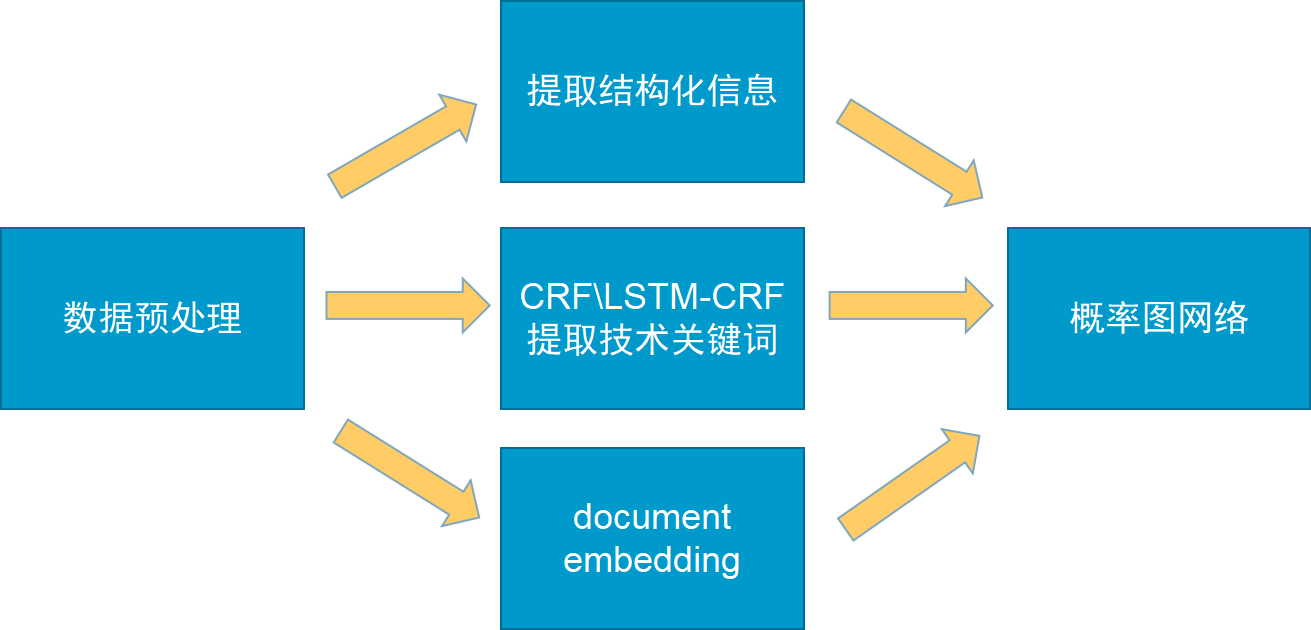
1 绪论

1.1 课题背景及目的

1.2 国内外研究状况

（从开题报告来）

1.3 论文构成及研究内容



1. 推荐算法的设计与实现

2.1 背景知识

2.1.1 CRF命名实体识别

2.1.2 LSTM/CRF 命名实体识别

2.1.3 RBM 的原理和应用

2.1.4 word embedding 和 document embedding

2.1.5 概率图模型的简介

2.2 数据预处理

2.2.1 数据的格式

2.2.2 数据的初步处理

2.3 特征提取

2.3.1 基于正则表达式的格式化信息提取

2.3.2 关键词提取

2.3.1.1 CRF

2.3.1.2 LSTM-CRF

2.3.3 RBM 提取one-hot编码的关键词特征

2.3.4 document embedding提取短文本信息

2.3.4.1 wiki 中文预料预训练模型

2.3.4.2 简历文本fine-tune模型

2.4 概率图模型的设计

2.4.1 模型的输入输出

2.4.2 模型架构

2.4.3 模型的训练

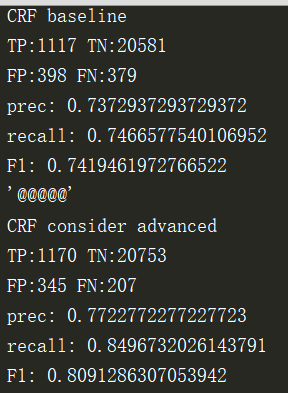
