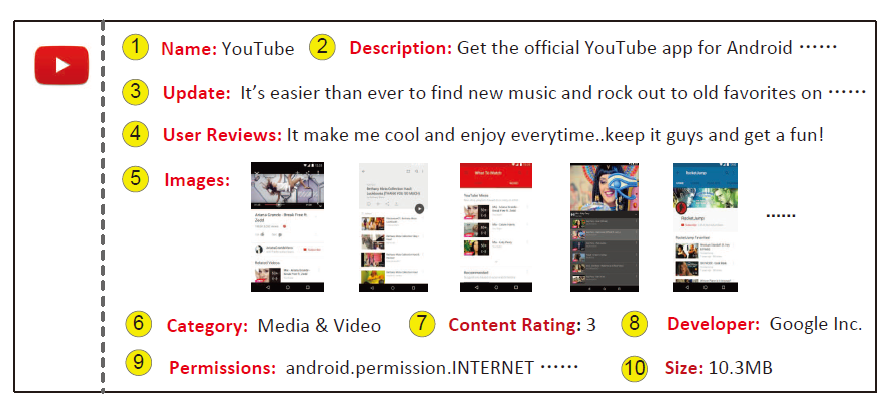
近期我主要阅读了一些推荐领域的论文，下面简单进行说明。

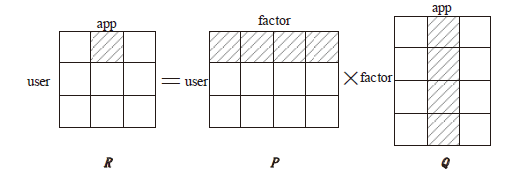
1. **Mobile App Tagging**

这是一篇16年WSDM上的论文，主要是针对于APP中普遍存在没有标签的这一现象提出。本文基于search-based 标记方法，该方法先被用于CV领域，思路是先通过相似性算法找到N个相关的内容，然后利用这N+1个内容进行标记。在训练时使用了Online learning方法，该方法将数据逐步输入，而不是直接使用整个数据集。本文对相似度计算考虑了10个不同维度的特征，如图所示：

由于模型年代的问题，该模型中的词嵌入使用的是word2vec，相对而言，效果必然会差一些。

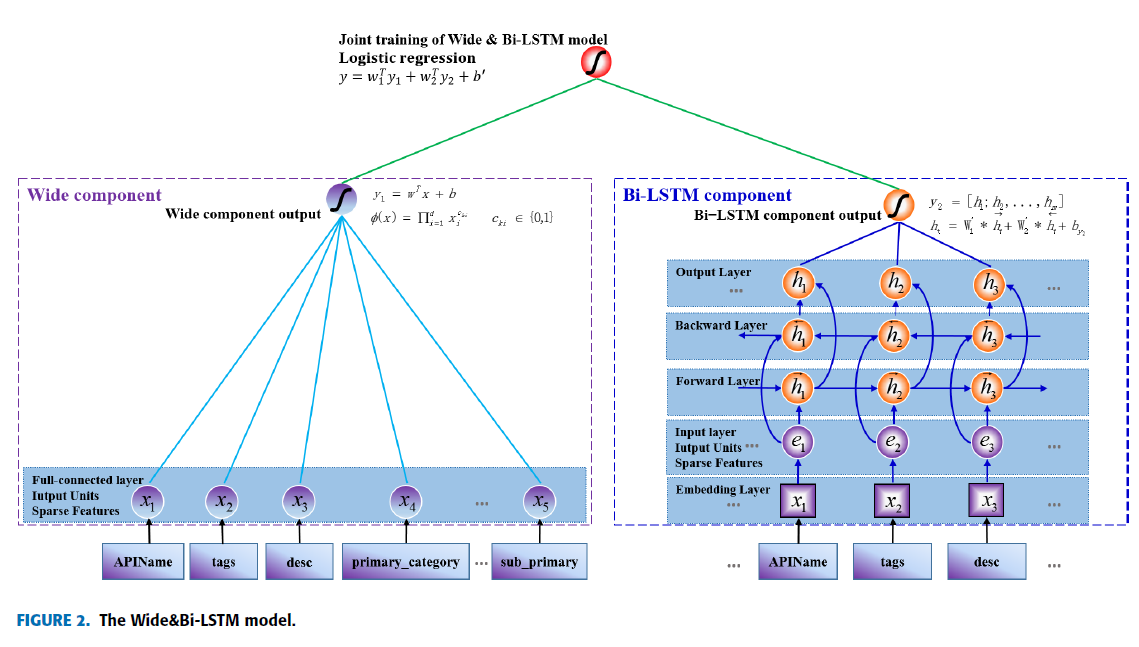
1. **A Hybrid Approach based on Collaborative Filtering to Recommending Mobile Apps**

这也是一篇2016年的论文，该论文提出了一种基于协同过滤算法的模型，该文章中提出读于APP而言，下载量不能很好的代表用户的偏好度，并提出了RDF（recency，frequency，duration）来衡量用户的真实偏好度。同时，该方法基于item进行推荐，所以该方法预测时，考虑了与该应用相似的几个应用。同时，本文考虑极端情况下，某些用户不存在使用记录的情况。协同过滤算法如下：



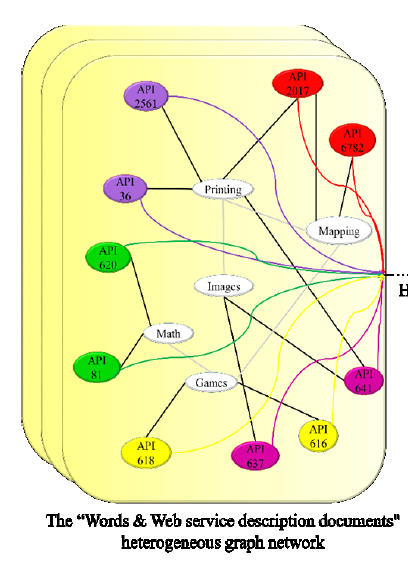
1. **Web Services Classification Based on Wide & Bi-LSTM Model和A Web Services Classification Method based on GCN**

第一篇是2019年IEEE Access上的论文，第二篇是2019年ISPA上的论文，这两篇论文呢都是针对web服务分类任务。前者提出web服务的描述信息一般简短，拥有的特征少，并且之前的方法未考虑到词序和上下文的信息。后者认为现有方法都忽视了web服务描述文本和web描述之间存在关联的问题，并提出了一种异质图。 前者模型如下：



该方法基于Google提出的Wide&Deep模型，是推荐领域经典的模型。本文模型也非常简单，就是一个MLP和BI-LSTM的结合。现在看来使用Transformer改造一下效果应该会好。

后者主要贡献是提出了异质图，如下：

该图包含有web服务节点和描述文本中进行分词处理后的word节点。模型就是GCN。

1. **Web Services Clustering Based on HDP and SOM Neural Network**

这是2018年的一篇文章，该文章比较简单，主要就是使用了两种新的方法进行结合，并对以往工作进行一些改进。主要是HDP未LDA改进了，使其不用手动调整主题数，和SOM进行该进是聚类算法可以自主学习类别数。本文有意思的一点是提出了对于文本信息少的时候，使用word2vec方法中的相似度计算，得到很多语义相近的词对描述文本进行扩充。