

基于用户行为序列特征增强的推荐算法研究

曹天若¹ 李景悦²

1 国防科技大学 长沙 410000

2 湖南师范大学 长沙 410012

(1608647231@qq.com)

摘要 随着互联网的迅猛发展,各种功能的 APP 层出不穷,人们已经可以在互联网上实现各种行为操作,各类商品、新闻、广告等信息流持续不断地产生和传播。与此同时,推荐算法领域的工程师们也在不断收集有用特征来迭代优化算法效果。从早期收集画像特征,演变到用户行为日志和历史行为统计,到目前的用户行为序列特征研究,目前推荐算法领域已取得一套完整的特征工程范式。随着用户的历史行为序列近年来被发现是非常重要的特征。但是,仅凭物品 ID 能获得的语义嵌入非常有限,也无法自动与其他相关信息进行交叉,其应用在算法效果收益方面也非常有限。自 2021 年底以来,语言模型的引入在学术界和工业界的应用已取得显著成果,工程师们在推荐算法领域也进行了一些尝试。文中基于语言模型提出了用户行为序列特征增强推荐算法,借助语言模型的语义分析和逻辑思考能力,采用用户行为序列特征的预训练表示学习来实现特征增强,最终提升推荐算法的模型排序能力。

关键词: 推荐算法;语言模型;序列特征;特征增强

中图分类号 TP181

Research on Recommendation Algorithm Based on Embedding User Behavior Sequence Feature Enhancement

CAO Tianruo¹ and LI Jingyue²

1 National University of Defense Technology, Changsha 410000, China

2 Hunan Normal University, Changsha 410012, China

Abstract With the rapid development of the internet, various functional apps have emerged, allowing people to carry out various activities online. All kinds of goods, news, advertisements, and other information continue to be generated and spread on the internet. At the same time, engineers in the field of recommendation algorithms are constantly collecting useful features to iteratively optimize the algorithm's effectiveness. From early collection of portrait features, evolving to user behavior logs, historical behavior statistics, and the current research on user behavior sequence features, the field of recommendation algorithms has now established a complete feature engineering paradigm. In recent years, it has been discovered that the historical behavior sequence of users is a very important feature. However, the semantic embeddings that can be obtained solely by item IDs are very limited and cannot automatically intersect with other related information, resulting in limited benefits in algorithm effectiveness. Since last year, the introduction of language models has achieved significant results in both the academic and industrial spheres, and engineers in the field of recommendation algorithms have also made some attempts. We propose a recommendation algorithm enhancement based on language models to enhance user behavior sequence features. By leveraging the semantic analysis and logical reasoning ability of language models, pre-training representation learning of user behavior sequence features is used to achieve feature enhancement, ultimately improving the effectiveness of recommendation algorithm models.

Keywords Recommendation algorithm, Language model, Sequence characteristics, Feature augmentation

随着互联网的蓬勃发展,以及各种网页、计算机应用、移动互联网终端和应用的普及,人与终端设备之间的信息交互日益频繁,信息爆炸现象日益明显。互联网是一个无所不包、无所不含的信息汇聚体,而互联网用户又很难在浩如烟海的信息海洋中快速找寻到自己需要和感兴趣的信息;网上的信息泛滥,存在大量用户不需要和不感兴趣的信息,用户如果单纯地使用搜索引擎检索信息,同一个关键字得到的结果几乎是相同的。然而,用户对信息的需求是多元化和个性化的,信

息及其传播是多样化的。于是,各大互联网公司纷纷通过收集画像信息和用户行为特征信息让推荐系统个性化地为用户推送内容、商品或服务,以满足用户个性化需求来提高用户参与度。这些推荐系统目前在电商、社交媒体、短视频和新闻平台上得到了广泛应用。

在现实世界中,用户在很多场景中的行为是序列化的,很多信息也呈现出序列相关性,这种序列数据在用户行为分析中也得到了极大的体现。例如:在电商平台中,通过用户成交

数据可以分析出用户存在长期和短期两种不同模式的兴趣偏好,并且存在一定的差异;在社交媒体中,比如微博数据往往是跟随实时话题以时间线的方式呈现,用户在当前时刻产生的微博内容往往在主题和情感上有一定的重叠。又如云学习平台,学生学习数据往往以周期形式变化,学生在不同学习阶段的学习内容和知识水平也会呈现周期性和问题相似性。因此,序列信息成为用户行为分析中十分重要的影响因素。

