

# 论文整理

2023年4月17日 13:34

WWW2023推荐系统涉及到的主题：时序推荐、基于图的推荐、可解释推荐、推荐系统中的bias、因果相关、公平性和隐私保护、强化学习、冷启动、跨领域、多任务、对比学习、多模态等。

## 基于多模态融合的推荐系统 2021.1 陆杨思旖

1、利用自注意力机制学习特征间的关系，提出了一个能同时利用文本模态和视觉模态进行推荐的算法框架。

**利用注意力机制搭建了一个基础的多模态推荐算法框架**，并在此基础上探索不同的融合位置对推荐结果产生的不同影响。

2、提出了一种能同时利用融合前的单模态高阶特征信息和融合后的多模态高阶特征信息的**中期融合策略**，并在搭建的基础算法框架上进行实现。

3、在融合两种模态特征前交换它们的特征信息（由于这种交换操作类似于在两种模态间实现一种交叉跨越和相互特征借鉴，因此这样的算法称为基于跨模态融合的推荐算法。）通过对两种模态特征**采用内积的方式**来计算跨模态权重，以此来挖掘到这两种模态特征间的直接关系并得到包含更多信息的跨模态特征表示。即通过拓展优化得到一个新的算法——**基于跨模态融合的推荐算法**。

针对如何利用多模态技术和知识图谱来增强用户表示这一问题，作者提出了一种基于知识图谱的多模态推荐算法，利用自注意力机制来计算目标物品和用户相关物品间的相似度，并在知识图谱中分别传播两种模态下的用户偏好信息，最后利用相加、拼接和跨模态融合的方式得到增强的用户表示。

作者认为现有的工作大多只关注如何通过引入多模态信息来增强物品的表示，或者只是将多模态特征映射到同一空间然后通过知识图谱传播用户单一的偏好信息，并没有真正地去捕获用户对不同模态特征的细粒度偏好。

作者是在基于知识图谱的经典单模态推荐算法RippleNet的基础上，提出一种新的基于知识图谱的多模态推荐算法。

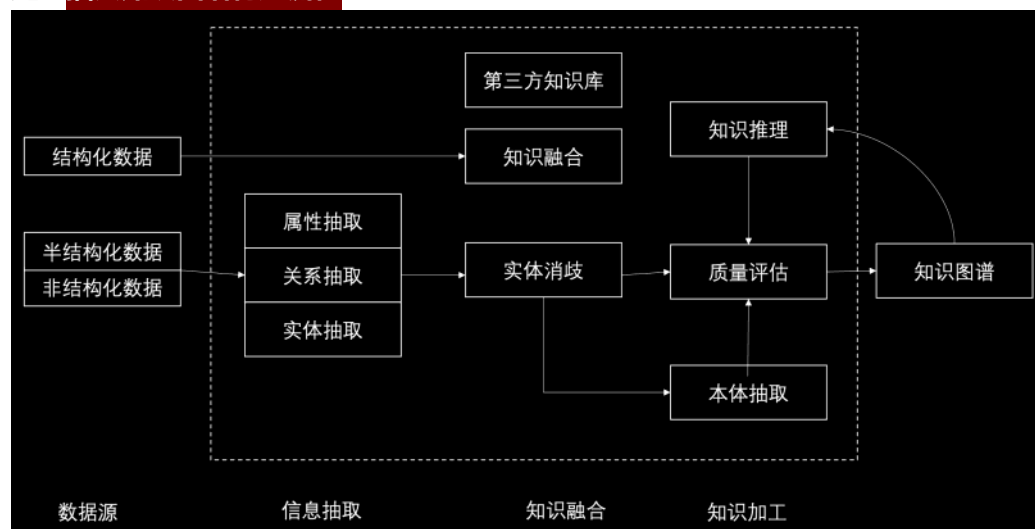
文章的最后作者提出了展望，包含了目前算法中依旧存在的问题或者缺陷。之后可以进行回顾。

