 个用户对第 j 个物品的评分。

具体就是找出两个低维的矩阵 P 和 Q , 使其乘积为原来的矩阵 R , 从而获得一个基础的训练模型:

$$P_{m \times n} = P_{m \times k} \times Q_{n \times k}^T \quad (2)$$

而 SVD 是经典的 MF 方法, 常用于通过矩阵分解来提取潜在的特征。在 SVD 中, 评分矩阵 R 通过奇异值分解被近似为三个矩阵的乘积:

$$R \approx U \Sigma V^T \quad (3)$$

式中 U 为用户矩阵, 每行对应一个用户的向量表示; Σ 衡量特征的重要程度; V^T 每列对应一个物品的向量表示。

SVD 的目标是通过最小化损失函数来学习用户和物品的隐向量, 损失函数为:

$$J = \sum_{(i,j) \in \text{Observed}} (r_{ij} - \hat{r}_{ij})^2 + \lambda (\|P_i\|^2 + \|Q_j\|^2) \quad (4)$$

式中 r_{ij} 是用户 i 对物品 j 的实际评分, $\hat{r}_{ij} = P_i Q_j^T$ 是预测评分; λ 是正则化系数, 用于避免过拟合。

传统的 CF 依赖于用户的交互数据, 通过计算用户间或物品间的相似性生成推荐, 其核心问题在于数据稀疏以及新用户和新物品缺乏历史交互信息的冷启动问题时难以进行精准建模和推荐, 例如 MovieLens 数据中用户平均评分仅占矩阵的 5% 以下, 传统方法会导致结果精度不高。为了解决这些问题, MF 算法通过潜在因子分解用户-物品评分矩阵, 捕捉低维隐语义关系, 而 SVD 方法因计算复杂度高, 难以扩展至大规模数据; 后文提到的 SVD++ 通过改进引入隐式反馈, 缓解稀疏性提高推荐质量。

2.2. 基于内容的推荐算法

基于内容的推荐算法(CB)通过分析用户的历史行为、偏好或已交互的内容, 计算用户与不同物品之间的相似性, 从而提供个性化推荐。构建 CB 算法模型通常包括三个关键步骤: 用户偏好特征的创建、不可度量因素的提取以及相似度的计算。根据图 2 和图 3 可以了解到其基本流程。

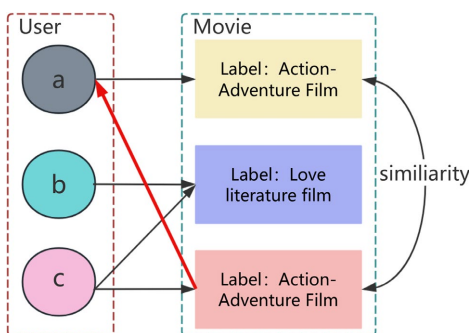


Figure 2. Basic principles of content-based recommendation algorithms

图 2. 基于内容的推荐算法基本原理

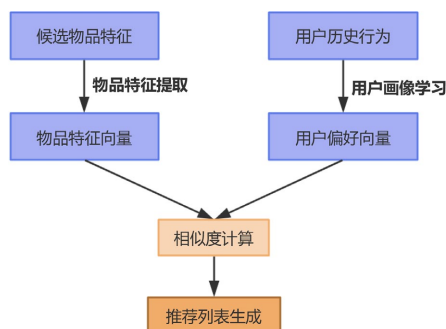


Figure 3. Basic workflow of content-based recommendation algorithm

图 3. 基于内容的推荐算法基本流程

CB 算法中只关注结构化数据, 最重要的步骤就是抽取物品和用户的特征, 需要充分了解用户和物品的特征信息。通过本文电影推荐举例, 除电影的分类、演员、导演等元数据外, 还可以通过运用 TF-IDF (Term Frequency-Inverse Document Frequency) 算法统计出权重较高的词汇, 通过计算特定单词在文档中出现的频率(TF)和在整个语料库中单词的普遍重要的程度(IDF)来量化文本信息的特征, 进行特征提取从而更精确地描述内容特性。

$$tf_{ij} = \frac{n_{ij}}{\sum_k n_{kj}} \text{ 即 } TF_i = \frac{m}{n} \quad (5)$$

式中 m 表示某一类中词条 t 出现的次数, n 表示该类中所有的词条目总和。

$$idf_i = \log \frac{|D|}{|\{j: t_i \in d_j\}|} \text{ 即 } IDF = \log \left(\frac{x}{y+1} \right) \quad (6)$$

式中 x 代表语料库中的文件总数, y 代表包含词条 t 的文件数目。

对公式(5)与公式(6)做进一步处理可以发现:

$$TF-IDF = TF \times IDF \quad (7)$$

由于 CB 算法无需大量用户历史行为数据, 能够在用户数据稀缺的情况下为新用户生成个性化推荐, 有效缓解了物品冷启动问题。但该算法也存在一定局限性, 如难以精准提取用户的深层次特征, 依赖人工特征设计无法有效挖掘用户潜在的、未显现的兴趣点, 仅能依据已有的用户特征做推荐, 推荐结果缺乏新颖性与惊喜感。

2.3. 深度学习推荐算法

深度学习是目前推荐系统中最常用的一种方法, 可以从用户商品中挖掘出潜在的信息, 并从高维数据中提取有效特征, 从而获得更多抽象表达提升推荐的准确性。深度推荐系统主要经历了 Embedding 与 MLP 结合的网络、基于特征组合的网络、用户行为序列建模的网络、融合知识图谱的网络四个阶段[22]。

