



搭建图书推荐系统

基于一个简单的原理构建一个书籍推荐系统：书籍有共同链接的是相似的，或者是有关联的。为了表示书籍，将使用神经网络实体嵌入（neural network entity embeddings）的概念，将每本书和每个Wikipedia链接（Wikilink）映射到 N 维的向量。

步骤：

1. 加载数据并清洗
2. 准备训练数据
3. 构建实体嵌入神经网络
4. 训练神经网络
5. 提取嵌入内容并找到最相似的书籍

数据来源：<https://dumps.wikimedia.org/enwiki/>

例子：[War_and_Peace](#)

训练集：(book, link, label)，将这一本书中有的链接，label为1，随机选择没的link进行组合，label为0。

分别训练 book 和 link 的embedding，再将它们相乘。标签比较。可分类，可回归。最终得到的模型参数就是书籍的embedding。输入一本书，可返回相似的10本书。