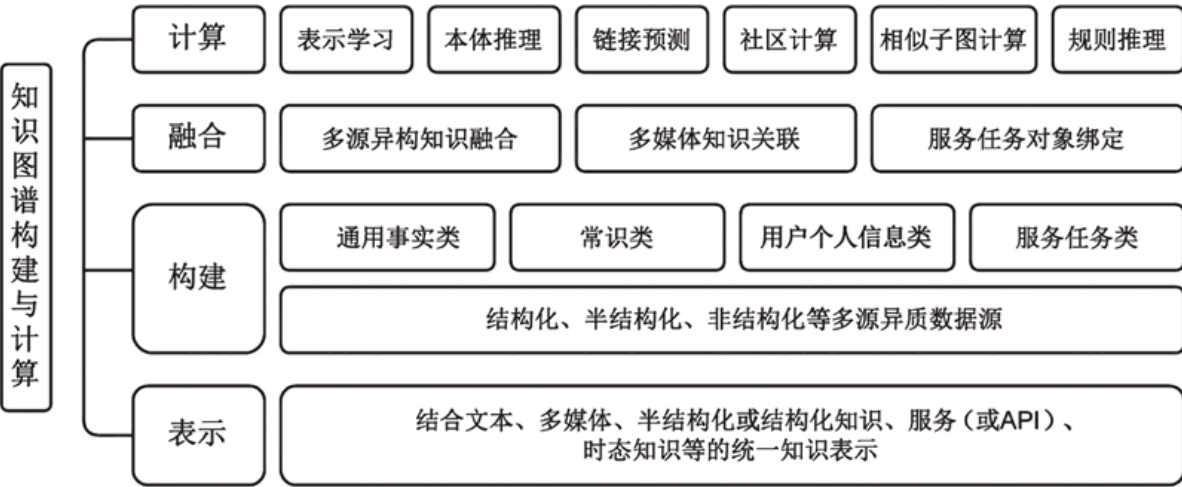


knowledge graph

知识图谱可以看作一类语义网络（Semantic Network）。语义网络是一种表示网络中概念（Concept）之间语义关系的知识库，通常是一个有向或无向图，由表示概念的结点和表示概念之间语义关系的边组成。

知识图谱本质上是一种叫作语义网络的知识库，即一个具有有向图结构的知识库，其中图的结点代表实体或者概念，而图的边代表实体/概念之间的各种语义关系。

知识图谱是符号学派的代表，可以帮助我们构建更有学识的人工智能，从而提升机器人推理、理解、联想等功能。而这一点，仅通过大数据和深度学习是无法做到的。多伦多大学的 Geoffrey Hinton 教授也提出，人工智能未来的发展方向之一就是深度神经网络与符号人工智能的深入结合。知识图谱是实现通用人工智能（Artificial General Intelligence, AGI）的重要基石。