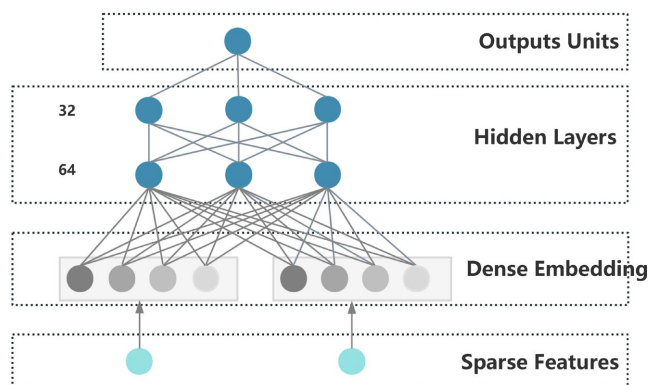


Figure 4. Embedding + MLP architecture
图 4. Embedding + MLP 架构

2016 年, Google 发布的深度学习推荐模式在业内引起了极大的兴趣, 其是建立在 Embedding + MLP 的框架(图 4)基础上, Embedding 核心思想是将高维、稀疏的离散特征映射到一个低维、连续的密集向量空间, 从而使神经网络能够更有效地学习特征之间的语义关系, 提供了很好的离散特征处理能力, 多层感知机(MLP, Multi-Layer Perceptron)擅长建模复杂的非线性关系, MLP 是一种包含多层隐藏层的神经网络。

络，通常由输入层、隐藏层和输出层组成。将 Embedding 与 MLP 结合，不仅能够有效捕捉离散特征的语义信息，还能通过 MLP 的非线性映射能力，实现对数据模式的更精细拟合。相比于传统基于 0-1 离散特征的模型，这种组合方式显著提升了模型的表达能力和预测性能。正因如此，Embedding + MLP 的架构已成为现代深度学习推荐系统的核心框架，并被广泛应用于各类推荐算法。

但其模型训练通常计算复杂度高、超参数众多与训练耗时长，提出了随机搜索、网格搜索、贝叶斯优化等超参数优化方法从而提高效率，提供更优质的用户体验。

2.4. 组合推荐算法

上述算法单一使用均有优缺点，考虑将推荐的多种推荐算法结合起来弥补单一算法的不足，提升系统的鲁棒性和准确性。本文采用转换型的组合方式，流程如图 5 所示，将目标用户分类，根据分类结果将每个用户类别输入到优化后的 CB、MF 以及深度学习 Wide&Deep 模型中进行个性化预测，以提供对其特定测试数据集的预测，最终系统将各个模型的推荐结果进行综合评估，以提供最优的推荐列表，从而提升整体推荐系统的精准度和用户满意度。

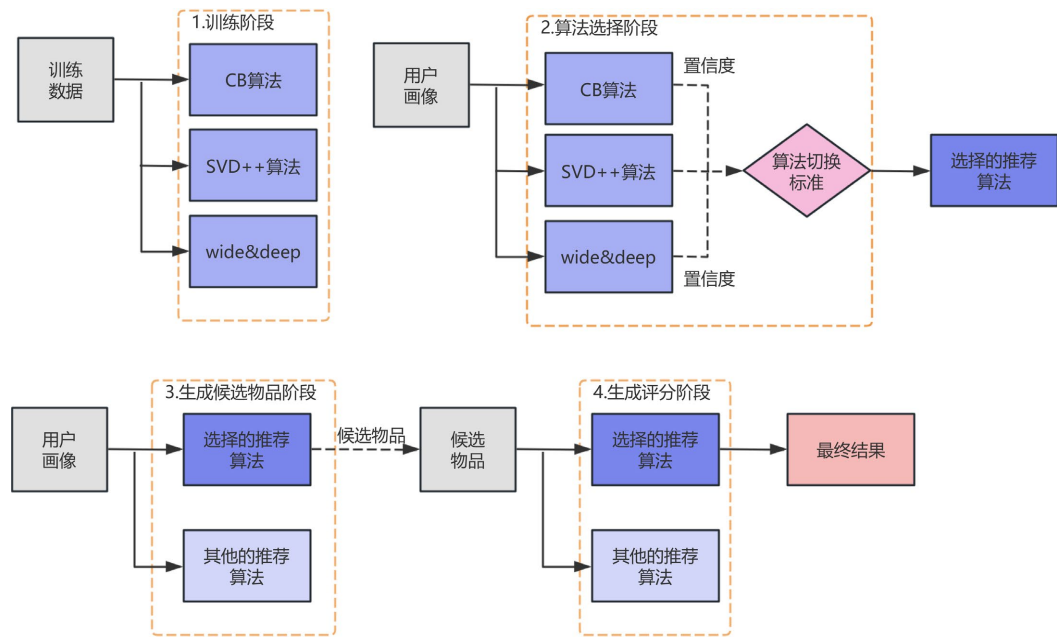


Figure 5. Workflow of the proposed transformer-based combination method
图 5. 本文转换型组合方式流程

3. 算法设计

3.1. 用户分组策略

本文数据集用户电影观看频次呈现典型的长尾分布，后续组合算法会根据数据集分类而采取不同的推荐算法，因此基于用户的评分次数构建用户活跃度的分类标准，对数据集进行新用户、中等用户与活跃用户的区分成为重点。选择在人为划定的分界点基础上进行决策树自动学习边界的分组方法，以保证每个类别的用户分布较为均衡，避免了如活跃用户过少的极端类别情况。相较于传统的人为划定的分界点即三等分数据集的分类方法，引入了决策树模型来自动学习最优分界点，基于信息增益(Gini)选择最佳的分割阈值，使得每个类别的纯度最大化，从而寻找最优的分类策略。

