**XXX大学XXX学院课程实验报告**

|  |  |  |  |  |  |
| --- | --- | --- | --- | --- | --- |
| 姓 名 |  | 学 号 |  | 学 院 |  |
| 年 级 |  | 专 业 |  | 邮 箱 |  |
| 课 头 号 |  | 课程名称 |  | 指导教师 |  |
| 实验类型 | □独立实验课 ☑含实验的理论课 □计划外自选开放实验  □学生自主式开放实验 □大学生科技竞赛 | | | | |
| 实验名称 | 科技文本挖掘概论与ItgInsight入门 | | | | |
| 实验日期 |  | 实验学时 | 2学时 | 成 绩 |  |
| **一、实验概况**  **（一）实验目的与内容**  1、掌握科技文本挖掘的概念、目的、操作流程。  2、掌握科学知识图谱、知识图谱、可视化、知识管理的关系，ItgInsight与科学知识图谱的关系，ItgInsight与知识管理的关系。  3、掌握科技文本分析软件工具/科学知识图谱软件工具的分类体系，各类软件工具的特点，闭源/商业软件与开源免费软件的差别。  4、掌握ItgInsight下载、安装、升级、资源获取方式；掌握ItgInsight页面颜色、语言设置，认识ItgInsight的页面构成，窗口显示和隐藏操作方法。  5、掌握ItgInsight数据读取、清洗、分析、报告的基本流程。  6、掌握ItgInsight过滤器的原理，常见科技文本数据源的过滤器；掌握分隔符的定义和设置；明确主题词的概念和提取来源；明确停用词表的作用。  **（二）主要仪器设备（含必要的元器件、工具）：**  1、操作系统  Windows 10及以上版本操作系统  2、工具软件  ItgInsight、Notepad++、Microsoft Office系列  3、数据集（以下数据集选一即可）   * 选定自己感兴趣的领域的SCIE/SSCI论文、知网论文或者其他数据集合； * 2012-2021年期间被SCIE/SSCI收录的图书情报领域的研究论文（Article）集合；   **（三）实验内容**  1、应用ItgInsight读取实验数据源，分别生成\*.dataset，\*.bigdataset，\*.xlsx数据集，观察文件构成；生成\*.itgn,\*.mod分析结果文件，观察并输出文件内容。  2、\*.dataset，\*.bigdataset文件的增量数据读取，数据补充，保存，二次打开利用。  3、\*.bigdataset文件，采用数据库管理软件进行打开，内容查看，二次利用。  4、\*.itgn，\*.mod文件保存，二次打开利用。  \*拓展题：  1、编制过滤器，读取软件文件“example\data\自然基金数据”的项目数据。  2、编制过滤器，读取软件文件example\data子目录下名为“example-data-patsnap”的专利数据，从题目和摘要中提取主题词。 | | | | | |
| **二、实验操作部分（可续页，附上文字说明和重要内容截图）**  1、  2、  3、  4、  5、  6、 | | | | | |
| **三、实验效果分析**  **（一）实验完成情况**  **（二）实验室完成（未完成）的总结** | | | | | |
| **四、实验反思与反馈**  **（一）本次实验的心得体会**  **（二）本次实验的提升建议** | | | | | |
| **教师评语** |  | | | | |