1. 如何搭建一个知识图谱

可以分为以下两大步骤：

- 知识提取
 - 信息抽取，获取三元组
 - 实体识别、实体链接、实体消歧、实体统一
- 图构建
 - 存储
 - 查询

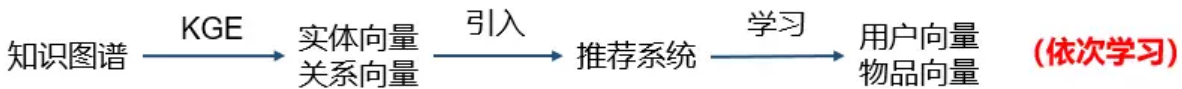
知识提取步骤是构建知识图谱的关键，三元组可以通过依存分析得到。

2. 知识图谱特征学习应用到推荐系统主要通过三种方式

依次学习、联合学习以及交替学习

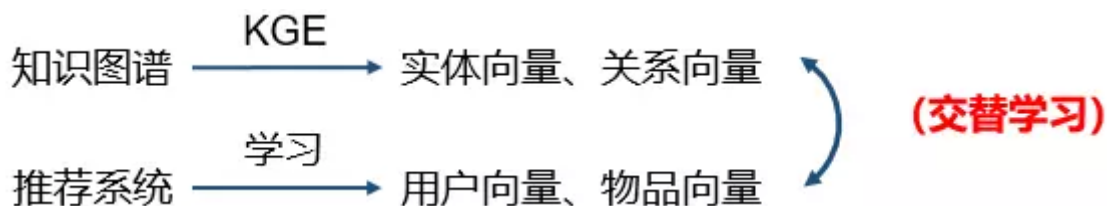
依次学习

首先使用知识图谱特征学习得到实体向量和关系向量，然后将这些低维向量引入推荐系统，学习得到用户向量和物品向量；



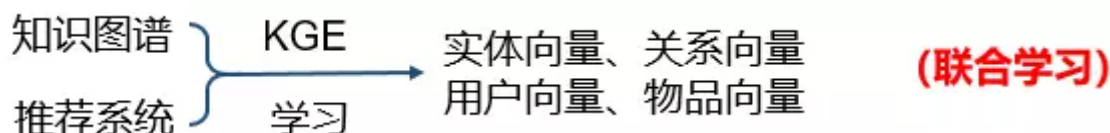
联合学习

将知识图谱特征学习和推荐算法的目标函数结合，使用端对端的方法进行联合学习：



交替学习

将知识图谱特征学习和推荐算法视为两个分离但又相关的任务，使用多任务学习的框架进行交替学习。

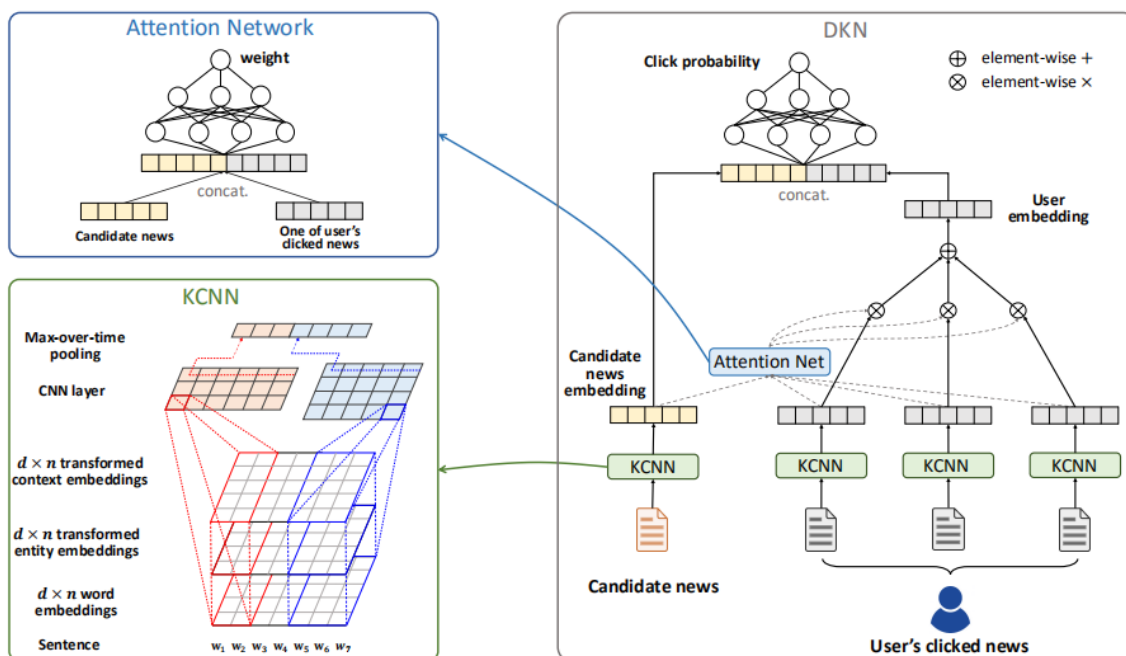


依次学习的案例---DKN

《DKN: Deep Knowledge-Aware Network for News Recommendation》

在新闻推荐领域存在三个主要挑战，第一，新闻具有时效性，传统的协同过滤方法不再有效；第二，用户在看新闻时会关注多个不同的主题，如何动态捕捉用户兴趣是一大难点；第三，新闻内容是高度精简的。

DKN的框架如下图所示，DKN接受一组候选新闻和用户历史数据，KCNN部分负责将知识引入推荐过程，attention部分对新闻赋予不同的权重，捕捉用户兴趣，和DIN一样。最后将得到的向量拼接，送入神经网络输出点击概率。



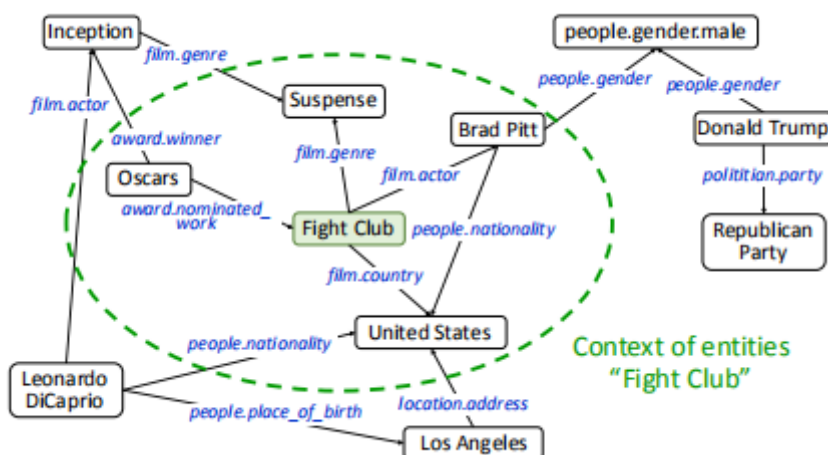
KCNN的输入特征包括三部分：新闻标题词向量 w_i 、实体向量 e_i 、上下文向量 \bar{e}_i 。