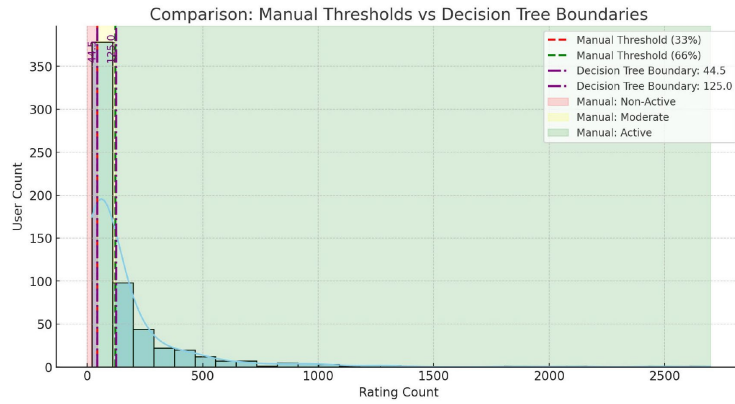


Figure 6. User grouping results
图 6. 用户分组结果

结果如图 6 表明, 决策树学习到的分界点与手动设定的阈值在一定程度上重合, 决策树能够基于数据自动优化分类标准, 比人工划分更具适应性。因此使用决策树分类模型自动学习最佳分界点, 分别为 44.5 与 125 部电影, 为了提高推荐系统的准确率, 避免现实应用中因无观影记录导致推荐质量极低的情况, 筛选用户至少对 5 部电影进行评分, 以提供基础的个性化推荐数据。进而根据用户 u 的观影次数 C_u 得到 $p\%$ 分位数 Q_p 将用户划分为三组。

$$\text{组别} = \begin{cases} \text{新用户} & \text{if } C_u \leq Q_{33}(C) \\ \text{中等用户} & \text{if } Q_{33}(C) < C_u < Q_{66}(C) \\ \text{活跃用户} & \text{if } C_u \geq Q_{66}(C) \end{cases} \quad (8)$$

3.2. 单模型优化方法

3.2.1. CB 模型 - 基于岭回归的特征偏好建模

回归分析是一种重要的预测建模技术, 通过回归模型预测用户对未评分物品的兴趣, 将用户和物品的特征作为输入, 学习用户对不同特征的偏好权重, 从而为用户推荐最符合其兴趣的物品。

基于 TF-IDF 构建电影特征矩阵, 用户偏好建模选择岭回归(Ridge)的方式, 为每一位用户单独训练 Ridge 回归模型, 避免群体偏好混淆, 提升个性化推荐精度。然而面对高维稀疏特征时, 普通线性回归可能会遇到过拟合和共线性问题, Ridge 通过在损失函数中加入了 L2 正则化项, 可以有效减少过拟合并提升泛化能力。

Ridge 回归目标函数是最小化均方误差与 L2 正则项, 通过网格搜索确定正则化强度:

$$J(\omega) = \sum_{i=1}^N (y_i - \omega \cdot x_i)^2 + \alpha \sum_{j=1}^M \omega_j^2 \quad (9)$$

式中 x_i 表示文本特征; y_i 为用户对物品的真实评分; ω 是回归系数; α 是正则化强度。

通过此操作该模型对每个用户都是个性化的, 能够更准确地预测用户对未评分电影的兴趣度, Ridge 回归通过正则化有效抑制高维稀疏特征导致的过拟合, 适用于新用户的小样本学习, 提高了推荐系统的准确性和稳定性。

3.2.2. MF 模型-SVD++隐式反馈融合

本文中用到了 SVD++ 算法来优化损失函数, 是一种基于隐语义模型的协同过滤算法。传统的奇异值分解(SVD)计算复杂度很高, 在实际推荐系统应用中会出现数据稀疏性问题与计算复杂度高的问题。为了

解决 SVD 的局限性, SVD++通过在用户侧增加了对“隐式反馈”信息的建模, 额外为每个用户加入其交互过的物品的影响, 也就是在评分矩阵 R 的基础上自动建模隐式反馈, 构建隐式交互矩阵, 使得显式与隐式反馈联合建模, 从而提高预测的准确性, 大大缓解了冷启动问题。基本原理:

$$\hat{r}_{ui} = \mu + b_u + b_i + \left(u_u + \frac{1}{\sqrt{|N(u)|}} \sum_{j \in N(u)} y_j \right) \cdot v_i^T \quad (10)$$

式中 μ 为全局评分均; $N(u)$ 表示用户 u 交互过的所有物品的集合; y_i 表示表示该物品对用户偏好的影响。

SVD++在构建了目标函数后, SVD++的目标是最小化损失函数(公式(10)), 一般采用随机梯度下降(SGD)和最小二乘法(ALS)的方法作为优化损失函数的方法, 本文中用到了 SGD 算法来优化损失函数, 更快收敛, 最终解在全局最优解附近。

$$J = \sum_{(i,j) \in \text{Observed}} (r_{ij} - \hat{r}_{ij})^2 + \lambda \left(\|P_i\|^2 + \|Q_j\|^2 + \sum_{j \in N(u)} \|y_j\|^2 \right) \quad (11)$$

式中第一项是评分预测的误差; 第二项是正则化项, 防止过拟合。

矩阵分解技术通过将用户和物品的评分矩阵分解为低维隐向量矩阵, 能够有效挖掘出用户和物品之间更深层次的关系, 能够有效地进行评分预测和补全, SVD++相较于传统的 SVD, 在原有的显式评分基础上引入了隐式反馈信息, 显著提高了推荐系统的准确性和鲁棒性, 有效的缓解数据稀疏和冷启动问题。