序列模式挖掘最早是由 Agrawal 等^[1]在 1995 年针对超市购物篮数据的分析中提出的。序列模式挖掘的目标是要找出序列数据库中所有超过最小支持度阈值的序列模式。最早的序列模式挖掘算法大多基于关联规则挖掘算法 Apriori。近年来,在序列模式挖掘的基础上,人们又提出了很多基于机器学习、数据挖掘技术的序列建模方法,如张量分解^[2]、隐马尔可夫^[3]、动态贝叶斯网^[4]和循环神经网络^[5]等。结合各个领域数据的独有特征,用户序列行为分析在不同的算法领域大放异彩。

为了方便持续进行领域研究,学术界和工业界联合将深度学习框架化和工业化。然而,深度学习的推广并不顺利,始终受限于计算速度,瓶颈也是算力不够。在框架化的早期,工业界对推荐系统的延迟要求极高,研究的推荐算法基本只适用于学术研究。因此,早期人们依然是使用简单的特征进行开发实验,不敢引入序列特征等复杂特征。随着近年来 GPU 等计算机设备的发展,目前也基本不受复杂特征限制,甚至在输入多模态特征的情况下也能实现较低的延时。近年来,随着深度学习的逐渐普及,自从谷歌 Google App Store 的推荐团队发表了 Wide and Deep 模型,使得国内推荐算法迎来了新的变革。传统深度学习做推荐就是将稀疏特征进行表示学习转变成稠密向量输入到神经网络。国内头部电商公司阿里巴巴率先引入序列模式研究,DIN^[6]引入注意力机制,捕获候选商品和用户浏览过的商品之间的兴趣。DIEN^[7]在 DIN 基础上引入序列概念,将用户历史行为序列纳入到网络内。DSIN^[8]将行为序列按照会话时间分割,以精细地刻画用户行为。以上算法都是基于 Wide And Deep 模型进行序列模式的研究、改进和提升。本文也是在算法基础上进行创新改进和对比实验。

随着 2023 年大语言模型(下称大模型)的横空出世,大模型在自然语言处理、计算机视觉等方面有着诸多成果。目前也有工程师正在研究将大模型应用于推荐算法之中,以发挥一些人工所不能解决的问题。本文提出了一种利用大模型自监督学习能力辅助优化序列特征预训练的推荐算法思路,最终实现推荐算法的效果提升。

1 问题定义

用户序列行为数据可以表示为一个 n 维的有限集 $\{x_1, y_1\}, \{x_2, y_2\}, \dots, \{x_n, y_n\}$,有限集合是根据事件发生的时间先后排列起来。例如,当某一个用户浏览电影网站时,产生的序列行为可表示为 $\{(a, click), (a, like), (b, click), (e, click), (b, like), (b, watch), (e, like)\}$ 。其中 a, b, e 代表不同的电影。 $click, like, watch$ 代表用户的行为动作:点击、收藏、观看。将 $\{x_n, y_n\}$ 单条样本成为事实之前的历史行为序列作为独立特征,经过特定的特征工程处理或者网络结构设置,输入

推荐算法模型。最后,通过对比不同推荐算法模型在指标上的差异,确定推荐算法的性能

2 相关研究

2.1 Wide and Deep 模型

此前人们对前馈神经网络(以下简称 DNN^[9])研究和应用已经非常深入和成熟,其主要原理是在单层神经网络模型中,增加了多层隐藏层来实现类似提炼功能,实验表明 DNN 模型具备较强的泛化能力。随后,Cheng 等^[10]通过分析后台收集的用户数据发现数据中存在一些有效的关联关系,将这种关联关系定义为网络的记忆能力,并将具备记忆能力的线性单元与 DNN 单元的特征表示能力相结合提出了 Wide and Deep 模型,其框架如图 1 所示。

