



西京学院
XIJING UNIVERSITY

《机器学习》课程 实验指导书

《机器学习》课程实验指导书

一、课程基本信息

课程名称：机器学习		课程编码：08090152001
课程性质： <input type="checkbox"/> 专业基础课程 <input checked="" type="checkbox"/> 专业方向课程 <input type="checkbox"/> 专业任选课程		
课程学分：3 学分	学时：总学时 48。其中，理论：32，实验：16	适用专业： 计算机科学与技术
先修课程：《高等数学》、《线性代数》、《概率论与数理统计》、《离散数学》、《数据科学导论》、《应用统计学》		
后续课程：《计算机视觉技术》、《深度学习技术》		
制定人：	审核人：	批准人：

二、实验教学任务

通过本课程的学习，使学生具有严谨、求真、求证的科学精神，具有较强的沟通交流能力及团队协作能力。了解常见的机器学习主流算法模型和未来技术的发展方向，掌握计算机学科的基本思维方法和研究方法，具有良好的科学素养和强烈的工程意识，并具备综合运用所掌握的知识、方法和技术解决复杂的实际问题及对结果进行分析的能力。培养以数据为驱动，算法为中心的问题求解能力，能够编写计算机软件程序解决行业实际问题。

三、实验项目内容与学时分配

序号	实验名称	实验类型	实验要求	学时
1	初识机器学习——运行环境设置及数据准备	演示性	必修	2
2	线性回归	设计性	必修	2
3	逻辑回归	设计性	必修	2
4	贝叶斯分类器	设计性	必修	2

5	模型评估与选择	设计性	必修	1
6	K-均值算法	设计性	必修	1
7	决策树	设计性	必修	2
8	神经网络与主成分分析	设计性	必修	2
9	支持向量机	设计性	必修	2
学时合计:16				

说明:

1. 实验设置要注意内容更新，体系设计科学合理，实验项目名称要准确规范。
2. 实验类型为：演示性、验证性、综合性、设计性。
3. 实验要求为：必修、选修。

实验一 初识机器学习——运行环境设置及数据准备

一、实验目的和要求

- (一) 掌握机器学习运行环境的搭建方法
- (二) 掌握对数据进行初步处理的方法

二、实验内容

- (1) Windows 操作系统下机器学习开发环境的安装与配置介绍 (MATLAB 或 Python)
- (2) 熟悉机器学习常用类库的基本操作、机器学习常用数据格式的创建与处理
- (3) 鸢尾花数据集的获取、加载、查看及可视化
- (4) 常用数据分析函数的使用与操作

三、实验仪器、设备

电脑一台，安装有 windows 7 旗舰版 (64 位) 及以上操作系统

四、实验原理

开发机器学习应用程序的一般步骤:

- (1) 收集数据。网络爬虫、设备实测数据、公开数据源等。
- (2) 准备输入数据。数据格式要符合要求。
- (3) 分析输入数据。是否有异常值、离群值等，数据可视化。
- (4) 训练算法。将数据输入到算法，从中抽取知识或信息。
- (5) 测试算法。评估算法的工作效果。

(6) 使用算法。将算法转换为应用程序，执行实际任务。

五、实验内容及实验步骤

(一) 机器学习常用开发环境的安装、配置 (Matlab 或 Python)

(二) 学习到 UCI、Kaggle、天池等网站下载鸢尾花数据集、波士顿房价数据集、泰坦尼克号数据集、MNIST 数据集等的方法

(三) 对给定的鸢尾花数据集进行加载、查看、图表化显示、预处理，学习掌握以上操作的命令、方法、步骤

(四) 尝试机器学习中的第一个“HelloWorld”程序，对鸢尾花数据集进行分类。

六、实验注意事项

(一) 掌握机器学习常用开发环境的安装、配置方法

(二) 初步掌握对数据集的加载、查看、图表化显示的命令操作

(三) 实验过程中认真细致，服从老师安排，遵守实验室纪律

七、实验报告要求

实验报告以书面形式提交。

实验报告主要内容：根据实验获得的体验及对实验问题的思考，以个人的观点总结归纳学到的软件环境安装配置方法、数据集分析、处理方法。

八、实验成绩考核

(百分制)