3.2.3. Wide&Deep 模型 - 贝叶斯优化超参数

Wide&Deep 是在上述 MLP 的基础上加入了 Wide 部分。是基于分类和回归的模型, 是将线性模型(Wide part)与 DNN 模型(Deep part)中的记忆能力与泛化能力结合起来(如图 7 所示)。

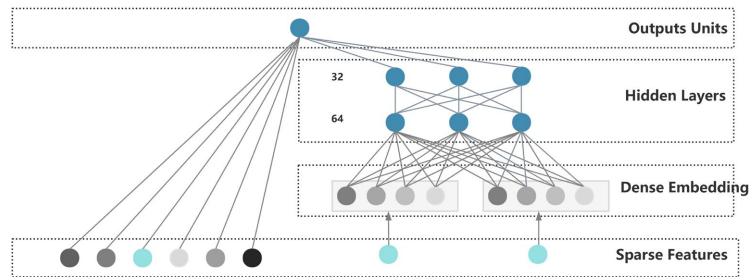


Figure 7. Basic workflow of the Wide&Deep recommendation algorithm
图 7. Wide&Deep 模型推荐算法基本的流程

在 Wide 端中, 输入的特征可以是连续的或稀疏的, 经过交叉后, 可以形成较高维度的离散特征, 针对已有模式的记忆, 有助于记忆特定的特征组合关系, 其经常提到的一种变换用来生成组合特征为:

$$\phi_k(x) = \prod_{i=1}^d x_i^{c_{ki}} \quad (12)$$

式中 d 代表输入 X 的维度, i 代表第 i 维的特征, k 代表第 k 个组合特征, c_{ki} 代表该第 i 维的特征是否参与构造第 k 个组合特征。

对于 Wide 部分的公式可以表示为:

$$y = W_{\text{wide}}^T \{x, \phi(x)\} + b \quad (13)$$

式中 W_{wide}^T 和 b 是模型参数, $\{x, \phi(x)\}$ 是特征集的组合。

Deep 端对应于前馈网络的 DNN 模型, 每一特征都有 embedding, 是一个低维的实数向量, 然后, 将其作为第一个隐藏层的输入并随机地初始化, 通常使用 ReLU 函数来激活隐藏层。通过 Embedding 层把用户 ID、物品 ID 编码成低维连续向量, 连接后再通过多层全连接网络提取更复杂的交互关系, 捕捉非线性特征, 对样本中未出现模式的泛化和模糊查询, 有助于泛化和表达复杂特征交互。前馈部分用公式(14)表示, 但是内嵌模型难以有效地学习低维表达, 因此会因过于泛化而给出与之无关的建议, 准确率无法保证。

$$a^{(l+1)} = f\left(W_{\text{deep}}^T a^{(l)} + b^{(l)}\right) \quad (14)$$

式中 1 代表第一层, $f(\cdot)$ 表示为激活函数, $W_{\text{deep}}^{(l)}$ 和 $b^{(l)}$ 是模型参数。

因此, 将一个线性模型和一个深度神经网络通过以下公式(15)结合起来, 形成了一个将浅层模型与深层次模型相结合的系统结构, 从而获得了 Wide& Deep 模型。

$$P(\hat{r}_{ul} = 1 | x) = \sigma\left(W_{\text{wide}}^T \{x, \phi(x)\} + W_{\text{deep}}^{(l)} a^{(l_f)} + b\right) \quad (15)$$

式中 $\hat{r}_{ul}$ 的输出在 0 到 1 之间, 表示预测值, $a^{(l_f)}$ 表示为最后一层的输出激活值。

在搜索 Wide&Deep 模型的超参数时, 采用了贝叶斯优化的策略。贝叶斯优化通过迭代学习, 根据已有的实验结果选择新的超参数组合, 逐步提升模型的性能。此基础上, 贝叶斯优化通过构建概率模型对目标函数进行建模, 并根据历史实验反馈预测每组超参数的效果使目标函数最小化, 因此贝叶斯优化能够在较少的试验次数中找到接近最优的超参数配置, 且具有更高的搜索效率。

3.3. 组合推荐框架

本文针对推荐系统中用户异质性和行为稀疏性的问题提出了一种基于转换型策略的组合推荐框架(如图 8 所示), 其核心思想是根据用户历史行为的密度分配推荐模型, 并通过多模型输出的置信度加权实现个性化推荐, 有效弥补单一模型的局限性。根据前面的用户分组策略对不同用户组别调用模型具体为以下内容:

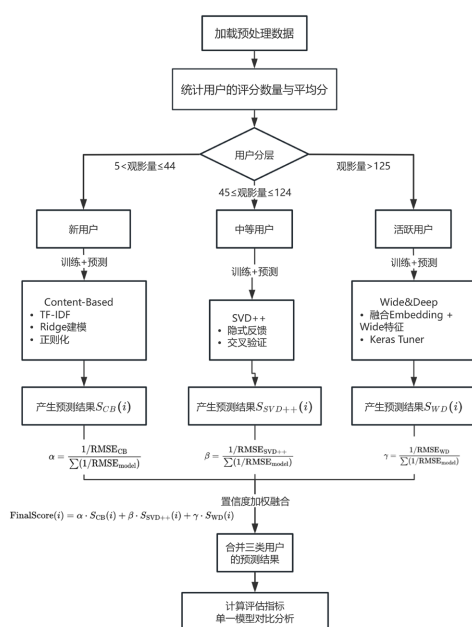


Figure 8. Combined recommendation algorithm flowchart

图 8. 组合算法流程图

对于新用户类别, 选择 CB 模型, 由于该分类的用户及其偏好信息有限, 而 CB 模型无需依赖大量历史数据, 其通过分析电影的特征进行个性化推荐, 规避了需要用户的历史数据的硬性要求, 因此适用。