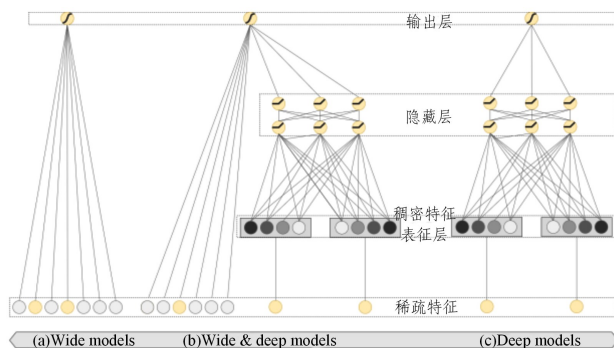


图 1 模型广义深度融合

Fig. 1 Wide and deep fusion model

如图 1 所示,图 1(a)所示的 Wide models 是广义线性模型。线性模型非常简单,特征值通过输入线性加权求和并添加模型偏置。模型训练的基本目的是为了获得对某些规则的记忆,也即是为特征列赋予合理的权重分。“记忆能力”可以被理解为模型直接学习并利用历史数据中物品或者特征的“共现频率”的能力。一般来说,协同过滤、逻辑回归等简单模型有较强的“记忆能力”。由于这类模型的结构简单,原始数据往往可以直接影响推荐结果,产生类似于“如果点击过 A,就推荐 B”这类规则式的推荐,这就相当于模型直接记住了历史数据的分布特点,并利用这些记忆进行推荐。其计算方法如下:

$$y = w^T x + b \quad (1)$$

其中, x 是一个 n 维的行向量; w^T 是一个 n 维的列向量, w^T 为向量转置操作; b 是一个标量,表示偏移量; y 是最终输出。原文中,线性模型的输入除了原始特征,还有转换特征。转换公式如下:

$$Q_k(x) = \prod_{i=1}^d x_i^{c_{ki}} \quad c_{ki} \in \{0, 1\} \quad (2)$$

其中, k 表示第 k 个变换特征, i 表示输入原始特征集合 X 的第 i 个特征, d 表示 X 的维度。当原始特征 x_i 是第 k 个变换特征的一部分时, $c_{ki} = 1$, 反之为 0。

图 1(c)所示的 Deep Models 部分就是前馈神经网络模型。对于高维稀疏的分类特征,首先会转化成低维的稠密的向量,然后作为神经网络隐藏层的输入进行训练。

$$a^{(t+1)} = f(w^{(t)} a^{(t)} + b^{(t)}) \quad (3)$$

式(3)是一个神经网络多层迭代公式, f 是激活函数,如 ReLU, $a^{(t)}$ 是上一层隐藏层的输出, $w^{(t)}$ 是训练参数, $b^{(t)}$ 是模型偏置。

Wide and Deep 模型是通过加权求和的方式将 Wide 侧和 Deep 侧输出组合起来,然后通过 logistic loss 函数联合起来训练。对于 Wide 的部分,一般采用 FTRL^[11] 优化器进行训练。对于 Deep 的部分原文采用 Adagrad 优化器进行训练。模型预测计算式如下:

$$p(y=1|x)=\sigma(w_{wide}^T[x,\mathcal{O}(x)]+w_{deep}^T a^{(f)}+b) \quad (4)$$

其中, Wide 部分特征输入包括原始特征和转换特征, $w_{deep}^T a^{(f)}+b$ 是 DNN 最后一层输出逻辑单元,两者逻辑单元输出之和作为最终输出, $\sigma(\cdot)$ 是 sigmoid 激活函数,转为预测分值, $p(y=1|x)\in[0,1]$ 。

2.2 历史行为序列特征

近几年,随着计算机显卡算力的快速发展,算法的输入从简单的数值特征和类别特征,扩充到 ID 类特征和 ID 类序列特征^[12]。随着自然语言处理(下称 NLP)领域的发展,推荐算法领域的研究人员也借鉴了 NLP 领域的建模思路,挖掘出一些重要的序列特征,这也引发了人们开始对时序算法的研究。他们发现用户的行为序列^[13]具有很高的特征重要性^[14]。但是,这些特征在深度模型输入中又分为随机 embedding 输入和有预训练 embedding 输入。预训练 embedding 输入这种方式如熟知的 word2vec^[15],将用户浏览的商品序列作为单词,进行预先训练。工程实验表明,预训练可以加快推荐算法模型的训练和收敛。

