【LLM论文阅读】Google Research：个性化语言提示的User Embedding模型

原创 方方 方方的算法花园 2024年11月02日 09:52 北京

0 论文概况

1. 论文名称：
User Embedding Model for Personalized Language Prompting 《个性化语言提示的用户嵌入模型》
2. 论文链接：https://arxiv.org/pdf/2401.04858
3. 论文作者所在机构：Google Research
4. 一句话概括：提出一种用户embedding模块（UEM），通过将用户历史转换为嵌入表示作为soft prompts输入语言模型，提升了模型处理长用户历史及理解用户偏好的能力。

1 论文出发点

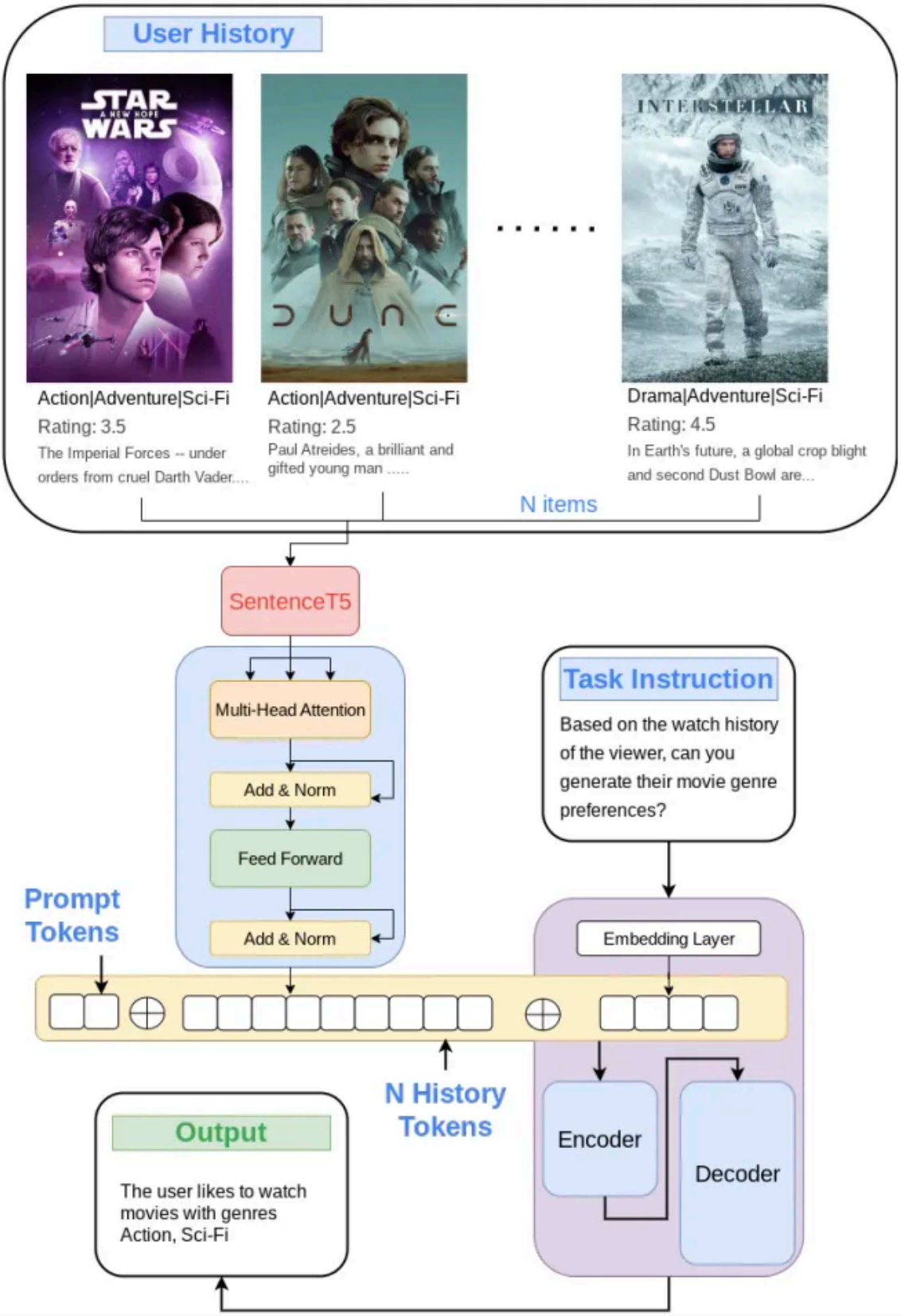
随着大语言模型（LLMs）在各种任务中的广泛应用，尤其是在理解用户偏好以生成推荐方面，**如何有效利用更长的用户历史成为关键问题。**当前研究大多集中在选取用户历史的代表性样本，而本研究旨在通过引入用户嵌入模块（UEM，User Embedding Module），将用户的整个历史以嵌入的形式进行压缩表示，从而更好地理解用户偏好并生成更精准的预测，同时探索这种方法在处理长用户历史时相较于传统文本方法的优势，以及对模型性能的影响。

