

数据挖掘实训方案

一、实训背景

随着互联网普及和移动互联网的突飞猛进，人们已然进入信息爆炸时代，人们已不再为如何获取信息而苦恼，但如何对已获取的信息进行有效处理，即在现有数据资源上进行“数据深加工”，进一步从中挖掘出有价值的内容却困扰着无数人。在此背景下，数据科学人才也顺势成为当今社会最受欢迎的群体之一，而其中又以数据分析与挖掘人才的需求最为迫切。

但现实是，一方面大量社会企业很难寻找到合适的数据分析及挖掘人才，另一方面高校对于如何培养出符合企业需求的人才又深感困扰，因此如何针对性提升学生数据挖掘实战水平，并使其更符合社会实际需要，是本次实训的主要目的。

二、实训意义

1、弥补学校在数据挖掘实战方面的教学短板，此次实训以企业真实案例为背景，让学生充分感受到真实的数据挖掘过程及其所需要的技能和知识点，对于学生的就业择业提供帮助。

2、本次实训可转换成学校教学成果，学校可以将此次实训过程吸收进往后的教学过程中去，或者成立学生工作室，促进学生在数据挖掘及建模方面的持续发展。

三、预期效果

- 1、学生对数据挖掘的概念及过程形成比较清晰的认识；
- 2、了解数据挖掘岗位目前的就业形式和前景，了解需要掌握的技能；
- 3、掌握一定的挖掘技能和工具，体验一个实际项目的全过程；
- 4、在项目经理的带领下能完成一个简化版的企业项目；
- 5、协助学生对自己后续的职业发展做出相对合理的规划与定位。

四、实训安排

实训共两周，分两个阶段，第一阶段 5 天时间，以工程师授课加学生上机为主，此阶段学生会接触到数据挖掘方面的专业知识、挖掘工具及实际应用等，课程设置详见附件（具体课程内容可根据学校具体需求进行一定调整）。

第二阶段 5 天时间，以学生自己动手做项目为主，工程师从旁指导协助，学生动手项目与第一阶段讲解的案例相似，是一个企业真实项目的简化版。

五、实训考核

为提升培训效果，促进学生对培训内容的吸收，我们特设立项目考核，学生 3 人一组在实训结束前须提交一份项目报告，并进行答辩。报告包含一份项目答辩 PPT，项目过程当中的程序代码、相应数据文件等内容。

每组答辩时间为 10 分钟，其中学生陈述 5 分钟，老师提问 5 分钟，评分标准如下表：

实训报告评分标准

学员姓名	分析报告的规范性、合理性等核心指标（40分）	PPT 质量、表达主次、演讲表述（30分）	基础理论知识（20分）	合理化建议、培训纪律（10分）	总分
学员 A					
学员 B					
学员 C					

六、实训地点

实训学校内或由实训学校指定的实训场地。

附件：实训课程表模版

日期	课程内容	课时数			备注
		课堂教学学时	是否多媒体教学	时间	
Day1	数据挖掘概论	3	√	9:00-12:00	
	数据探索/数据预处理	3	√	14:00-17:00	
Day2	R 基础介绍、数据的导入导出	3	√	9:00-12:00	
	R 数据结构及数据基本管理	3	√	14:00-17:00	
Day3	R 控制流、基础绘图	3	√	9:00-12:00	
	数据挖掘常用算法介绍	3	√	14:00-17:00	
Day4	案例：企业全真项目	6	√	9:00-12:00 14:00-17:00	
Day5	项目案例上机实验	6	√	9:00-12:00 14:00-17:00	
Day6~ Day9	学生实操、项目报告撰写	24	√	9:00-12:00 14:00-17:00	
Day10	项目答辩	6	√	9:00-12:00 14:00-17:00	