对于中等活跃用户类别, 选择 SVD++模型, 这种方法不仅能处理用户历史行为数据, 还能借助其他用户的数据进行预测, 从而有效提升准确性。新用户矩阵太过空余无法给出有效的建议, 因此这个模型最适合中等活跃用户。

对于活跃用户类别, 选择 Wide&Deep 模型, 活跃用户观影历史较为丰富, 交互数据较多, 能提供更完整的用户偏好信息, 便于挖掘用户长期兴趣模式, 更好地捕捉复杂的特征交互关系, 以实现更精准的个性化推荐。

对多模型输出进行置信度加权合并, 计算方法:

$$\text{FinalScore}(i) = \alpha \cdot S_{\text{CB}}(i) + \beta \cdot S_{\text{SVD++}}(i) + \gamma \cdot S_{\text{WD}}(i) \quad (16)$$

式中 $S_{\text{model}}(i)$ 为模型对电影 i 的预测评分, 权重通过验证集模型性能动态分配: $\alpha = \frac{1/\text{RMSE}_{\text{CB}}}{\sum (1/\text{RMSE}_{\text{Model}})}$,

$$\beta = \frac{1/\text{RMSE}_{\text{SVD++}}}{\sum (1/\text{RMSE}_{\text{Model}})}, \quad \gamma = \frac{1/\text{RMSE}_{\text{WD}}}{\sum (1/\text{RMSE}_{\text{Model}})}。$$

4. 实验与结果分析

4.1. 数据集与实验设置

本文数据来源为 MovieLens 的数据集中的 ml-latest-small 数据集, 其中数据集包含 610 个用户、9724 部电影以及其评分类别交互信息。数据集中分为四个不同的数据表: Rating 表, 内容总共包含着用户 id, 电影 id, 评分和观看时间; Tags 表, 内容总共包含着每一个用户对于每一个电影的分类; Movie 表, 内容总共包含着电影 id、电影片名和电影类型; Links 表, 内容包含通过网站 id 在对应网站上找到对应的电影链接的方法。

由于在后续问题的研究时, 会存在将电影根据评分排序进而形成推荐列表的情况, 过滤低频用户即观影次数小于 5 次的用户, 避免数据过于稀疏导致训练偏差。为了评估推荐模型的性能, 用户分组按照 4:1 比例划分训练、测试集, 保证模型能够在足够的数据量上进行训练, 并在未见数据上进行评估。根据图 9 可知, 数据的平均评分大致集中在 3.5 到 4 之间, 由于截止值设置为 4 太严格, 会导致大量的信息流失, 因此设定评分阈值为 3.5, 大于视为正向反馈, 以保证推荐质量。

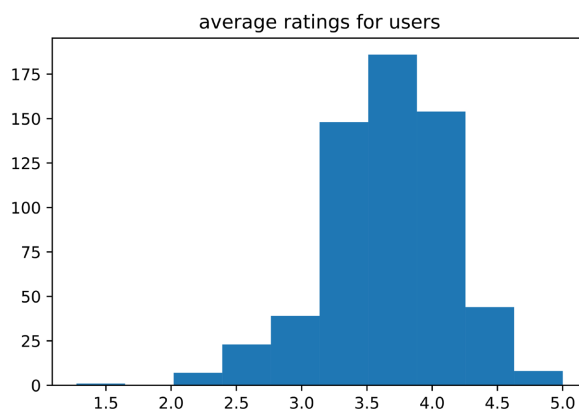


Figure 9. User average rating
图 9. 用户平均评分

在新用户模型中, 通过实验结果可以看出, RMSE 随着 α 因子的增大而显著提高, 由此确定超参数为 $\alpha=1$ 时模型效果最佳。从计算结果中可以看出, 模型的 RMSE 以及准确率仍然较差, 意味着预测评分与用户实际打分的平均误差较大, 以及推荐的电影与用户偏好的相关性较差, 根据上述问题后续还需进一步对模型进行组合优化。

在中等活跃用户模型中, 根据优化过程中计算出的 RMSE 值, 潜在因子数量和正则化参数的优化对模型性能有显著影响。通过调整潜在因子的数量和正则化参数, 得到了最佳的模型超参数配置。如图 10 与图 11 所示, 随着潜在因子数量的增加, RMSE 显著下降, 最终在潜在因子数量为 3、正则化参数为 0.25 时达到了最小, 因此其为最优选择。

在活跃用户模型中, 通过对模型超参数如嵌入维度、密集层数量、dropout 率等的优化, 使用贝叶斯优化不断地调整超参数的组合情况, 寻找 RMSE 最小的超参数组合。最终实验结果配置参数为: 用户嵌入维度为 80, 电影嵌入维度为 50, dropout 率为 0.3, 密集层数量为 2, 学习率为 0.0012。

