

机器学习技术与实践

每个“机器学习技术与实践”开放性实验的编写都包含了五部分内容，即实验目的、实验学时、实验内容、实验环境、实验步骤。其中实验目的介绍了本实验的教学目标，学生该达到的能力目标；实验学时则根据情况列出课程实验的课时；实验环境说明了实验实施所需要的软硬件平台；实验内容主要是对相关的 Jupyter notebook 文件内容的概述；实验步骤则简要描述课程知识点与 Jupyter notebook 文件操作的每一步输入及输出情况。在该课程中也给出了四个基于汇萃人工智能教学实验平台的验证性实验。

实验环境：汇萃人工智能教学实验平台，基于 notebook 交互式环境、sk-learn 及其数据集

实验一 线性回归算法基础

线性回归算法一般作为机器学习课程的入门算法，是学生学习机器学习课程比较早接触的算法之一。所谓的回归简单讲就是从结果来反推原因的过程，是通过对大量事实所呈现的结果，来推断系统可能存在的规律；本实验将对线性回归算法进行相关的学习。

实验内容：根据教学平台的线性回归算法实验，进行线性回归算法的基本原理的了解，并通过 python 实现对参数的计算，最终展示回归的结果。

实验二 线性回归项目实战

前面了解了线性回归算法的基础，并通过简单的 python 代码实现了最小二乘法的系数计算，本实验将继续对线性回归算法进行实战，并学习通过 sk-learn 来实现线性回归算法。

实验内容：根据教学实验平台进行线性回归算法的实战，学习 python 手工实现多元线性回归的算法，并通过 sk-learn 库实现对乐高套件的回归分析。

实验三 k 近邻算法

K-近邻算法也是机器学习算法中比较基础的算法，算法的思路也比较简单，就是根据样本本周边离它最近的 k 个邻居的类别来投票决定样本的类别，本实验将对 k-近邻算法做原理性的介绍，并通过 python 及 sk-learn 实现相应的算法。

实验内容：根据教学实验平台的 k-近邻算法实验，进行原理的学习与 python 编程，并在实验的最后部分通过“手写数字分类测试”的案例学习 sk-learn 中 k-近邻算法的使用。

实验四 梯度下降法

梯度下降在机器学习及深度学习中是非常重要的内容，它本身并不是机器学习的一个算法，但是它在求解目标函数的过程中起到了及其重要的作用；本实验将对梯度下降的基本原理进行回归，并通过 python 来编写简单的梯度下降的算法模拟。

实验内容：根据教学实验平台的梯度下降实验，学习梯度下降的基本思路，并通过 python 实现梯度下降的基本算法模拟。

实验五 逻辑回归基础

逻辑回归主要是用来做二分类预测，对于多分类的问题，则需要做一定的变化；本实验将通过对逻辑回归算法的回顾加深对其的理解，并通过 `python` 实现模型的构建。

实验内容：根据教学实验平台的逻辑回归算法实验，进行逻辑回归基本原理的回归，了解其基本的算法和损失函数；通过 `python` 实现简单的模型构建。

实验六 逻辑回归项目实战

在逻辑回归基础实验里对逻辑回归做了简单的介绍，在逻辑回归中比较难的是相关的原理的推导，公式比较多；本实验将继续对逻辑回归做介绍，并通过 `python` 代码和 `sk-learn` 库实现实验数据的分类。

实验内容：根据教学实验平台的逻辑回归实战，进行逻辑回归的算法回顾，并通过 `python` 和 `sk-learn` 实现逻辑回归的分类实践。

实验七 决策树原理

决策树(Decision Tree)是机器学习算法中一类很重要的算法，也是机器学习的十大算法之一，在实际使用中决策树可以用于分类，同时决策树也可以用于回归分析，本实验将对决策树进行相关的学习。

实验内容：根据教学实验平台的决策树实验，进行决策树的学习与 `sk-learn` 中的算法使用。

实验八 决策树项目实战

在决策树原理实验中对决策树做了原理性的介绍，本实验将对决策树做进一步的介绍，了解特征选择及决策树的生成和裁剪工作，并通过 `python` 及 `sk-learn` 实现相应的算法展示。

实验内容：根据教学实验平台的决策树实验，进行决策树的学习与 `python` 编程，内容主要是对决策树原理的回顾，通过 `python` 手工实现决策树算法(基于信息熵)，并在实验的最后部分通过“眼科医生判断患者是否需要佩戴隐形眼镜”的案例学习 `sk-learn` 的决策树编程方法。