1. 推荐算法的评估

3.1 技术关键词的结果分析

3.1.1 评价指标

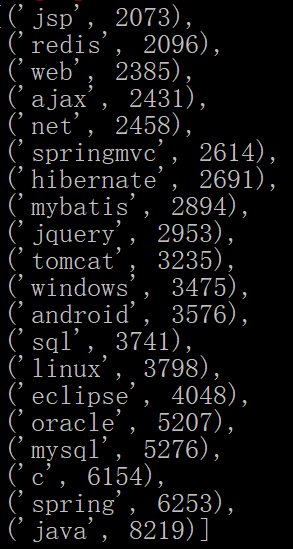
3.1.2 关键词提取结果

extract from 5636 JL data from 51job





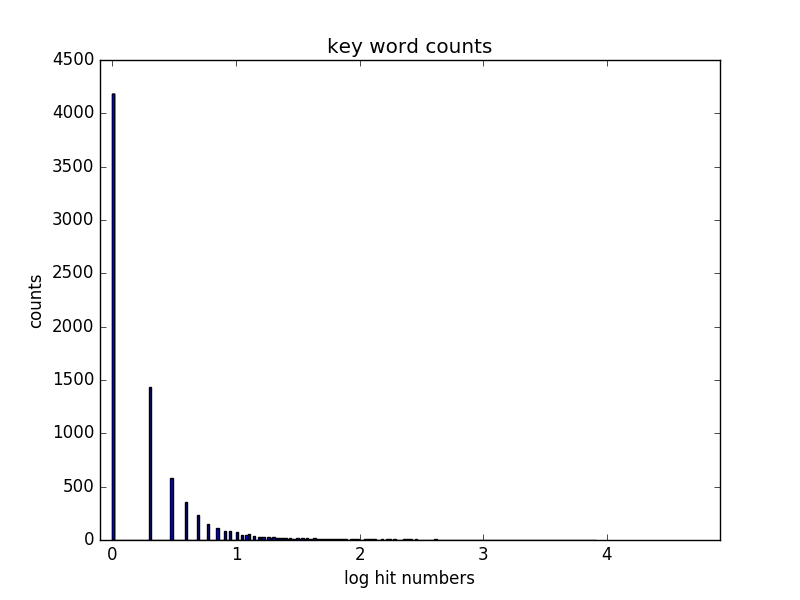
关于keyword 的统计：



一共找到8251个关键词，平均一个关键词hit 23次

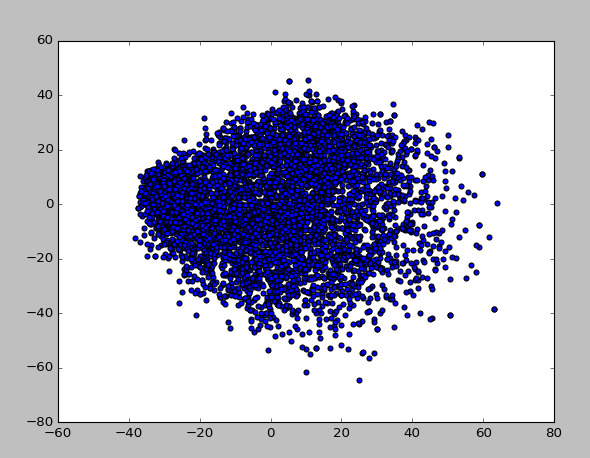
困难：错误拼写：

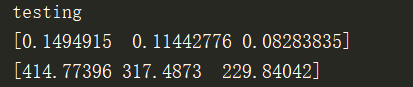


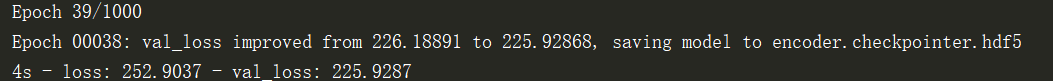


remove all the word which hit times<=3, left 2058 key words

auto encoder之后，保留最大的两个主成分：







3.1.3 LSTM的结果：

3.2 推荐算法在测试数据集上的结果

3.2.1 评价指标

3.2.2 结果与分析

1. 结论与展望
2. 致谢
3. 参考文献