

个人信息				
姓 名：杨海天	性 别：男			
出生年月：1988 年 4 月 7 日	学 历：硕士			
政治面貌：党员	籍 贯：河南新乡			
专 业：计算机应用技术	研究方向：文本挖掘、信息检索			
外语水平：CET-6 480	邮 箱：584530361.qq.com			
手 机：13342289551	求职方向：研发工程师			
毕业时间：2014 年 7 月	身高：178cm			
教育背景				
➤ 2011.9- 至今	大连理工大学	计算机应用技术	硕士在读	
➤ 2006-2010 年	洛阳师范学院	计算机科学与技术	本科	
专业技能				
<div>➤ 熟练掌握JAVA编程技术和面向对象思想以及J2EE平台开发技术；</div> <div>➤ 具有大型 web 开发经验，具有真实的 B/S 应用开发经验</div> <div>➤ 研究方向为信息检索和文本挖掘，掌握常见的文本挖掘方法和自然语言处理技术；熟悉搜索引擎平台搭建的框架和流程，熟悉常用的推荐算法，基于用户的协同过滤算法和基于物品的协同过滤算法；</div> <div>➤ 熟悉MySql数据库的操作，熟练编写SQL语句；</div> <div>➤ 熟练使用JDBC连接数据库进行数据的CRUD操作；</div> <div>➤ 熟练使用MyEclipse、PowerDesigner、Rational Rose等开发工具；</div> <div>➤ 熟练使用Hibernate完成对象到关系型数据库的映射，使用Hibernate技术实现DAO，掌握持久化对象生命周期管理，使用Hibernate进行事务管理，Hibernate查询语言；</div> <div>➤ 熟悉Tomcat 6.0 web服务器，熟练掌握Struts2的基本开发流程，使用Struts2进行MVC模式的Java Web 开发；</div> <div>➤ 熟练编写基于Spring 轻型JavaEE容器的应用，实现与Struts2、Hibernate的整合；</div> <div>➤ 熟练掌握 JSP 常用标签、EL 表达式及 JSTL 标签库的使用；</div> <div>➤ 有基于 SVN 版本控制系统的项目开发经验，熟悉 JUnit 4 单元测试</div>				
研究生参与代表性项目				
<div>➤ <b>XX 数码网购平台</b></div> <div><b>项目概述：</b>该系统一个类似众多购物网站的系统，并根据现有的网上购物系统的现状而设计开发的电子商务平台。主要完成：对各类电子产品进行品牌，类别，最新商品，最热商品进行分类与显示；商品管理员对商品、订单、用户的管理；注册用户的在线购物、管理购物车、查看订单。</div> <div><b>责任描述：</b>参与了该系统的详细需求设计，对各个模块的具体实现细节做了深入分析与设计，并完成了该系统的管理员模块、商品管理模块、订单模块；对各模块都加入了增、删、改、查的功能，方便管理员的操作。</div> <div><b>项目技术：</b>项目采用 MVC 设计模式，基于 Struts2+Hibernate+Spring 整合。View 层用 JSP，Controller 用 Struts2，持久层采用 Hibernate 技术；用 Ajax 实现异步刷新功能，解决了全屏刷新资源消耗大，及出空白页面等待的问题。</div> <div><b>XX 公司 OA 办公系统</b></div> <div><b>项目概述：</b>该项目中的权限管理模块采用的是树的展现形式，用到的 zTree 框架插件生成树。以及通过采用 sturts2 与 ajax 相结合的方式，做到了完全的无刷新展现。采用 JS 面向对象方式编写。在审批流程模块中使用的是 jbpm 工作流框架，利用其 API 完成了审批流程的功能。在查询过程中采用的 hibernate 的缓存机制和抓取策略进行数据查询的优化。其中还利用 spring 的 AOP 做了一个异</div>				

常中转站，避免了当 service 层发生异常时跳转到 Sturts2 的错误模板页面。在代码编写的过程中采用了面向接口编程的方式。对公共代码的重用做到了很大程度的封装

**项目收获：**我在此项目中主要做了部门管理，岗位管理和用户管理。熟悉了数据库的多对多，一对多设计。

➤ **基于 VSM 模型离线信息检索系统，垃圾邮件过滤系统**

**项目概述：**搭建一个简单的信息检索系统，包括分词，去停用词，词干化和倒排索引的建立都是独立的完成。建立向量空间使用 VSM 模型进行计算文档相似度。特征选择使用信息增益的方法和经典的 IFIDF 方法。实验结果是信息增益的方法好点，本项目是我们实验室入学时每个新生都必须独立完成的项目。垃圾邮件过滤系统是用贝叶斯分类做的，不依赖外部工具，通过训练集训练特征，同时经过互信息进行特征选择，从而进行垃圾邮件的分类。

**项目收获：**通过基于 VSM 模型离线信息检索系统的训练使我对整个检索系统的整体流程有了深刻的认识。对基础的 NLP 知识有了初步的认识。通过垃圾邮件的分类对分类的任务有了初步的了解。

➤ **基于问句类别的问句检索系统**

**项目概述：**基于问句类别信息问句检索的方法，主要利用问句的类别信息对语言模型进行平滑，同时融入问句的语义信息来解决问句检索问题。该项目最终在 Yahoo! Answers 上抽取的真实标注数据集上进行，对比实验结果表明，本系统所采用的方法在性能上得到了较好的结果。并且该项目所采用的算法已经写成论文，投到期刊 KAIS(知识和信息系统)，目前在二审。

**项目收获：**熟悉了传统的语言模型，修改了原始语言模型的公式融入例如问句的类别信息和语义信息

➤ **基于特征融合入的问句检索系统**

**项目概述：**基于特征融合的问句检索的方法，主要利用问利用答案特征、词序特征、统计特征和语义特征相结合来解决问句相似度计算问题。该项目最终在 Yahoo! Answers 上抽取的真实标注数据集上进行，对比实验结果表明，本系统所采用的方法在性能上得到了较好的结果。并且该项目所采用的算法已经写成论文，投到 CCIR2012，目前已录用。

**项目收获：**熟悉了特征选择算法，尝试了各种特征融合的方法。熟悉了 wordnet 的 API 和句法分析器的 API，能够准确的获取语义信息和句法特征

## 学术情况

- <<基于特征融合的社区问答问句相似度计算>>被 CCIR2012 录用，第一作者
- <<一种基于主题类别信息问句检索的新方法>>投计算机应用与软件
- <<Improving Question Retrieval by Using Topic Information and Answer Information>>投 KAIS
- <<Improving Question Retrieval by Exploiting Category and Answer Information>>投 KAIS
- 研究模型：分类：SVM 聚类：LDA, K-means, KNN, 网页排序算法：PageRank, HITS

## 获奖情况及社会活动

- 2013 年 6 月参加百度电影推荐大赛
- 2012 年 11 月获得大连理工大学优秀研究生
- 2012 年 9 月获得大连理工大学研究生二等奖学金
- 2012 年 9 月~2012 年 12 月大连理工大学计算机组成原理助教
- 2011 年 9 月获得大连理工大学研究生二等奖学金

## 自我评价

- 自信、忠实、敬业、能吃苦，适应能力强、可以加班和适应出差工作
- 具有很强的团队精神，有良好的组织、协调和沟通能力，有强烈的集体荣誉感；
- 自学能力强，喜欢钻研新技术，有比较强的动手能力和很好的分析问题与解决问题的能力