## 基于知识图谱的个性化推荐方法研究 2021.6.1 宋秀来

- 1、设计了一种探索用户偏好传播的方法。利用知识图谱中的三元组展示了多步骤的相互关系，该关系是用户-项对的一条或者多条路径。
- 2、设计了一种知识感知路径可解释性推荐方法。知识感知路径模型包括三层：嵌入层、LSTM层、池化层
- 3、设计了多模态特征交互深度融合推荐方法。将多模态的特征数据用不同的方法进行处理，将标题数据转换为序列特征；视频和音频使用pca降维，输入到两个不同的DNN层，然后结合原始特征数据输入到深度融合模型（xDeepFM）中。

文章思路首先是分析了知识图谱路径查询相关技术和原理，为后文研究知识感知路径算法提供支撑。

### 这里插入知识图谱构建流程



构建知识图谱的数据源有多种：

- 1、结构化数据（可以直接输入到知识融合中）
- 2、非结构化数据和半结构化数据（需要进行信息抽取）

构建过程中：需要对数据进行相应的处理：

- 信息抽取（对于非结果化和半结构化数据获取对应的三元组信息，形成基础的知识表示）
- 知识融合（将获取的新三元组信息进行合并）
- 知识加工（对融合的知识或者本体抽取的信息进行质量评估，选取质量合格的三元组信息融合到知识库中）

一轮构建结束，需要对知识图谱进行一直迭代，让知识图谱信息更加丰富。

### 知识图谱路径查询

作用：根据数据集中的user-item对图谱进行路径查询，获取对应的数据。

方法：带标签的路径查询、元路径查询、关键字查询、社团搜索查询

### 知识感知路径可解释性推荐算法

将user和item作为实体，找到得到最高的路径作为一个结果进行可解释性推荐

## 基于多模态特征交互深度融合推荐算法

特征的构建：

- 1、运用深度学习技术提取原始数据中含有语义的隐式特征（视频、音频、文本等）
- 2、自动学习多个相关特征之间的交互关系

本文首先将多个模态的特征数据利用不同的方法进行处理进行最终交互，然后深度融合模型以xdeepfm为基础。

特征交互方法：

- 1、FM
- 2、DeepFM

## 基于多模态知识图谱的推荐系统 孙睿 2021.3

提出了MKGAT模型（多模态知识图注意力网络）

该模型利用**改进的图神经网络将多模态知识图谱中的多模态融合过程转化为图的信息传播过程**。进一步丰富推荐系统中物品和用户的表征。

本文构建了两个含有多模态知识图谱的推荐系统数据集，且在这两个数据集上进行了详细的试验。

前人的多模态知识图谱表示学习方法还只是通过考虑单一三元组的方式来训练嵌入向量，并没有加入显示的多模态的信息融合过程。例如，toy story既和toy story图像实体相连，又和toy story的文本实体相连，因此只考虑单独的三元组，难以学到即包含图像信息，又包含文本信息的目标实体嵌入向量。本文创新性的将知识图谱中多模态的信息融合过程转换为图上的信息传播过程，利用了信息在多模态知识图谱上的传播，在传播过程同时考虑了多模态信息融合，以获得更好的知识图谱实体嵌入向量。