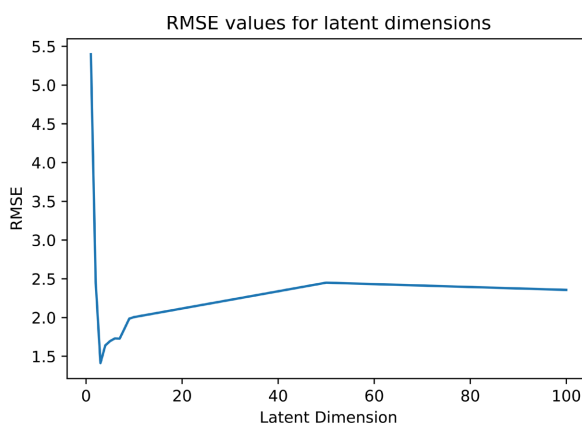


Figure 10. RMSE value for latent dimensions

图 10. 潜在维度的 RMSE 值

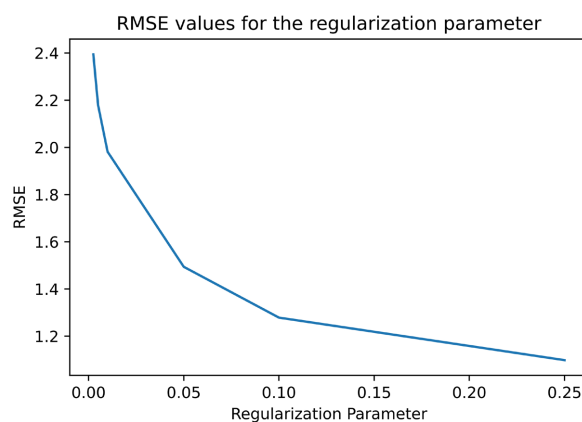


Figure 11. RMSE value for regularization parameters

图 11. 正则化参数的 RMSE 值

4.2. 对比实验设计

本文通过对比基线模型与最终组合模型在多个评估指标下的表现, 将组合模型与基于岭回归的特征偏好建模的 CB 模型、隐式反馈融合的 SVD++模型与贝叶斯优化的 Wide&Deep 模型进行全面的实验对

比。在以下评估指标进行对比分析, 进而验证用户分组策略和置信度加权融合机制的优化方法对于提升推荐系统的整体表现起到作用。

主要评估指标是 RMSE (Root Mean Square Error) 均方根误差, 可以反映推荐系统在预测用户评分时的准确性, 其表达式为公式(17), RMSE 通过平方误差加权了较大偏差的影响, 对较大误差具有较高的敏感度, 通常用来评估预测模型的整体精度。

$$\text{RMSE}(X, h) = \sqrt{\frac{1}{m} \sum_{i=1}^m (h(x_i) - y_i)^2} \quad (17)$$

式中 m 为一次训练选取的样本量。

二级评估指标是召回率和准确率, 是二分类问题常用的指标。其中召回率(Recall_ k)是指推荐系统中能够推荐出的所有相关物品中, 实际被正确推荐的比例; 准确率(Precision_ k)是指在全部返回的结果中, 预测正确的相关结果所占的百分比, 即从所有推荐的内容中衡量真正相关的推荐。其表达式分别为公式(18)与公式(19)。

$$\text{Recall}_k = \frac{TP_K}{TP_K + FN_K} \quad (18)$$

$$\text{Precision}_k = \frac{TP_K}{TP_K + FN_K} \quad (19)$$

式中 k 代表推荐的物品数量。

第三个评估指标是归一化折损累计增益(NDCG), 是一种用于衡量推荐系统排序效果的重要指标, NDCG 考虑了推荐结果中物品的排名信息, 更高排名的物品赋予更高的权重, 以反映出推荐系统在推荐位置的预测能力。

$$\text{NDCG} = \frac{\text{DCG}}{\text{IDCG}} \quad (20)$$

4.3. 结果分析

本文针对不同用户类别实现组合推荐系统, 使用了 RMSE、MAE、Precision_10 (前 10 个推荐的精度)、Recall_10 (前 10 个推荐的召回率)和 NDCG_10 (归一化折损累计增益)指标进行分析。