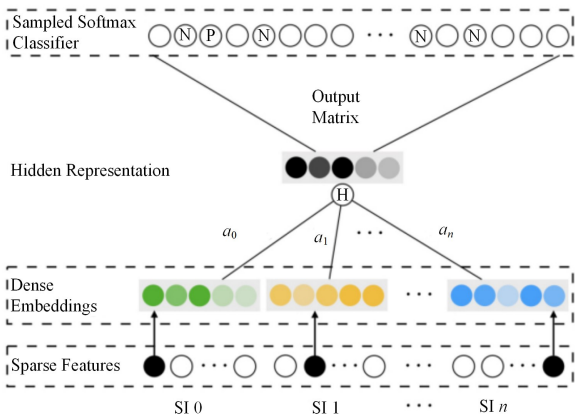


图 2 EGES 框架图

Fig. 2 Framework diagram of EGES

Wang 等^[16]提出了 EGES 模型,将商品 ID 和商品的附加信息联合用户的操作进行列表分割,通过 self-attention 机制对商品特征信息加权池化(下称 pooling),然后用 word2vec 方式实现特征表示的预训练。

图 2 为 EGES 框架图,SI 表示 side info,即附加信息。SI 0 则表示商品自身,即商品 ID。

$$H_v = \frac{\sum_{j=0}^n e^{a_j^v} w_j^v}{\sum_{j=0}^n e^{a_j^v}} \quad (5)$$

其中, a_j^v 表示注意力权重,即商品 v 第 j 个特征的贡献值,表示为 $e^{a_j^v}$ 主要是为了权重大于 0。 $\sum_{j=0}^n e^{a_j^v}$ 则是为了求和方便归一化。整个式子是注意力加权 pooling 方式,计算出最终的商品的最终表示向量 H_v ,该向量整体作为特征输入点击率模型做初始化,也可以使用对应的特征向量 w_j^v 作为模型的初始化,可操作性很强。

2.3 语言模型文本表示学习

通用人工智能迅猛发展,人们的学习、生活、工作产生了巨大变化。通用人工智能在很多方面能够帮助到我们,如创作文案、创作图像,甚至做一些繁杂的科研提炼和筛选工作都起到了很好的效果。例如 OpenAI 基金会的 chatgpt, text-embedding-ada-002,但这些模型目前仅商用,模型训练的细节暂未公布。学术界研究人员更倾向于开源模型,如最近开源大模型新秀 Wang 等^[17]提出的 mistral-7b 是目前学术和开源效果显著的模型,e5-mistral-7b-instruct 则是一个开源的 embedding 模型,可以将任意语言转换为 4 096 维度的向量。2023 年底,开源界大模型代表公司 jina 发布了 jinaai/jina-embeddings-v2-base-en^[18],它是世界上首个与 OpenAI 的专有模型 text-embedding-ada-002 相匹配的开源 8K 文本嵌入模型。jina 发布的该模型可以将英语转换为 768 维度的向量。该模型适用于处理长文档的多种用例,包括长文档检索、语义文本相似度、文本重排序、推荐、RAG 和 LLM 基于生成式搜索等。本文选择该模型作为推荐算法特征增强辅助工具,利用其逻辑思考能力和语义分析能力来提升特征预训练表示学习的效果。将采用式(6)生成特征表示向量。

$$H_v = \text{text_embedding}(t_v) \quad (6)$$

其中, text_embedding 是大模型文本词嵌入模型, t_v 是物品 v 对应的序列特征, H_v 是物品 v 输出的表示向量,它的维度为 n 维。

3 基于语言模型序列特征增强推荐算法

本文借助大模型在文本处理、文本总结、逻辑思考和联想等方面具有突出的能力,优化行为序列特征预训练过程,提出了基于语言模型序列特征增强推荐算法。算法的框架如图 3 所示。

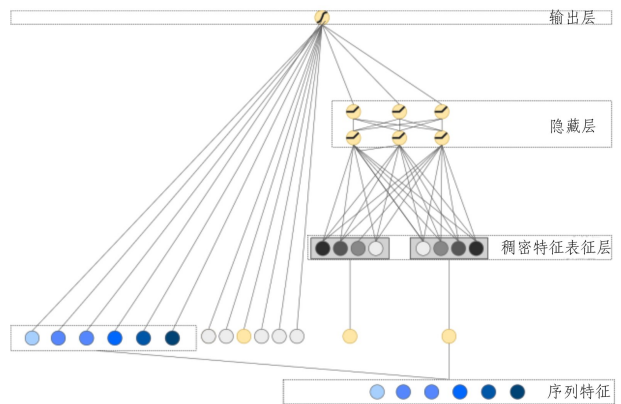


图 3 基于语言模型序列特征增强推荐算法

Fig. 3 Recommendation algorithm based on language model sequence feature enhancement

