**学习进度**

本周学习任务：

学习Part I：Item Recommendation from Implicit Feedback；Deep Learning for RSs

下周学习任务：

学习Part I：Deep Learning for RSs

1. **基于隐式反馈的项目推荐**

**基于隐式反馈的项目推荐属于Top-N推荐任务，亦可理解为上下文感知的排序问题**，即：在给定上下文的情况下，应该怎样有序推荐某些商品。上下文不仅可以表示用户，还可以表达用户时间及用户之前选择项目的顺序等隐式反馈信息。同时，此类问题中的上下文和项目输入表示更为灵活，可以是用户或项目id，也可以是图像或文本描述等。从隐式反馈中学习的核心挑战在于：需要针对专门的优化目标，设计能够高效处理大型项目目录的算法；而且，隐式反馈将无交互记录视为弱负反馈的不确定性，也将导致学习成本高昂。

基于隐式反馈的项目推荐的**评分预测规则**，如下所示。其中，和是可学习的嵌入模型（亦可称为：双塔模型或双编码器）。比如，可以是一个多层感知机，可以根据项目各种类型的特征生成一个维的嵌入向量。



问题在于，根据给定上下文和预测规则，完成包含若干项目的有序推荐工作后，**如何评估项目推荐的质量？**

简单的想法是，根据隐式反馈标记的项目标签，使用精确度计算前个推荐项目中正样本的比率，来度量本次推荐工作的质量。然而，此类评价指标对推荐顺序不敏感。

由此，**平均精度均值（Mean Average Precision，MAP）**通过对所有用户推荐的AP值取平均的方式，实现了关于推荐顺序质量的度量。公式如下所示，表示推荐项目数，表示第个推荐项目的相关性分数（命中为1，未命中为0），表示目标用户数。



此外，**归一化折损累计增益（Normalized Discounted Cumulative Gain，NDCG）**通过对所有推荐项目的反馈标签进行加权和归一化的方式，衡量了推荐的排名质量，公式如下所示。其中，DCG按照推荐顺序依次累加推荐项目的反馈标签值，并额外除以对应位置价值的惩罚因子，使得靠后的推荐项目标签值置信度更低。此外，由于DCG属于累计值，故此需要额外除以完成归一化操作。IDCG是指理想情况下的最大DCG值，|REL|表示推荐结果按反馈标签值从大到小的顺序排序后，取得的前个结果。



值得注意的是，相较于NDCG，AP通过每个位置的准确率作为对应相关性得分的惩罚因子，关注到了排序性能的同时，却忽略了不同位置上相关性得分的差异，这可能会给相关性得分高的结果过大的惩罚。而NDCG则考虑到了相关性得分和位置的权重，对于排名靠前的结果更为敏感，然而其计算复杂度相对较高。

1. **深度学习推荐方法**

目前，使用深度学习方法构建推荐系统的趋势，备受追捧。深度学习推荐方法能够引入非线性并学习潜在高阶交互，使得交互建模环节更为便捷。而且，深度学习方法本身，能够更轻易地完成多类型数据的表示学习。同时，使用序列模型构建推荐系统，能够更好地考虑时间动态及用户兴趣的多样性。此外，推荐模型能够通过额外生成关于推荐项目描述的方式，使得深度学习推荐方法具有一定程度的可解释性。

然而，深度学习方法需要海量可用数据用于训练、调试及测试，才能保证模型的性能。同时，深度学习推荐方法优点的体现，皆需要付出额外的时空复杂度。值得注意的是，当前许多关于深度学习推荐方法的工作，缺乏一致的评估及数据分割标准。比如，数据集的选择和划分具有一定程度上的随机性，某些工作使用不恰当的评估方法夸大特定问题等。

在使用深度学习方法构建推荐系统的过程中，可以完成交互建模及用户建模工作。其中，交互建模是为了根据用户的历史交互行为来预测用户对特定物品的感兴趣程度，而用户建模则是为了更好地理解和描述用户特征及兴趣。值得注意的是，交互模型及用户表示模型是相互依存的关系，用户表示模型可以为交互建模提供关键的输入或上下文，而交互模型的结果又可以进一步改进用户建模的准确性。