本文多模态知识图谱的构建流程

也是模型MKGAT的输入：协同多模态知识图谱

采用的数据集：movieLens（电影）和Dianping（美食）

## 多模态知识图谱构建与应用研究综述 陈烨 2021.12

多模态知识图谱基本概念

多模态知识图谱的构建工作，总结了两种主要方法的思路

分析了多模态知识图谱的构建和应用中的关键技术和相关工作：例如多模态信息提取、表示学习和实体链接

### 1、知识图谱

知识图谱一般采用自动或者半自动技术从结构化、半结构化以及非结构化数据资源中抽取知识，并存入基于逻辑划分的数据层和模式层，是一个迭代更新的过程。

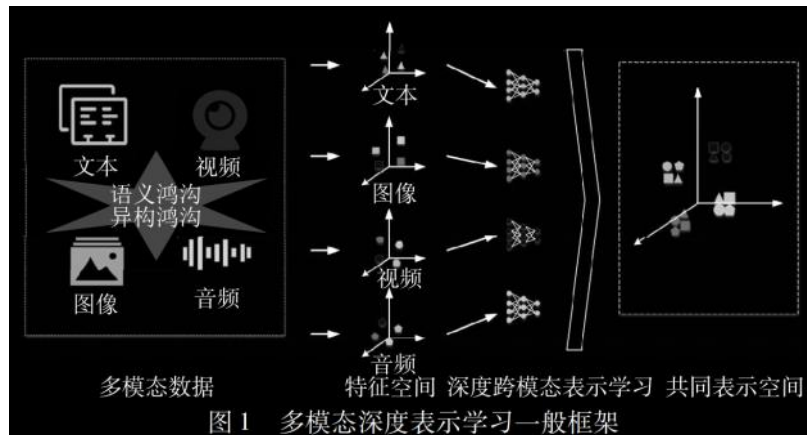
### 2、多模态学习

本文讨论的多模态数据可理解为描述同一对象的多媒体数据，多模态数据虽然在底层表征上是异构的，但是相同实体的不同模态数据在高层语义上是一致的。

多模态数据之间由于其本身结构特点，技术研究主要面临两大挑战：

- a、语义鸿沟：计算机表示系统与人类认知系统对同一个概念形成不同描述的差异
- b、异构鸿沟：指图像、文本等不同媒体的数据具有不同的特征表示形式，它们的相似性难以直接度量。

对多源异构数据的挖掘分析可被理解为多模态学习，其任务是通过学习多个模态数据中的信息，实现各个模态信息的转换和交流。



用合适的神经网络学习文本、图像、音频、视频等多模态数据在相应特征空间的表示，随后将各模态的表示作为输入，继续构建更深层的神经网络结构，利用深度跨模态表示学习构建的神经网络融合各模态的语义信息 得到在共同表示空间中各模态的高层语义表示。

多模态数据数据量大、数据分布稀疏，多模态数据集的构建更加困难，指代同一实体的不同模态数据通常需要更加昂贵的人工标注。

### 3、多模态知识图谱

定义：

MMKG: multi-modal knowledge graphs 从[链接预测](#)和[实体匹配任务](#)出发定义多模态知识图谱，是包含了所有实体的数值特征和图像，以及多个知识图谱之间的实体对齐的知识图谱。

多模态知识图谱在传统知识图谱的基础上，构建了多种模态下的实体，以及多种模态实体间的语义关系。

包含文本和图像等多种数据类型的知识图谱，将视觉或文本信息引入到知识图谱中，将图像或者文本作为实体或者实体的属性。



额外信息：Richpedia（东南大学开源：首个多模态开放知识图谱）

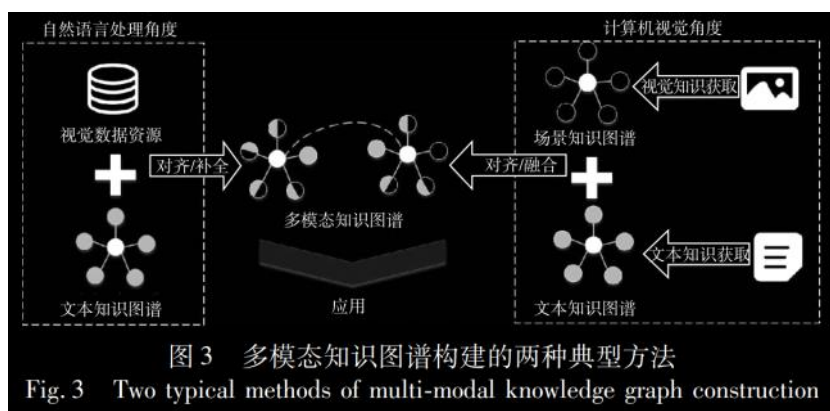
richpedia包含了全面的图像实体以及实体之间的关系，可以广泛应用于新闻阅读、时事推荐、商品查询等场景中，多模态知识图谱补全技术也可以通过远程监督补全多模态知识图谱，完善现有的多模态知识图谱，利用动态更新技术使其更加的完备。