序列特征经过两种特征增强方式输入到 Wide and Deep 模型。1)序列特征是物品 ID 时间顺序组成的列表,将序列单元(即物品 ID)字符串类型按字典顺序排序,得到对应序列单元 w 的索引 i 。 n 维序列特征 $[w_1, w_2, \dots, w_n]$ 经过转换变成 $[i_{w_1}, i_{w_2}, \dots, i_{w_n}]$,为了转换为 Wide 侧输入,也就是对应索引取值 1,其他为 0。2)序列特征 W 通过结构为任务 $task$, 查询 $query$, 结果 $document$ 的序列单元 $[w_1, w_2, \dots, w_n]$ 构造语言模型的对话。

Task:你是推荐系统方面的专家,请帮忙揭示下面用户对商品 ID 兴趣的关系?

query: 用户浏览过: $[a, b]$ 。

Document: 用户对 e 也十分感兴趣。

$$H_w = \text{llm}(\text{task} = \text{Task}, \text{query} = \text{Query}, \text{document} = \text{Document}) \quad (6)$$

通过将用户行为序列 $W = [a, b]$ 和目标商品 e 构造成语言模型上下文对话关系, 并增加用户画像信息, 来获得增强的序列特征预训练表示向量 H_w , 最后将其输入到 DNN 侧。经过实验证明, 通过以上两种方式的序列特征处理, 引入了语言模型的上下文分析处理能力, 本文提出的基于语言模型序列特征增强推荐算法极大地提高了算法的准确度。

4 实验

4.1 实验数据

本文实验中使用数据集为公开数据集 MovieLens 100 K, 这个已经成为研究推荐系统和协同过滤算法的基准数据集之一, 因为它方便易用、具有丰富的信息和足够的规模, 可以用于评估各种推荐算法的性能。MovieLens 100 K 数据集包含 100 000 个评分记录, 一共有 943 个用户对 1 682 个项目的评分, 所有评分值范围为 $[1, 5]$, 评分均为整数。每个用户至少对 20 部电影进行评分, 评分越高表示用户越喜欢这部电影。但是, 本文采用时序的方式进行训练, 训练的目标是用户的评分。

表 1 MovieLens 数据描述

Table 1 Description of MovieLens datasets

数据集	100 K	97 171
用户数	943	943
物品数	1 682	1 682
交互次数	100 000	100 000
用户平均交互数	106.4	106.4
物品平均交互数	59.5	59.5
稀疏度	93.7%	93.7%

因为算法设计需要, 行为序列特征的最小长度为 2, 因此数据集经过处理略微有所损失, 剩余 97 171 条评分记录, 占原始数据集的 97%。

4.2 评价指标与对比算法

4.2.1 平均绝对误差 (Mean Absolute Error)

均方误差 MAE (Mean Absolute Error) 又被称为 L1 范数, 取值范围为 $[0, +\infty)$, 当预测值与真实值完全吻合时等于 0, 即完美模型; 误差越大, 该值就越大。MAE 的值越小, 说明预测模型拥有更好的精确度。其计算式如下:

$$MAE = \frac{1}{m} \sum_{i=1}^m |y_i - f(x_i)| \quad (7)$$

其中, m 是样本数, y_i 是真实值, $f(x_i)$ 是预测值。推荐算法模型预测的是 0 到 5, 共 5 个分类, 但是为了清楚了解算法的预测准确度, 将用户的评分和预测的评分计算 MAE 整体误差。

4.2.2 精确率 (precision)

表 2 列出了检测结果情况。

表 2 检测结果情况表

Table 2 Description of prediction

检测结果	英文缩写	含义
真阳性	TP	正确地检测到阳性结果
假阳性	FP	错误地检测到阳性结果
真阴性	TN	正确地检测到阴性结果
假阴性	FN	错误地检测到阴性结果

如式 (8) 所示, 精确率定义为预测为正的样本中有多少是真正的正样本, 它是针对我们预测结果而言的。Precision 又称为查准率。本文因为是多分类模型, 采用了统一用宏平均 Macro-average 方式计算精确率、召回率和 F1-score。

$$\text{precision} = \frac{TP}{TP + FP} \quad (8)$$

4.2.3 召回率 (recall)