1. **交互建模**

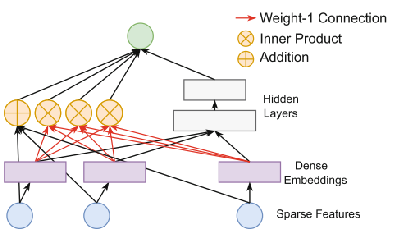
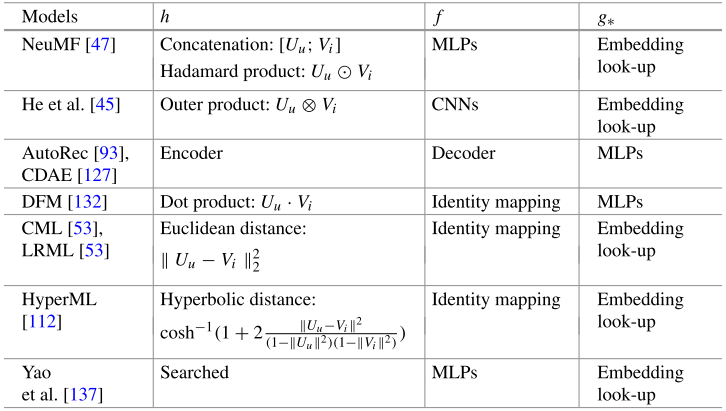
交互建模属于推荐系统的核心问题，使用深度学习方法构建一个好的交互模型，除了需要具备准确预测用户对于特定物品感兴趣程度的能力之外，还需要具备自动学习特征交叉组合的能力。由此，衍生出用户-物品交互建模及特征交互建模这两个子问题。

**（1）用户-物品交互建模**

针对用户-物品交互数据，深度学习推荐方法可以构建模型，表示嵌入模型，表示评分映射模型，表示用户与物品间的交互连接方式。其中，关于用户或物品的嵌入表示学习，不仅可以使用用户或物品的标识符，还可以合并每个用户或物品的所有历史交互来丰富表示。比如，将用户评价过的所有物品作为输入，从而获得对应的用户嵌入表示。

用户-物品交互建模方式的部分示例，如下左图所示。其中，**NeuMF**的用户-物品交互连接，存在将拼凑或逐元素相乘两种方式；**AutoRec**等模型使用编码-解码器结构，完成交互连接及评分映射；**CML**等模型依赖用户与物品间的距离度量，完成交互连接。

值得注意的是，由于不同推荐场景下的交互模式可能非常不同，使得能够灵活适应各种应用场景的通用交互算法，设计起来非常困难。故此，Yao等人提出可以根据包含内积、外积、欧氏距离、拼接及最大小值等不同交互连接方式的**搜索空间**，然后使用AutoML在搜索空间中自动寻找出最优交互函数。



**（2）特征交互建模**

特征交互建模，主要关注模型自动学习样本特征交叉组合的能力。过去，常使用因子分解机FM，通过任意特征分量间内积的方式，实现特征交叉的自动学习。现在，则在FM的基础上进一步改进，比如使用洛伦兹距离代替内积将FM推广到非欧几里德空间的**LorentzFM**，或者是使用循环卷积等运算来代替内积的**HFM**。

更进一步，**DeepFM**（如上右图所示）通过结合多层感知机（MLP）及线性模型的方式，构建一个宽且深的深度学习框架，用于学习样本特征的交叉组合。其中，深度部分（MLP）用于捕捉非线性的高阶特征交叉组合，宽度部分（线性模型）用于捕捉线性的低阶特征组合。DeepFM模型的显著特点是，深度和宽度部分共享同样的输入特征向量。然而，MLP隐式地模拟特征相互作用的方式，不一定对所有类型的交叉特征都有效。由此，**xDeepFM**可以通过结合广义交叉网络及MLP的方式，实现矢量特征间的交互，从而有效打破MLP关于特征相互作用隐式建模的限制，进一步提高关于特征组合的学习能力。此外，如若需要考虑不同特征组合的重要性，可以使用引入注意力机制的深度因子分解推荐算法。比如，**AFM**使用参数化的注意力网络，为因子分解机FM中每对特征组合计算一个对应的注意力分数，然后通过加权求和的方式，重新分配注意力到更重要的特征组合身上。

值得注意的是，相较于传统的因子分解机FM算法，前述若干深度因子分解推荐算法的复杂度更高，使得此类方法在处理大型数据集时，可能会面临计算上的挑战。

1. **用户建模**

在实际推荐场景中，由于交互数据本身的动态性及用户兴趣的不确定性，使得传统机器学习方法缺乏可用性。因此，在用户建模的过程中，可以通过时间序列及用户多元兴趣建模的方式，分别额外考虑用户历史交互的时间顺序及用户兴趣的多样性。

1. **Temporal Dynamics Modeling**

在用户建模的过程中，考虑用户历史交互的时间顺序，是由于用户的短期交互能够对当前决策产生更大的影响，而且用户在某些时间语境中（比如，新年等节日）更可能会做出特定的决策，此外商品的受欢迎程度本身也是随时间发生变化的。引入时间序列的深度用户建模方法，主要可以分类为基于序列、卷积及自注意力等，相关工作如下所述。

在基于序列模型的方法中，循环推荐网络**RRN**使用LSTM捕获用户和项目的长短期时序效应。然而，RRN存在用户和项目需要分开建模的缺陷。更进一步，可以改进序列模型中的门控单元。比如Guo等人提出的**PcGRU**是一种改进的GRU，它依赖于位置感知及控制机制，可以更精细地理解用户兴趣的动态演化过程。

