

homework1 小论文片段

刘万宇

May 2024

1 Introduction

在现今信息技术爆发式发展的时代，人们分析信息的技术也是日新月异。但是当我们使用各种各样的技术对信息进行分析的时候，应该去注重信息的各个方面，这样才能够得到全面的分析结果。同理，为了让某个用户得到一份满意的服务匹配答卷，我们就需要对该用户进行个性化分析。但是，这些方法很难去真正考虑到用户的各个模态信息，举这样一个例子：某位用户想要看电影，他非常喜欢动作片巨星成龙，尽管他自身并不是很喜欢看动作片，但是最后他却很乐意为自己喜欢的演员买单。然而对于另外一位用户，尽管他很喜欢演员成龙，但是他非常不喜欢动作片这类的影视作品，所以就算这部动作片是由成龙主演的，他也不会去看这部影片。

知识图谱作为用于知识信息表示的载体，通过“实体—关系—实体”形式的三元组来存储一条知识，通过构建这样海量的三元组，使得我们可以在日常生活中搜索各类需要的信息。而使用知识图谱来存储多模态信息，通过将每个节点作为一种实体，每条边表示一个关系这种图结构，可以完美地解决多模态信息的异构性这个问题，就像图 1 所显示的那样 不仅如此，在大数据时代，对于一些漏失错误的用户数据，一般情况下，我们需要有应对手段去处理以防止可能出现的错误结果。但我们若是构建如图二所示的用户知识图谱，由于该用户没有上传照片数据导致图片信息缺失，但是得益于信息存储是知识图谱这种结构，所以对用户数据信息的分析不会因为缺失数据产生比较大的影响，因为整个图结构依然是完整的。而对于错误数据的处理，我们只要以用户为中心，检索与用户相邻的节点这类简单方法就能对该用户的错误信息进行修正。如图 2 所示：用户 2 的生日是 2001/2/08，那么，其年龄应该是 23 并非 22，基于这样的结构我们就能很快使用合适的方

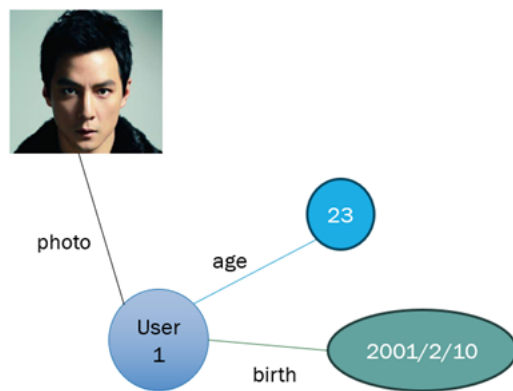


图 1: 用户知识图谱的一个简单示例

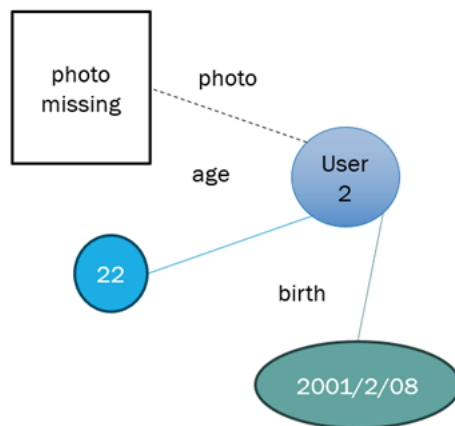


图 2: 用户知识图谱对于用户缺失错误数据的示例

法对其进行我们所构建用户知识图谱中的存在的错误信息进行纠正。由此在某种程度上，这种结构能体现出各个模态之间的互补性。

2 Related Work

略

3 Proposed Method

3.1 总体架构

图 3 展示了我们所构建的用户多模态知识图谱的总体架构。对于一个用户来说，我们的所构建的用户多模态知识图谱由两个部分组成。左边橙色虚线框中是该用户个人的多模态知识图谱，主要用于对用户个人一些多模态信息进行存储，用户个人的知识图谱主要以“用户为中心”，首先“引出”用户与各个模态部分的联系，与用户自身相衔接的始终是各个模态的 summary node, summary node 起到类似于“引路标”的作用，帮助我们指引到用户某个模态信息的具体偏好或是用户的某个具体属性。Harry Li 等人在不久前的报告中曾指出 KG Consumer 更加偏好简单的可视化结构 (如：简单的表格而非复杂的 NLD 结构)，并且我们认为当我们对用户各个模态信息进行分析的时候，我们可以先不去考虑各个具体信息之间的联系，我们需要得到的仅仅是受分析用户的个人属性，以及该用户所喜欢或是厌恶的事物，当我们得到这些比较直接的多模态信息之后，之后可以再使用如维基百科等大规模知识图谱进行分析。基于以上这两点，当我们构建用户个人的知识图谱的时候，我们可以不用考虑各个模态信息的具体联系。例如：某位用户的国籍是中国，出生地是北京，那么，在构建用户个人层面的知识图谱的时候，中国与北京两个实体之间不需要一条表示“...the capital of”的关系边相连，当我们通过该用户的知识图谱表示提取到了“中国”，“北京”这两个实体之后，再对这两个实体进行分析。