上图中参照Richpedia中的形式化定义，遵循RDF框架，扩展图像模态作为实体在知识图谱中的存在，并强调图像模态实体之间的关系发现。

### 3.1 多模态知识图谱的构建

基于图数据库视角的构建

基于知识图谱视角的构建



多模态知识图谱构建的两个典型思路：

- 1) 自然语言处理角度出发：在文本知识图谱构建的基础上，给实体补充视觉信息，工作实质是知识图谱的补全，并在图像间做视觉关系的发现和跨模态实体链接
- 2) 计算机视觉角度出发：建立在场景图谱生成基础之上，是一种分布式的构建方法，将视觉知识与外部文本知识图谱相桥接。

## 4、多模态知识图谱关键技术

多模态信息抽取、多模态表示学习、多模态实体链接

### 4.1 多模态信息抽取

信息抽取的主要目标：从无结构的生文本中抽取结构化、半结构化或者非结构化的数据。对于不同模态的信息可以采取先进行单模态抽取，再进行多模态处理的方法。



图像模态信息抽取

在图像特征提取的角度下，CNN通过卷积和池化操作，产生图像模态的矩阵表示，图像特征的向量表示工作由全连接层或者全局均值池化层的输入完成。通过增加网络的深度或对卷积和池化操作进行变形，后续研究提出了更加复杂和搞笑的卷积神经网络，陆续有VGG、NIN、GoogLeNet等，对图像模态的特征提取能力进行了拓展，极大提高了图像识别的精度。

表2 部分经典卷积神经网络统计		
Tab.2 Survey of several classical convolutional neural networks		
CNN 网络	方法	效果
LeNet-5 <sup>[53]</sup>	利用卷积和池化提取图像特征,将结果输入到全连接层生成图像特征向量	在手写体数字和字母识别中得到极高精度,是现代卷积网络的基础
AlexNet <sup>[54]</sup>	使用 ReLU/normalization/dropout 增加网络深度,得到包含图像深度语义信息的特征表示	IISVRC 2012 冠军,top-5 错误率 16.4% ,8 层神经网络
VGGNet <sup>[55]</sup>	利用小卷积核和多个卷积层来捕获精细的图像特征	IISVRC 2014 年度亚军,top-5 错误率 7.3% ,19 层神经网络
NIN <sup>[56]</sup>	使用 MLPconv 算法改进卷积层,使用全局平均池代替全连接层	获得高度非线性图像的矩阵表示,提高了网络的泛化能力
Google InceptionNet <sup>[57]</sup>	使用不同大小的卷积内核的 inception 模块	IISVRC 2014 年度冠军,top-5 错误率 6.7% ,22 层神经网络
ResNet <sup>[58]</sup>	融合身份映射和残差映射	2015 年 IISVRC 冠军,top-5 错误率 3.57% ,152 层神经网络
CapsNet <sup>[59]</sup>	输入和输出均由胶囊神经元向量形式构成,每个元素是图像中实体特征的参数表示,通过动态路由算法将相邻的胶囊层连接起来	避免了图像中实例丢失方向和空间信息

4.2 多模态表示学习

文本模态表示

提取文本特征的神经网络主要包括：前馈神经网络、RNN（循环神经网络）、LSTM（长短期记忆网络）及其变体。近年来，基于自注意力机制的Transformer及其变体BERT具有比RNN更好的文本建模效果，逐渐取代RNN成为主流的文本特征提取方法。