在基于卷积模型的方法中，比如，用于顺序感知推荐的分层时间卷积网络**HierTCN**，它主要由GRU和时序卷积网络组成，前者用于对会话期间的长期兴趣建模，后者用于学习动态用户嵌入。同样，二维卷积网络**CosRec**，通过提取学习用户行为序列中的局部模式及上下文信息，高效地完成了用户兴趣及行为规律的建模表示。

在基于自注意力的方法中，Atrank率先利用自注意力网络，捕捉了用户历史行为的时间动态。Xu等人则**通过自定义的时间核学习自注意力模型中的函数时间表征**，考虑到了用户历史行为的时间跨度（若用户上一个动作的发生时间和当前时间跨度较大，则该动作的置信度可能很低）。而Sun等人采用类似于BERT的模型，使用**深度双向自注意力**模块，主要考虑了从左到右及从右到左的双向时序依赖关系。

此外，如若面临缺失用户标识符或系统只能记录当前会话中活动的情形，则需要考虑基于会话的深度用户建模方法。比如，**GRU4Rec**采用门控单元GRU完成会话中关于时间顺序模式的建模，用于预测下一个事件。令人惊讶的是，基于卷积的顺序学习架构有时候比基于序列的模型更有效，基于卷积神经网络的用户建模方法显示了更有希望的结果。

1. **Diverse Interest Modeling**

过去，常将所有的兴趣编码成潜在因子，并没有考虑到用户可能会因为不同的原因，对商品的某些方面感兴趣。由此，在用户建模的过程中，需要使得模型意识到并有能力区分用户兴趣的多样性。

比如，Li等人提出了一种**带有动态路由的多兴趣学习方法**，它可以从用户历史交互中解耦用户兴趣的多面性，从而实现更丰富的用户表示。其中，动态路由属于软聚类方法，它将用户的历史行为划分为若干群组，每个群组对应一个特定的潜在兴趣。此外，Ma等人使用**解耦的变分自编码器**，来推断用户兴趣的宏观解耦（高级意图，比如购买手机）和微观解耦（低级因素，比如颜色和大小等）。显而易见的是，通过学习用户兴趣的解耦表示，能够使得推荐列表更具可控性的同时，也提高了关于用户表示的可解释性。

1. **内容表示学习**

尤其在传统的基于内容的推荐系统中，根据各种类型（文本、图像及音频等）的数据，学习用户或物品在同一潜在因素空间中的映射表示，并非易事。现在，深度学习方法能够更轻易地提取和利用各种类型数据的特征，用于提高深度推荐系统的性能。

1. **文本特征提取**

项目的文本描述（比如，论文摘要、电影情节摘要及新闻内容等），是推荐系统中使用最多的文本数据之一。为了利用这些文本描述，结合自编码器及BPMF的协作深度学习模型**CDL**，在忽略上下文信息的前提下，通过词袋模型完成了输入表示。Bansal等人则提出了一种**端到端的协同过滤模型**，它利用GRU将与项目相关的文本编码为嵌入向量，用于冷、热启动推荐。

此外，针对基于评论的推荐问题，特殊之处在于：文本评论中蕴含的丰富语义信息，无法简单地通过隐式交互数据完成传递。故此，为了更好地利用评论信息，Zheng等人提出了一种**深度合作神经网络**，实现了用户和物品评论的联合建模表示（每个用户用他写过的所有评论表示，每个物品用它收到的所有评论表示）。Tay等人提出的**多指针共同关注网络**，能够筛选出更相关的评论用于对象表示，而非盲目地拼接所有评论。而Guan等人设计了一种**面向因素的推荐方法**，能够更好地理解评论中用户对于项目不同方面的差异性偏好。

1. **图像特征提取**

图片往往更能够吸引用户的注意力，视觉特征或许能够更好地捕捉用户的潜在偏好。具体而言，可以使用卷积神经网络等方法提取图像特征，并将图像特征融入到推荐系统中，以提高推荐的准确性和个性化程度。比如，He等人将视觉特征融入BPR框架中，提出了**VBPR**模型用于Top-N推荐任务。Geng等人则在Pinterest上使用卷积神经网络，学习基于图像的视觉特征表示。

1. **音视频特征提取**

针对考虑节奏、旋律及音色等音频信息的音乐推荐任务，Wang等人使用深度信念网络**DBN**自动提取音乐特征完成了音乐推荐。同时，视频也可以被转换成一系列帧和音频波，用于改进推荐。比如，Xu等人提出的**关键帧推荐系统**，使用了一个由5个卷积层和3个全连接层构成的卷积神经网络，完成了视频帧的表示学习。