2 论文贡献点

1. 提出新的用户嵌入模块（UEM）
能够高效处理自由形式文本的用户历史，将其压缩并表示为embedding，作为 soft prompts 输入语言模型（LM），从而增强了模型对长用户历史的处理能力，相比传统文本方法能处理更长历史，显著提升预测性能，模型在F1分数上有高达0.21和0.25的提升。
2. 创新的个性化 soft prompts 方法
基于用户历史生成个性化 soft prompts，使模型能更好地理解用户偏好，在任务中最大化标签的可能性，提高了模型对用户个性化需求的捕捉能力。
3. 深入的实验研究与分析
（1）通过在MovieLens数据集上的实验，详细评估了不同方法和模型设置对用户偏好理解任务的影响，包括历史长度、语言模型选择、用户embedding模块大小等因素的消融实验，为模型优化提供了依据。
（2）提出将任务视为多标签分类问题并采用加权精度、召回率和F1分数评估，相比传统指标能更细致地理解任务性能，尤其在处理类别不均衡的电影类型数据时表现出优势。
4. 研究的拓展性和前瞻性
（1）讨论了未来可探索使用更参数高效的方法（如LoRA）与UEM结合来优化模型训练和服务，以及将该方法扩展到多模态信号的可能性，为后续研究提供了方向。

（2）尽管当前研究存在一定局限性，但为进一步研究提供了基础，如进一步优化文本提示、改进用户嵌入表示、在更多任务上进行测试等，有望推动该领域的发展。

3 UEM 具体方案



1.任务框架与模型概率建模

（1）遵循T5的text-to-text方法，将所有任务定义为基于输入的文本生成任务。对于给定的查询输入token序列X，模型输出Y的概率表示为 $Pr_\theta(Y|X)$，其中θ为模型权重。

2. soft-prompting生成与模型输入构建

（1）用户嵌入模块（UEM）处理用户历史

将文本形式的用户历史$H = \{h_i\}_{i=1}^{p}$通过SentenceT5转换为embedding$U = \{u_i\}_{i=1}^{p}$，每个历史项***ui***是标题与类型、评分、描述三个不同embedding的组合，整体历史表示为$U \in \mathbb{R}^{p\times 3s}$（***s***为SentenceT5的embedding维度）。这些embedding在UEM中的Transformer网络中进行处理，并通过线性投影层将维度从***3s***映射到***e***（与语言模型embedding维度一致），得到$Pr_{UEM}(U) \in \mathbb{R}^{p\times e}$。

（2）结合任务级soft prompt

引入k个任务级 soft prompts $P_e \in \mathbb{R}^{k\times e}$，将用户提示$PrUEM(U)$和任务提示$P_e$与输入embedding$X_e$（由语言模型对输入x embedding得到，$X_e \in \mathbb{R}^{n\times \epsilon}$，n为token数）进行拼接，得到统一的embedding矩阵$[P_e; Pr_{UEM}(U); X_e] \in \mathbb{R}^{(k+p+n)\times e}$，该矩阵输入语言模型，最大化Y的概率并更新模型参数。

3. 模型训练与实验设置

（1）数据集与数据处理

使用MovieLens数据集结合电影描述，对用户历史数据进行格式化处理，如将标题和类型格式化为 “The movie {movie_title} is listed with genres {genres}” 等形式。数据集划分为训练集（117k）、验证集（5k）和测试集（5k）。

（2）模型选择与训练参数

实验主要使用FlanT5系列模型，训练10k步，批量大小为128，text历史模型学习率为1e - 2，embedding历史模型学习率为5e - 3。用户embedding模型包含3个Transformer层、12个注意力头、768维嵌入和2048维MLP层，添加65M参数，任务级soft prompts使用20个token。

4. 评估指标

将模型输出通过一个转换函数提取电影类型，把任务视为多标签分类问题，采用加权精度、召回率和F1分数来评估模型在理解用户偏好任务中的性能，以更细致地评估模型在不同类型上的表现和整体任务性能。

END

#LLM学习 12　LLM与推荐 15　LLM论文阅读 13

#LLM学习 · 目录

上一篇 · 【LLM论文阅读】Google Research：REGEN数据集 && CF与LLM融合框架