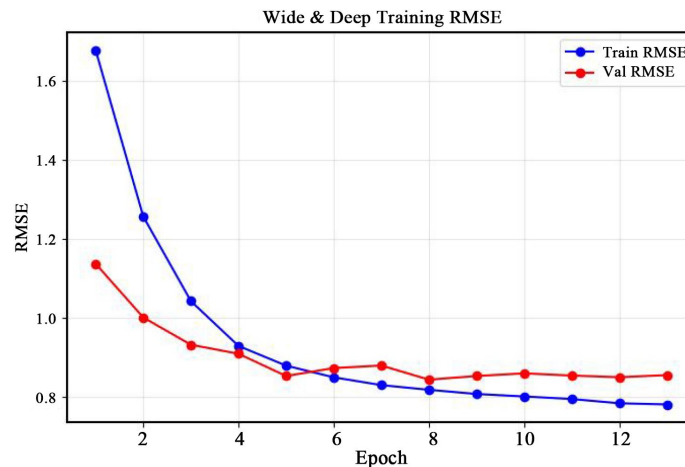


Figure 12. Decreasing trend of RMSE

图 12. RMSE 下降趋势

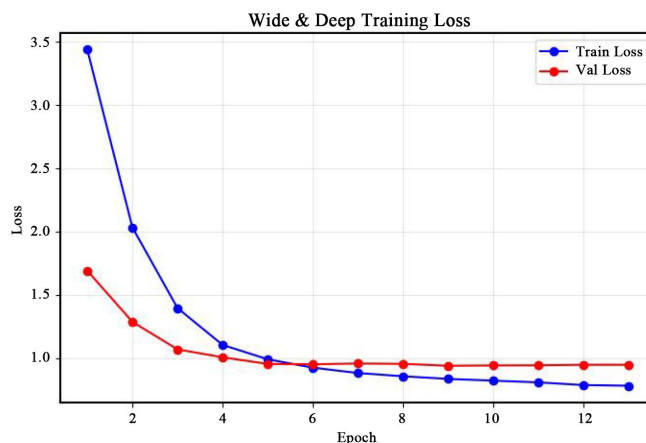


Figure 13. Trend of training and testing loss

图 13. 训练损失与测试损失变化趋势

在训练过程中, 对 Wide&Deep 模型的损失函数(RMSE Loss)和 RMSE 进行了跟踪分析。图 12 展示了训练过程中 RMSE 的变化, 训练 RMSE(蓝色曲线)和测试 RMSE(红色曲线)在前 6 个 epoch 内迅速下降, 表明模型正在有效地学习数据特征, 预测误差快速降低。在 epoch 6 之后, 训练 RMSE 继续缓慢下降, 而测试 RMSE 开始波动并趋于稳定, 训练 RMSE 始终略低于测试 RMSE, 表明模型仍具有一定的泛化能力, 并未出现严重的过拟合。图 13 展示了训练损失(RMSE Loss)的变化, 训练损失(蓝色)和测试损失(红色)在前几个 epoch 内快速下降, 表明模型成功学习到了数据的模式。训练损失最终稳定在 0.38 左右, 而测试损失稳定在 0.43 附近, 训练损失始终低于测试损失, 在 6 个 epoch 之后逐渐贴近, 表明模型的学习趋于稳定, 再一次证明没有明显的过拟合。

从表 1 的实验结果可以明显看出, 组合模型(Combined Model)在多个关键指标上均表现优异, 相较于单一模型的基线, 具有显著的优势。组合模型的 RMSE (0.6864)和 MAE (0.5087)均为最低, 相较于其他单一模型, 组合模型的预测误差最小。这表明, 该模型能够更加准确地预测用户评分, 减少了推荐系统的误差, 提高了用户体验。相较于 SVD++(RMSE = 0.8786), 组合模型的 RMSE 下降了 21.9%, 表明其预测更精准。Wide&Deep 模型的 RMSE 高达 0.9525, 组合模型有效融合了多个方法的优点, 使得预测更加稳健。组合模型在 Precision_10 (0.7794)和 NDCG_10 (0.7947)上达到了最高值, 准确率最高意味着在前 10 个推荐的电影中, 真正符合用户兴趣的电影占比更高, 比单一基线模型 SVD++(0.7553)和 Wide&Deep (0.7262)均有提升。归一化折损累计增益(NDCG)衡量的是推荐结果的排序质量, 组合模型在这一指标上超越了所有单一模型, 表明其推荐的电影不仅是用户喜欢的, 而且排序更符合用户偏好。尽管组合模型的 Recall_10 (0.3808)略低, 但它在 Precision_10 上更优, 这说明组合模型倾向于提供更加精准的推荐, 而非仅仅扩大召回的电影集合。这种精准性的提高对于推荐系统来说尤为关键, 因为用户更倾向于接受高质量的推荐, 而非数量庞大但质量不高的推荐结果。

Table 1. Calculated value of evaluation index

表 1. 评估指标计算值

Model	RMSE	MAE	Precision_10	Recall_10	NDCG_10
Content-Based	0.9061	0.7004	0.6776	0.4429	0.7448
SVD++	0.8786	0.6737	0.7553	0.4867	0.7807
Wide&Deep	0.9525	0.7334	0.7262	0.4730	0.7636
Combined Model	0.6864	0.5087	0.7794	0.3808	0.7947

组合模型在 RMSE、MAE、Precision₁₀ 和 NDCG₁₀ 方面均优于其他单一模型, 展现出了更强的预测准确性、推荐质量以及用户满意度。尽管 Recall₁₀ 略低, 但整体而言, 它在不同用户群体上的个性化推荐能力最强, 真正实现了高效、精准的推荐策略。

5. 结论与展望

为了优化推荐效果, 本文采用了 Ridge 回归优化基于内容的推荐算法, 结合隐式反馈优化矩阵分解算法(SVD++), 并通过贝叶斯优化超参数提升 Wide&Deep 模型性能。实验结果表明, 组合推荐算法在 RMSE、MAE、Precision₁₀ 及 NDCG₁₀ 等评估指标上均优于单一推荐模型。特别是组合模型在 RMSE 指标上的表现显著优于单一模型, 表明该方法在评分预测的准确性具有优势。然而, 实验结果也表明, 尽管组合模型在 Recall₁₀ 值仍有优化空间, 这表明模型在推荐物品的多样性方面仍存在改进的余地。