如式 (9) 所示, 召回率定义为样本中的正例有多少被预测正确, 它是针对我们原来的样本而言的。Recall 又称为查全率。

$$\text{recall} = \frac{TP}{TP + FN} \quad (9)$$

4.2.4 F1-score

F1 值是精确率和召回率的加权调和平均数, 精确率和召回率都是越高越好, 但两者往往是矛盾的。因此常用 F1-score 来综合评价分类器的效果, 它的取值范围为 0 到 1, 越接近 1 效果越好。

$$F1\text{-score} = \frac{2 * \text{precision} * \text{recall}}{\text{precision} + \text{recall}} \quad (10)$$

4.2.5 对比算法

基线模型为 Wide and Deep 模型, 在 Wide 侧和 Deep 侧均加入序列特征, Wide 侧按照序列特征增强方式 1) 加入, Deep 侧则是采用序列单元随机 embedding, 并通过 mean pooling 方式输入到 DNN 中。实验 A 组: 为 Wide and Deep 模型的 Wide 侧和基线组一致, Deep 侧采用了语言模型简单增强序列特征, 即仅和序列特征增强方式 2) 序列单元 (物品 ID) 给语言模型分析上下文, 输出序列特征表示学习向量。实验 B 组: 为 Wide and Deep 模型的 Wide 侧和基线组一致, Deep 侧如序列特征增强方式 2), 但是增加了附加信息。利用语言模型分析行为序列、用户画像、物品信息融合分析上下文, 输出序列特征表示学习向量。3 组算法分别在以上 4 类评价指标上进行评估比较。

4.3 实验环境

本文在华为笔记本 Matebook X Pro 上完成对比实验。笔记本配置如表 3 所列。

表 3 实验笔记本配置列表

Table 3 Experiments computer system info

指标类型	指标值
CPU 型号	12th Gen Intel ^(R) Core ^(TM) i7-1260P 2.10 GHz
内存	16 GB
操作系统	Windows 11 家庭中文版 22H2 64 位
运行环境	Window 11 wsl ubuntu 22.04 子系统
编译环境	Python 3.8.18
编译 IDE	PyCharm 2023.1.2

4.4 实验参数

本文使用的 Wide and Deep 模型的模型设置完全一致, Wide 侧特征输入是用户行为序列经过方式 1) 生成与 movie id 数量一致的 1 682 维的 multi-hot 向量。Deep 侧特征输入是 62 个连续型特征和经过方式 2) 生成的行为序列特征表示学习向量, 向量维度是 768, 总共是 830 维, DNN 隐藏层神经元个数为 $\{256, 64, 16\}$, 因此 DNN 隐藏层输出是 16 维的向量。经过参数调优, 实验最终的 learn_rate 设置为 3×10^{-3} , epoch 为 50 轮, 采用交叉熵 loss 评估, 用 SGD 优化器来优化

模型参数。本文采用了随机数种子,保证实验结果能够复现。

如表 4 所列,实验组 A 的评价指标大幅优于基线模型,实验组 B 的参数在精确率、F1-score 和 MAE 上大幅优于实验组 A,召回率略低于实验组 A。因此,通过实验证明,本文提出基于语言模型对序列特征有较强的增强作用,并且在将用户画像、商品信息和行为序列融合的实验组 B 效果更佳。实验表明,语言模型对行为序列有一定的分析能力,提高了行为序列预训练表示学习快速收敛,不仅可以加快模型参数学习,而且提高了推荐算法的准确率,降低了预测误差。

表 4 实验效果评价

Table 4 Metrics of experiments

	Precision	Recall	F1-score	MAE
基线模型	0.1690	0.1674	0.1682	1.1742
实验组 A	0.1699	0.1679	0.1689	1.1263
实验组 B	0.1708	0.1678	0.1693	1.1209