图像模态表示

卷积神经网络针对图像处理取得了优异的成绩

多模态知识图谱表示学习相关工作

基于特征的方法

将多模态信息作为实体的辅助特征来处理，基于翻译的模型TransE将实体和关系投影到相同连续的低维向量空间，并且将关系解释为头和尾实体之间的翻译

基于实体的方法

将不同的模态信息作为结构化知识的关系三元组，而不是预定的特征。  
VTransE、TransAE、

4.3 多模态实体链接

实体链接（entity linking，EL）任务是指将从给定资源中抽取的实体对象链接到知识图谱中对应的实体中。

基于知识图谱的多模态可解释服装推荐 汪淳

2022.5.20

本文提出了一种基于知识图谱的多模态可解释推荐算法（MEFR：Multi-modal Explainable

## Fashion Recommendation based on Knowledge Graph)

本文为了应对异构的多模态数据，分别引入了面向属性和面向视觉的协作知识图，并设计了MEFR的三个关键组件：基于属性知识图谱的表示学习、基于视觉的表示学习和多模态增强的偏好预测。

挑战：

- a、如何有效地利用异构多模态数据（图像和属性）来增强用户偏好和服装表示学习
- b、如何自适应的权衡多种多样的服装属性信息在表征服装商品以及用户偏好上的不同贡献，从而提高模型的推荐性能以及可解释性
- c、现实世界的属性分布通常是不均匀的，例如：几乎所有电商都有“类别”和“颜色”属性，但是只有鞋子拥有“鞋跟”属性，这就导致了不同类型的属性在数据中出现次数的差异是巨大的。因此，如何正确处理建模过程中不均衡的属性分布，从而准确地建模服装和用户偏好。

## 可解释推荐

**出现的原因：**为了让用户更好地信任我们的推荐系统，一种比较好的方法是为用户推荐物品的同时给用户推荐的理由，向用户解释清楚是基于什么原因给他推荐的。

**推荐解释：**为用户提供推荐的同时，给出推荐的理由。

**推荐解释价值：**增加认可度和接受度，没有解释和理由的推荐是缺乏足够说服力的。有一定的推荐解释可以提升用户对推荐系统的信任度和接受度，进而提升用户对推荐产品的满意度。并且对于很多行业，是必须具备解释能力的，例如风控、金融、医学等。

## 互联网推荐产品的推荐解释模型

对于推荐模块，可以在模型训练过程中生成解释（事先解释），也可以在推荐结果生成后做解释（事后解释）

在模型训练过程中生成解释

- 1) 将推荐和解释看成是两个优化目标，通过协同训练来同时优化两个目标
- 2) 将解释过程嵌入到推荐过程中，解释和推荐过程耦合在一起，融为一体，集成为一个模型训练。（让系统更加复杂，建模过程难度大，训练时间长，但是可以提升整个模型的可解释性）

在推荐结果生成后做解释

基于给用户的推荐结果，从中找出用户跟推荐物品之间的某种内在联系，基于该关系做推荐解释。

该方案基本将推荐过程和解释过程解耦，工程实现上更加的简单，也更容易让用户理解和接受。

## 推荐解释的形式

一般为文字，不同产品的形态可以采用不同的形式，例如电台，可以采用音频的方式

## 推荐解释现状

研究和投入较少，因为都将精力放到了提升推荐系统的精准性上  
有些算法模型解释性好，有些模型解释性不好。

### 构建可解释性推荐系统的方法

#### 1) 基于用户关系来做推荐解释

基于真实社交关系做推荐解释

基于用户行为建立两个用户之间的关系，进而做推荐解释

基于用户画像做推荐解释

#### 2) 基于物品相似关系做推荐解释

基于内容特征

基于用户行为构建物品之间的关系

#### 3) 基于标签

通过标签建立用户与物品之间的关系

通过用户自身的标签

通过物品自身的标签

...

