论文整理：综述

面向深度学习的多模态融合技术研究综述

面向深度学习的多模态融合技术是指机器从文本、图像、语音和视频等领域获取信息实现转换与融合以提 升模型性能，而模态的普遍性和深度学习的热度促进了多模态融合技术的发展。在多模态融合技术发展前期，以提升深度学习模型分类与回归性能为出发点，阐述多模态融合架构、融合方法和对齐技术。重点分析联合、协同、 编解码器 3 种融合架构在深度学习中的应用情况与优缺点，以及多核学习、图像模型和神经网络等具体融合方法 与对齐技术，在此基础上归纳多模态融合研究的常用公开数据集，并对跨模态转移学习、模态语义冲突消解、多模态组合评价等下一步的研究方向进行展望。

多模态融合技术的主要 目标是缩小语义子空间中的分布差距，同时保持模态特定语义的完整性，例如利用多模态融合特征，提 高 视 频 分 类 [ 5 ] 、事 件 检 测 [ 6 - 7 ] 、情 感 分 析 [ 8 - 9 ] 、跨 模 态 翻译[10]等跨媒体分析性能。

融合的架构：联合、协同。编码器-----------模态映射，将不同模态的信息映射到一个语义空间，实现模态的对齐。

*读后感：多模态融合技术主要是针对于如何将不同模态的信息进行应用，主要包括于不同模态信息的表示不同，语义空间维度也不相同，如何将他们使用统一的方式进行表示。但训练多模态模型工程较大，需要大量数据，我的目标是使用一个已有的多模态模型，应用于自己的下游任务（多模态新闻推荐），经过搜索了解到，多模态对齐任务可能还是需要做。*

*Next：如何将新闻的文本图片信息进行提取、融合（实践，多模态大模型LLM与AIGC前沿技术实践）*

推荐系统综述

随着网络应用的不断发展，网络资源呈指数型增长，信息过载现象日益严重，如何高效获取符合需求的资源成为困扰人们的问题之一。推荐系统能对海量信息进行有效过滤，为用户推荐符合其需求的资源。对目前推荐系统研究现状进行了详细描述，介绍了基于内容的推荐、协同过滤推荐、混合推荐传统的三种推荐方式的研究进展和优缺点，并重点分析了基于卷积神经网络(CNN)、深度神经网络(DNN)、循环神经网络(RNN)、图神经网络(GNN)四种常见的深度学习的推荐模型的研究进展、优势、难点以及未来在应用中改进的方向，归纳整理了推荐领域常用的数据集，同时分析对比了传统的推荐算法和基于深度学习的推荐算法的差异。最后，总结了二者在实际应用中的代表性的推荐模型，讨论了推荐系统面临的挑战和未来的研究方向。

*基于内容的推荐（以用户历史选择中的内容参考，挖掘内容关联性高的项目）、基于协同过滤的推荐（评分矩阵的方式进行推荐）、混合推荐（前两种融合方法结合）深度学习的推荐：wide&deep RNN cnn gnn 新闻推荐的五步：用户特征的计算、用户和文章匹配度的计算、对推荐的新闻进行排序、去重、计算可能植入的广告*

*读后感：文章对推荐系统的研究进行了总结，深度学习的推荐系统可以处理大规模的数据集，具有更好的推荐效果。推荐过程有候选集生成（多分类），排序，我的研究内容是排序阶段，因此不考虑候选集生成。*

多模态推荐系统综述

推荐系统(RS)已经成为在线服务不可或缺的工具。它们集成了各种深度学习技术，可以根据标识符和属性信息对用户偏好进行建模。随着短视频、新闻等多媒体服务的出现，在推荐的同时了解这些内容变得至关重要。此外，多模态特征也有助于缓解RS中的数据稀疏问题。因此，多模态推荐系统(multimodal recommendations System, MRS)近年来受到了学术界和业界的广泛关注。在本文中，我们将主要从技术角度对MRS模型进行全面的综述。本文首先总结了MRS模型的一般流程和面临的主要挑战，然后分别从特征交互、特征增强和模型优化三个方面介绍了现有的MRS模型。为了方便那些想要研究这个领域的人，我们还总结了数据集和代码资源。最后，我们讨论了一些有希望的发展方向，并对本文进行了总结

多模态推荐的三大挑战：

1. 如何融合不同语义空间中的模态特征，得到每种模态的偏好-----特征交互
2. 如何在数据稀疏条件下获得推荐模型的综合表示----特征增强
3. 如何优化轻量级推荐模型和参数化模态编码器------模型优化

特征交互的方式：桥接（建立多模态消息传递通道---图）考虑采用知识图谱的方式进行特征交互。 融合（通过注意力机制来生成融合权重） 过滤（去噪）

*读后感：多模态推荐的主要任务是如何利用推荐系统的多模态数据，主要在于用户-物品的交互数据，其本质还是在于多模态融合技术。这篇文章比较完整的讲述了多模态推荐系统的融合方式，提供了多模态推荐的开源数据集及框架MMRec、Cornac。我的工作主要在于应对挑战1，选择合适的特征交互方式。*

*Next：选择合适的特征交互方式对新闻文本图片不同模态特征进行融合。考虑使用MMRec多模态工具箱，能否在其上改自己的代码。*

多模态融合与多模态推荐的区别

我：多模态融合是多模态推荐的一部分，将推荐系统的多模态数据进行融合。

百度：多模态融合：主要是将不同模态的数据进行融合，以获取更全面、更准确的信息，提高系统的性能。它主要关注数据的整合和融合。多模态推荐是推荐系统的一个子集，它专门处理来自多个模态的数据，并考虑如何利用这些多模态数据来提高推荐的准确性和效果。