参考文献

- [1] Chen, T., Yin, H., Ye, G., Huang, Z., Wang, Y. and Wang, M. (2020) Try This Instead: Personalized and Interpretable Substitute Recommendation. *Proceedings of the 43rd International ACM SIGIR Conference on Research and Development in Information Retrieval*, Virtual, 25-30 July 2020, 891-900. <https://doi.org/10.1145/3397271.3401042>
- [2] Cheng, H., Koc, L., Harmsen, J., Shaked, T., Chandra, T., Aradhye, H., et al. (2016) Wide & Deep Learning for Recommender Systems. *Proceedings of the 1st Workshop on Deep Learning for Recommender Systems*, Boston, 15 September 2016, 7-10. <https://doi.org/10.1145/2988450.2988454>
- [3] Covington, P., Adams, J. and Sargin, E. (2016) Deep Neural Networks for Youtube Recommendations. *Proceedings of the 10th ACM Conference on Recommender Systems*, Boston, 15-19 September 2016, 191-198. <https://doi.org/10.1145/2959100.2959190>
- [4] He, X., Liao, L., Zhang, H., Nie, L., Hu, X. and Chua, T. (2017) Neural Collaborative Filtering. *Proceedings of the 26th International Conference on World Wide Web*, Perth, 3-7 April 2017, 173-182. <https://doi.org/10.1145/3038912.3052569>
- [5] He, X., Zhang, H., Kan, M. and Chua, T. (2016) Fast Matrix Factorization for Online Recommendation with Implicit Feedback. *Proceedings of the 39th International ACM SIGIR conference on Research and Development in Information Retrieval*, Pisa, 17-21 July 2016, 549-558. <https://doi.org/10.1145/2911451.2911489>
- [6] Huang, P., He, X., Gao, J., Deng, L., Acero, A. and Heck, L. (2013) Learning Deep Structured Semantic Models for Web Search Using Clickthrough Data. *Proceedings of the 22nd ACM international conference on Information & Knowledge Management*, San Francisco, 27 October-1 November 2013, 2333-2338. <https://doi.org/10.1145/2505515.2505665>
- [7] Koplon, R. and Sontag, E.D. (1997) Using Fourier-Neural Recurrent Networks to Fit Sequential Input/output Data. *Neurocomputing*, **15**, 225-248. [https://doi.org/10.1016/s0925-2312\(97\)00008-8](https://doi.org/10.1016/s0925-2312(97)00008-8)
- [8] Koren, Y. (2008) Factorization Meets the Neighborhood: A Multifaceted Collaborative Filtering Model. *Proceedings of the 14th ACM SIGKDD International Conference on Knowledge Discovery and Data Mining*, Las Vegas, 24-27 August 2008, 426-434. <https://doi.org/10.1145/1401890.1401944>
- [9] Liu, H., Wu, F., Wang, W., Wang, X., Jiao, P., Wu, C., et al. (2019) NRPA: Neural Recommendation with Personalized Attention. *Proceedings of the 42nd International ACM SIGIR Conference on Research and Development in Information Retrieval*, Paris, 21-25 July 2019, 1233-1236. <https://doi.org/10.1145/3331184.3331371>
- [10] Pazzani, M.J. and Billsus, D. (n.d.) Content-based Recommendation Systems. In: Brusilovsky, P., Kobsa, A. and Nejdl, W., Eds., *The Adaptive Web*, Springer, 325-341. https://doi.org/10.1007/978-3-540-72079-9_10
- [11] Sarwar, B.M. (2001) Sparsity, Scalability, and Distribution in Recommender Systems. University of Minnesota.
- [12] Sarwar, B., Karypis, G., Konstan, J. and Riedl, J. (2001) Item-Based Collaborative Filtering Recommendation Algorithms. *Proceedings of the 10th International Conference on World Wide Web*, Hong Kong, 1-5 May 2001, 285-295. <https://doi.org/10.1145/371920.372071>
- [13] Wang, C. and Blei, D.M. (2011) Collaborative Topic Modeling for Recommending Scientific Articles. *Proceedings of the 17th ACM SIGKDD International Conference on Knowledge Discovery and Data Mining*, San Diego, 21-24 August 2011, 448-456. <https://doi.org/10.1145/2020408.2020480>
- [14] Wang, H., Wang, N. and Yeung, D. (2015) Collaborative Deep Learning for Recommender Systems. *Proceedings of the 21st ACM SIGKDD International Conference on Knowledge Discovery and Data Mining*, Sydney, 10-13 August 2015, 1235-1244. <https://doi.org/10.1145/2783258.2783273>
- [15] Wu, C., Ahmed, A., Beutel, A., Smola, A.J. and Jing, H. (2017) Recurrent Recommender Networks. *Proceedings of the*

-
- Tenth ACM International Conference on Web Search and Data Mining*, Cambridge, 6-10 February 2017, 495-503. <https://doi.org/10.1145/3018661.3018689>
- [16] Yin, H., Cui, B., Huang, Z., Wang, W., Wu, X. and Zhou, X. (2015) Joint Modeling of Users' Interests and Mobility Patterns for Point-of-Interest Recommendation. *Proceedings of the 23rd ACM International Conference on Multimedia*, Brisbane, 26-30 October 2015, 819-822. <https://doi.org/10.1145/2733373.2806339>
 - [17] Zhang, F., Yuan, N.J., Lian, D., Xie, X. and Ma, W. (2016) Collaborative Knowledge Base Embedding for Recommender Systems. *Proceedings of the 22nd ACM SIGKDD International Conference on Knowledge Discovery and Data Mining*, San Francisco, 13-17 August 2016, 353-362. <https://doi.org/10.1145/2939672.2939673>
 - [18] Zhang, H., Shen, F., Liu, W., He, X., Luan, H. and Chua, T. (2016) Discrete Collaborative Filtering. *Proceedings of the 39th International ACM SIGIR conference on Research and Development in Information Retrieval*, Pisa, 17-21 July 2016, 325-334. <https://doi.org/10.1145/2911451.2911502>
 - [19] Zhang, J., Lin, Y., Song, Z., *et al.* (2018) Learning Long Term Dependencies via Fourier Recurrent Units. arXiv: 1803.06585. <https://doi.org/10.48550/arXiv.1803.06585>
 - [20] Zhang, S., Yao, L., Sun, A., *et al.* (2019) Deep Learning Based Recommender System: A Survey and New Perspectives. *ACM Computing Surveys (CSUR)*, **52**, 1-38. <https://doi.org/10.1145/3285029>
 - [21] Zheng, L., Noroozi, V. and Yu, P.S. (2017) Joint Deep Modeling of Users and Items Using Reviews for Recommendation. *Proceedings of the Tenth ACM International Conference on Web Search and Data Mining*, Cambridge, 6-10 February 2017, 425-434. <https://doi.org/10.1145/3018661.3018665>
 - [22] Zhou, G., Zhu, X., Song, C., Fan, Y., Zhu, H., Ma, X., *et al.* (2018) Deep Interest Network for Click-Through Rate Prediction. *Proceedings of the 24th ACM SIGKDD International Conference on Knowledge Discovery & Data Mining*, London, 19-23 August 2018, 1059-1068. <https://doi.org/10.1145/3219819.3219823>