利用知识图谱增强算法解释能力

论文：Explainable recommendation: a survey and new perspectives. 2020,14

对于目前可解释推荐的研究进行分类

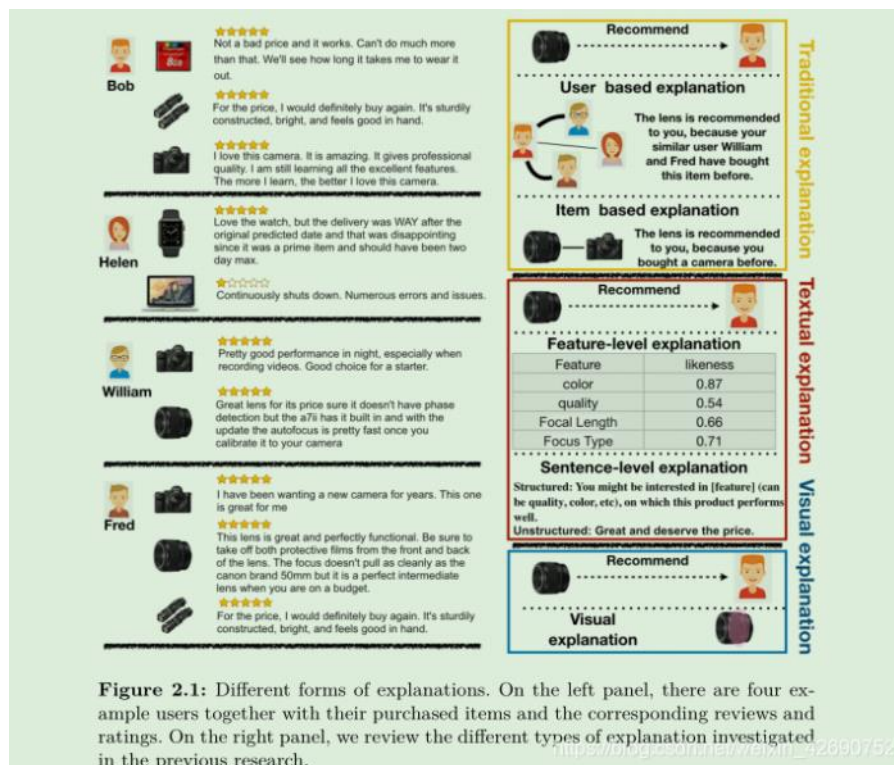
a、生成的推荐的类型（文本、视觉等）

b、用于生成解释的推荐模型或者算法（MF、基于图、深度学习、知识图谱、关联规则和因果模型等）

本论文中主要介绍了六种解释的类型：基于用户、物品的解释，基于内容的解释，文本解释，可视化解释，社交解释和混合解释

基于用户和基于物品的解释、文本解释和可视化解释用例





这里因为基于用户、基于物品和基于内容的形式是比较传统的推荐算法，所以不详细解释

## 文本解释

电子商务中的用户评论和社交网络帖子，这些信息对于获取用户更加全面的偏好有很大的价值，可以用来产生细粒度跟高和更加可靠的推荐解释。

文本解释可以分为特征层面（展示产品特征和用户对这些特征的可能的的情感）和句子层面（用句子告诉用户为何给他推荐这一产品）

## 可视化解释

在上述例子中，为了告诉用户推荐这个镜头的原因是它的颈圈外观，推荐的解释中对这一区域进行了高亮

## 社交解释

# 基于知识图谱的深度推荐系统研究 王越群 2022.6 博士论文

知识图谱融合至推荐系统策略：联合融合方法（近年来的热门）

基于知识图谱作为辅助信息的多模态多任务深度推荐模型

SIMKR模型：该模型对item属性采用多模态特征学习的方式进行属性的分类讨论。