# BigRLab线上服务总结

## Knn求最近邻

使用Annoy tree

## Article

Jieba提供的服务有：

* 分词 wordsegment
* 关键词抽取 keyword

Article service 提供knn服务，求最相似的文章的label和score

## Xgboost

用于做特征工程

两个model，model1用于得到leaf编号，model2用于预测带leaf编号特征的数据。详见文档 xgboost.docx

## 打标签

包括tagging和改进的tagging2

基于关键词权重和共现词权重。

## 话题预测 topics