3.2 自然语言规约

略

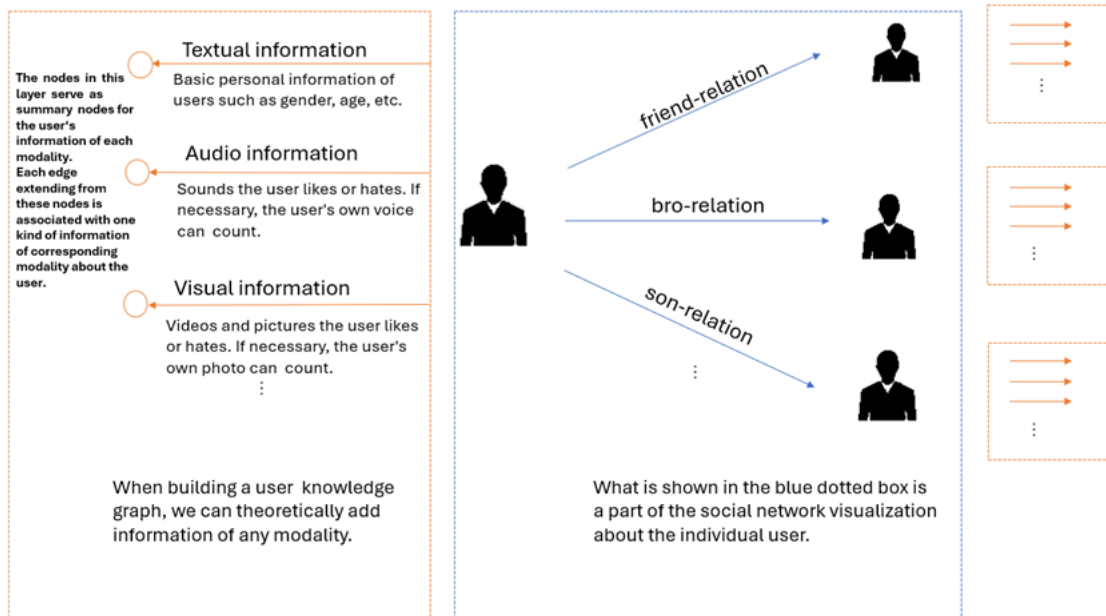


图 3: 用户知识图谱的总体架构

3.3 存储结构

略

3.4 TransU Model

Jason Weston 等人提出的 TransE 模型是知识表示模型中最有影响力的模型之一，TransE 模型最大的优势在于简单且容易扩展，但是其模型的缺点也很明显：比如存在多个实体在 embedding 空间中竞争一个点的情况，不适合处理一对多的关系，并且完全没有考虑语义信息，因而后来产生了一系列模型对 TransE 进行改进和补充。对于构建用户多模态知识图谱也是一样，用户喜欢的某样事物可能不止一样，但是我们却注意到了以下几点：

1. 所有构建知识图谱的三元组都是单向的，也就是说，我们只需要关注用户对于各个除自己外的实体的偏好以及用户自己的个人属性，不需要去考虑除用户自己外的实体对用户个人的影响。
2. 从 PKG 的实用性以及从简性考虑，在构建 PKG 的时候，我们将非用户实体之间的联系舍弃了，而非用户实体之间的联系可以从 PKG 中提取出来

之后再考虑。

3. h_u 在整个用户个人多模态知识图谱中是独特的实体即用户本身，换句话说，我们所形成的用于构建的用户个人多模态知识图谱的三元组 head 是可以不用变化的。

4. 基于第 3 点，当对于某用户喜欢的某类事物不止一样时，那么我们形成用户偏好三元组 (h_u, r_u, t_u) 中的 h_u 与 r_u 是可以不变的，那么，我们只需要变化 t_u ，就可以形成新的用户偏好三元组。

考虑到以上几点，当我们在进行所构建的用户个人的知识图谱便十分简单且明了，同时，由于非用户的实体之间的联系被舍弃了，因此我们在 Embedding 的时候便不必对各个模态的信息进行同维度投射的处理。我们受到了 Andrea Frome 等人在构建 DeViSe model 时所使用的一种方法的启发，Andrea Frome 等人为了得到近义词的词向量表示从而使用了 Skip Gram model，但是，他们训练 Skip Gram model 的目的并不是为了使用这个模型，而是为了得到经过训练之后的隐层矩阵。同理，我们使用的模型并不是为了得到每个实体 Embedding 的结果，而是为了得到各个关系以及 Summary Node 的 Embedding。当得到了 Summary Node 以及各个关系的 Embedding 之后，我们会使用偏好三元组对每个非用户实体进行构造而非直接的 Embedding 化操作。该公式表示我们构造的思想：

$$h_u \cdot M + k \cdot r = e_k \quad (1)$$

其中， h_u 是用户在社交网络中的 Embedding，由对该用户的社交网络进行 Embedding 操作得到；M 是模态投影矩阵，可以将代表用户的 Embedding 投影至某一个模态维度中，在例图中，我们可以将 $h_u \cdot M$ 理解成为该模态下的 Summary Node；r 是某种关系的 root embedding，k 是一个线性变化量，当满足该关系的用户实体不止一个时，k 会根据用户对于某种事物喜欢不同程度进行相应变化， e_k 则是最后我们所构造 PKG 中关于非用户实体的向量表示。

4 Expriment

略