实验九 朴素贝叶斯

朴素贝叶斯也是比较经典的机器学习算法，同时也是为数不多的基于概率论的分类算法，在本实验中将学习贝叶斯的决策理论，并通过 `python` 及 `sk-learn` 实现基本的文本分类。

实验内容：了解贝叶斯决策理论及一些常用的概率公司，并掌握朴素贝叶斯的推断方法；通过 `python` 及 `sk-learn` 学会文本分类。

实验十 主成分分析技

主成分分析技术主要是用来降维，特别是在数据量大、维度高的情况下，如果不采用一些降维的手段，那么系统的计算量将及其庞大，通过 `PCA` 等技术可以降低数据的维度，减

少系统的计算量，提升系统的效率。本实验将对 PCA 进行学习。

实验内容：学习 PCA 的基本原理，通过 python 代码模拟 PCA 的思想。

实验十一 支持向量机原理

SVM 又叫支持向量机，其实就是有向量支持的一种机器学习算法，在本实验中将对 SVM 做原理性的了解，并通过 sk-learn 中的模型库来实现相关的代码测试。

实验内容：通过学习 SVM 的基本原理，了解其背后的数学理论，并通过 sk-learn 中的库函数实现测试数据的分类

实验十二 集成算法介绍

所谓的集成学习是指通过集合多个学习器的方法，通过一定的策略用多个学习器来完成任务，从而达到比单个学习器更优的效果，集成算法一般有：Bagging、Boosting、Stacking 三大类；本实验将对集成学习做总体介绍，后续的实验将分别介绍三大类的集成算法。

实验内容：了解集成算法的基本思想及其特点，了解三大类的基本区别，并通过 sk-learn 中的库来实现集成算法的演示。

实验十三 Bagging 算法（1）

前面讲解了集成算法的基本思想，本实验将对集成学习中的 Bagging 算法做介绍，并通过 sk-learn 库函数来进行代码操作。Bagging 主要是借助于有放回的采样，用来训练多个不同的学习器，最后对多个不同的学习器的结果进行投票得出最终的分类结果。

实验内容：了解集成算法中 Bagging 算法的基本思想，并学会操作 sk-learn 中的函数库。

实验十四 Boosting 算法（2）

Bagging 算法实验（1）主要讲解了集成算法的基本思想，本实验将对集成学习中的 Boosting 算法做介绍，并通过 sk-learn 库函数来进行代码操作。Boosting 主要是通过增强不同模型间的关系，每次都做一些提升，最终达到较好的学习效果。

实验内容：了解集成算法中 Boosting 算法的基本思想，并学会操作 sk-learn 中的函数库。

实验十五 Stacking 算法

本实验将对集成学习中的 Stacking 算法做介绍，并通过 sk-learn 库函数来进行代码操作。Stacking 主要是通过组合模型的方式来实现集成算法，最终达到较好的学习效果。

实验内容：了解集成算法中 Stacking 算法的基本思想，并学会操作 sk-learn 中的函数库。

基于汇萃人工智能实验平台的验证性实验

实验一 轮廓位置与计数

实验内容:找到给定图像中与已注册的轮廓相似的物体,并输出所有目标物体的坐标和角度。即使目标物体存在欠缺、重叠以及表面的变化,仍然能够被检测。

实验二 印刷检测

实验内容:找到给定图像中与已注册的轮廓相似的物体。即使目标物体存在欠缺、重叠以及表面的变化,仍然能够被检测。在目标物中检测出比背景亮或暗的块状物,输出其缺陷个数和缺陷面积。

实验三 字符识别

实验内容:通过对要进行字符识别的物品进行“图像采集、图像预处理、行分割、列分割”,切割出检测范围内的字符信息,然后通过字符训练注册后对照识别出图像内的字符串。

实验四 药瓶良品抓取

实验内容:通过对药瓶进行“图像采集、图像预处理、良品药瓶注册学习训练”后识别出良品药瓶,然后进行“机械手标定、添加机械手抓取工具“实现良品药瓶抓取。

尊敬的老师,如果您对我们的课程实验感兴趣,需要完整的课程实验资料,请填写“实验资源申请表”,并发邮件联系我们,我们将会尽快将您所选资源发送给您。谢谢。

联 系 人: 张延迪

E-mail: zhangyd@hc-vision.cn

联系电话: 17681806123

杭州汇萃智能科技有限公司教育事业部

附件：

课程资源申请表

尊敬的老师，感谢您对我们公司的支持！您可根据教学需要选择我司的课程资源，并认真填写下表，我们会尽快将您所选资源发送给您。

姓名	
学校	
院系	
职称/职务	
教龄	
E-mail	
手机号	
所授课程	

您对我司课程资源有什么建议和意见？