- w_i : 标题词向量根据语料库预训练得到
- e_i

实体向量由知识图谱特征学习算法得到，具体过程如下：

- 1.使用实体链接技术将新闻文本中的实体和知识图谱中的实体相关联，消除歧义；

- 2.基于识别出的实体构造一个子图，并从原始知识图中提取它们之间的所有关系链接。注意，所识别的实体之间的关系仅是稀疏的并且缺乏多样性。因此，我们将知识子图扩展到已识别实体的一跳内的所有实体。
- 3.根据构造的子图，用TransE、TransH、TransR等方法得到embedding向量 e_i
- \bar{e}_i :为了更好的理解实体在知识图谱中的位置信息，作者提出了一个额外的语境向量。一个实体的语境“context”定义为一跳邻居节点，语境的直观解释如下



每个输入向量作为一个通道，类似图像中的RGB通道，即多通道。注意，这里没有直接拼接原始向量和实体向量，因为1) 拼接会破坏词和实体之间的关系；2) 词向量和实体向量是用不同方法学习得到的，直接在单个通道进行操作显然不合理；3) 拼接隐式的要求词向量和实体向量有相同的维度，实际上它们之间有差异。

依次学习过程：

构建知识图谱、学习实体特征向量、将实体特征向量和原始向量共同做为输入进行多通道卷积、利用attention机制捕捉兴趣、最终向量送入到网络输出点击概率。

KOPRA

mind数据集解释:<https://juejin.cn/post/7028114250577117220>

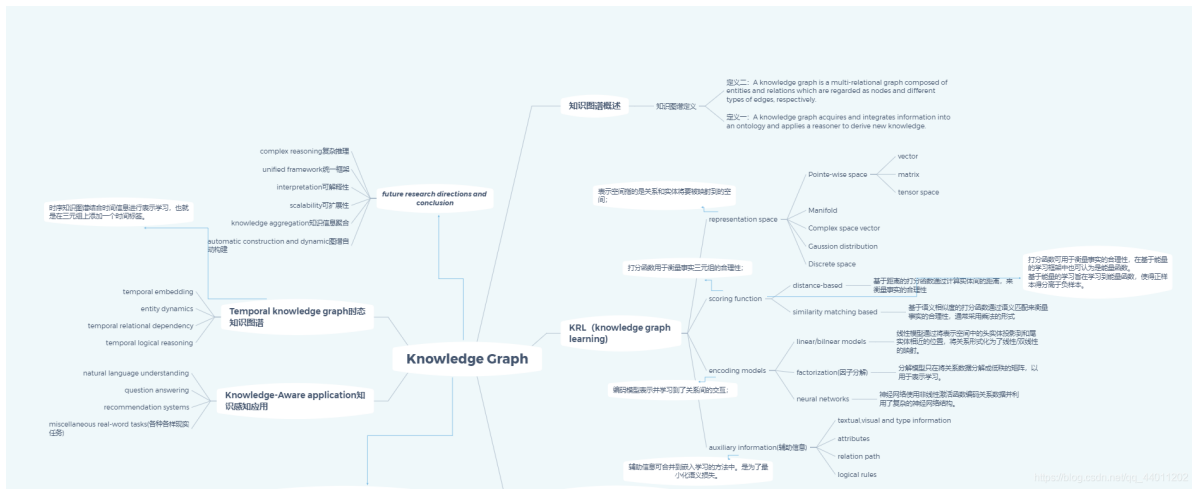
https://blog.csdn.net/qg_27590277/article/details/112211183

<https://www.jiqizhixin.com/articles/2020-07-15-5>

循环图卷积网络

循环图卷积技术是一种基于图卷积神经网络（GCN）的模型，它可以处理具有周期性结构的数据。在循环图卷积技术中，每个节点都会将自己的特征与邻居节点的特征进行组合，然后将结果传递给下一个节点。这个过程会一直重复，直到所有节点都被处理完毕。这种重复的过程就像是一个循环，因此被称为循环图卷积技术。

知识图谱论文综述



https://blog.csdn.net/qz_44011202



https://blog.csdn.net/qz_44011202

<https://www.cnblogs.com/0111B/articles/14655352.html>

<https://blog.csdn.net/byn12345/article/details/106128275>

<https://www.jianshu.com/p/aaae8b8b5bbe>