个性化新闻推荐方法研究综述

个性化新闻推荐是帮助用户获取其感兴趣的新闻信息和缓解信息过载的重要技术。近年来，随着信息技术和社会发展，个性化新闻推荐得到了日益广泛的研究并在改善用户的新闻阅读体验方面取得了显著成功。本文旨在对基于深度学习的个性化新闻推荐方法进行系统性综述。首先，分类介绍了个性化新闻 推荐方法并分析各自特点及影响因素;然后，给出了个性化新闻推荐的总体框架并对基于深度学习的个性 化新闻推荐方法进行了分析总结;在此基础上，重点综述了基于图结构学习的个性化新闻推荐方法，包括基于用户-新闻交互图、知识图谱和社交关系图的新闻推荐;最后，分析了当前个性化新闻推荐所面临的挑战，探讨了如何解决个性化新闻推荐系统中数据稀疏性、模型可解释性、推荐结果多样性和新闻隐私保 护等问题，并在未来研究方向中展望了更具体可操作的研究思路和方法。

个性化新闻推荐的时间、位置、社交网络因素------新闻推荐受多种因素影响，可以对每种因素进行分析建模。个性化新闻推荐目的在于挖掘用户与新闻的关系，准确的匹配用户兴趣和候选新闻。新闻建模---自然语言处理（Bert）、用户建模---用户历史行为（LSTM）。未来研究方向：多源异构、多源融合、可解释性。

*读后感：个性化新闻推荐属于推荐系统的个例，偏向于自然语言处理，对于新闻文本的建模，文章主要是对基于图结构方法的个性化新闻建模进行分析。图结构的个性化新闻推荐能够利用图的边和节点特征，在处理多源数据上具有优势，如社交网络、位置信息。新闻推荐是常见的一种推荐系统，文章是否好出待商榷。论文首先针对新闻的多模态交互，再引入知识图谱。*

*Next：针对新闻的多模态数据，如何对新闻及用户进行建模。*

*Deep Learning for click-through rate estimation*

*点击率(CTR)估算是各种个性化网络服务的核心功能模块，包括网络广告、推荐系统、网络搜索等。从2015年开始，深度学习的成功开始有利于CTR估计的性能，现在深度CTR模型已经广泛应用于许多工业平台。在本调查中，我们对CTR估计任务的深度学习模型进行了全面的回顾。首先，我们回顾了浅CTR模型向深CTR模型的转变，并解释了为什么深CTR模型是发展的必然趋势。其次，重点研究了深度点击率模型的显式特征交互学习模块。然后，作为具有丰富用户历史的大型平台的一个重要视角，讨论了深度行为模型。此外，还介绍了近年来出现的深度CTR结构设计的自动化方法。最后，对调查结果进行了总结，并对该领域的发展前景进行了展望。*

*LR POLY2 GBDT FM W&D deepFM autoFIS*

*双塔 单塔模型*

推荐系统研究进展与应用

推荐系统一直以来都是很活跃的研究方向。回顾了近 5 年来推荐系统涉及的数据集、算法、效果评估的研究发展状况，给出了推荐系统在互联网应用场景中的系统架构，从工业知识与经验以及工业生产工艺与工艺参数这两方面对推荐系统 工业生产中的最新应用做了介绍。分析了推荐系统中“信息茧房”问题产生的原因，展望了未来推荐系统在“信息茧房”、可解释性、隐私性等方面的研究。

总结：推荐算法主要是针对用户和物品的交互数据进行建模，涉及到非常多的流程：召回、过滤、排序、混排、强规则。（我的工作不在召回，过滤，应该在后面三个。对于不涉及的流程，就不再考虑（合理），抓住工作重点）。对什么样的数据进行怎么样的处理（数据的形式，需要提出的特征，计算的方式）？用户数据，图文信息流，用户行为数据，新闻分类数据。。。推荐算法的发展从传统的到深度学习的，逐步解决用户冷启动的问题与数据稀疏的问题，从低阶特征到高阶特征，挖掘更多特征的关联性，从短期到长期特征。多模态是现在的一个研究热点，图片和文本存在一定的关联，可以互相辅助，目前关于多模态的研究跨模态对齐、多模态融合、多模态迁移，多模态对话，自监督的多模态等研究领域，多模态融合主要在于解决模态间的语义冲突，重复，噪声。主要步骤分为：数据预处理、特征提取、模态融合。（我虽然要做前期的工作，但主要方向是在模态融合上）

基于深度学习的多模态新闻推荐系统研究与实现

研究内容：如何提取图像的语义信息并将其整合到自然语言处理模型中，一般情况下是直接提取图像的ROI，但ROI的数量往往较少，会带来语义信息的丢失。文章是基于transformer和Bert模型上的改进。

＠3.4 今天花时间把相关的几篇论文结构摘抄在纸上了，明天根据每篇论文的引言理一理他的详细结构。开始引言的撰写。

VLSNR：文章提出了一个自己的数据集，模型引入了GRU网络，获取新闻时序特征。贡献

我们提出了VLSNR框架，该框架集成了时间序列中新闻的视觉和文本信息，以学习点击偏好趋势。

VLSNR能够充分表征新闻的整体特征，更好地模拟用户的行为。V-MIND的MIND数据集，这是规模最大的。