结束语 本文提出了一种基于语言模型序列特征增强推荐算法。本文主要贡献包括:1)充分利用用户的历史行为序列特征在用户之间的相互影响,提高推荐算法排序的准确性;2)综合用户属性特征、电影描述特征、用户历史观看行为序列,提出了基于语言模型序列特征增强推荐算法,在保证个性化的同时考虑用户历史行为信息和上下文之间的交叉。在 movielens 100K 数据集上的实验结果表明,与普通的 Wide and Deep 深度学习方法相比,基于语言模型序列特征增强推荐算法能有效推荐算法的排序准确性。然而,由于用户行为序列因为时间一定会不断变长,未来考虑将语言模型与推荐系统相结合,通过语言模型不断循环迭代用户历史行为序列特征预训练表征,让行为序列不断处于动态计算之中,以此优化推荐系统的排序准确性。

参考文献

- [1] AGRAWAL R, SRIKANT R. Mining sequential patterns[C]// Proceedings of the Eleventh International Conference on Data Engineering. IEEE, 1995:3-14.
- [2] KULESHOV V, CHAGANTY A, LIANG P. Tensor factorization via matrix factorization[C]// Artificial Intelligence and Statistics. PMLR, 2015:507-516.
- [3] ZUCCHINI W, MACDONALD I L. Hidden Markov models for time series: an introduction using R[M]. Chapman and Hall/CRC, 2009.
- [4] CHAPELLE O, ZHANG Y. A dynamic bayesian network click model for web search ranking[C]// Proceedings of the 18th International Conference on World Wide Web. 2009:1-10.
- [5] TAN Y K, XU X, LIU Y. Improved recurrent neural networks for session-based recommendations[C]// Proceedings of the 1st Workshop on Deep Learning for Recommender Systems. 2016:17-22.
- [6] ZHOU G, ZHU X, SONG C, et al. Deep interest network for click-through rate prediction[C]// Proceedings of the 24th ACM SIGKDD International Conference on Knowledge Discovery & Data mining. 2018:1059-1068.
- [7] ZHOU G, MOU N, FAN Y, et al. Deep interest evolution network for click-through rate prediction[C]// Proceedings of the AAAI Conference on Artificial Intelligence. 2019, 33(1):5941-5948.
- [8] FENG Y, LV F, SHEN W, et al. Deep session interest network for click-through rate prediction[J]. arXiv:1905.06482, 2019.
- [9] ZHANG J, ZHENG Y, QI D, et al. DNN-based prediction model for spatio-temporal data[C]// Proceedings of the 24th ACM SIGSPATIAL International Conference on Advances in Geographic Information Systems. 2016:1-4.
- [10] CHENG H T, KOC L, HARMSSEN J, et al. Wide & deep learning for recommender systems[C]// Proceedings of the 1st Workshop on Deep Learning for Recommender Systems. 2016:7-10.
- [11] MCMAHAN H B, HOLT G, SCULLEY D, et al. Ad click prediction: a view from the trenches[C]// Proceedings of the 19th ACM SIGKDD International Conference on Knowledge Discovery and Data Mining. 2013:1222-1230.
- [12] COVINGTON P, ADAMS J, SARGIN E. Deep neural networks for youtube recommendations[C]// Proceedings of the 10th ACM Conference on Recommender Systems. 2016:191-198.
- [13] SUN G F, WU L, LIU Q, et al. Recommendations Based on Collaborative Filtering by Exploiting Sequential Behaviors[J]. Journal of Software, 2013, 24(11):2721-2733.
- [14] PI Q, BIAN W, ZHOU G, et al. Practice on long sequential user behavior modeling for click-through rate prediction[C]// Proceedings of the 25th ACM SIGKDD International Conference on Knowledge Discovery & Data Mining. 2019:2671-2679.
- [15] MIKOLOV T, CHEN K, CORRADO G, et al. Efficient estimation of word representations in vector space[J]. arXiv:1301.3781, 2013.
- [16] WANG J, HUANG P, ZHAO H, et al. Billion-scale commodity embedding for e-commerce recommendation in alibaba[C]// Proceedings of the 24th ACM SIGKDD International Conference on Knowledge Discovery & Data Mining. 2018:839-848.
- [17] WANG L, YANG N, HUANG X, et al. Improving text embeddings with large language models[J]. arXiv:2401.00368, 2023.
- [18] GÜNTHER M, ONG J, MOHR I, et al. Jina embeddings 2: 8192-token general-purpose text embeddings for long documents [J]. arXiv:2310.19923, 2023.



CAO Tianruo, born in 1999, postgraduate. His main research interests include recommend system, knowledge graph and natural language processing.



LI Jingyue, born in 2002. Her main research interests include recommend system and natural language processing.