**Collaborative Knowledge Base Embedding for  
Recommender Systems**

# ABSTRACT

在不同的推荐技术中，由于用户项目交互的稀疏性，协同过滤通常受到有限的性能的影响。为了解决这些问题，通常使用辅助信息来提高性能。由于Web上信息的快速收集，知识库提供异构信息，包括具有不同语义的结构化和非结构化数据，可由各种应用程序使用。在本文中，我们研究如何利用知识库中的异构信息来提高推荐系统的质量。首先，通过利用知识库，设计三个组件，分别从结构内容，文本内容和视觉内容中提取项目的语义表示。具体来说，我们采用异构网络嵌入方法，称为TransR，通过考虑节点和关系的异质性来提取项目的结构表示。我们应用堆叠去噪自动编码器和堆叠卷积自动编码器，它们是两种基于深度学习的嵌入技术，分别提取项目的文本表示和视觉表示。最后，我们提出我们的最终综合框架，称为协作知识库嵌入（CKE），共同学习协作过滤中的潜在表示以及知识库中项目的语义表示。为了评估每个嵌入组件以及整个系统的性能，我们对来自不同场景的两个realworld数据集进行了广泛的实验。结果表明，我们的方法胜过多种广泛采用的最先进的推荐方法。

# INTRODUCTION

由于信息的爆炸性增长，推荐系统在网络服务方面发挥越来越重要的作用。在不同的推荐策略中，利用历史交互或偏好的协同过滤（CF）方法取得了重大成功[23]。然而，当用户项目交互非常稀疏时，CF方法通常会受到有限的性能的影响，这对于诸如在线购物（其中项目集非常大）的场景是非常常见的。此外，CF方法不能推荐新项目，因为这些项目过去从未收到用户的任何反馈。为了解决这些问题，结合协同过滤和项目内容等辅助信息的混合推荐系统通常可以获得更好的推荐结果，近年来越来越受欢迎[2]。

在过去的几年中，越来越多的语义数据被公布遵循链接数据原则1，通过将统一的全球数据空间中的人员，书籍，音乐，电影和地理位置等不同主题领域的各种信息连接起来。 这些异构数据彼此相互联系，形成了一个名为知识库的庞大信息资源库。已经建成了几个典型的知识库，包括学术项目如YAGO2，NELL3，DBpedia4和DeepDive5，以及商业项目，如微软的Satori6和Google的知识图表7。 使用来自知识库的异构连接信息可以帮助开发关于难以发现的数据的问题的见解单个域[6]。 到目前为止，信息检索[9]，社群检测[25]，情绪分析[4] - 仅举几例成功利用知识的值得注意的应用程序基础。