机 器 学 习 导 论

实

验

指

导

书

实验一 鸢尾花分类实验

1. 实验名称：

鸢尾花分类实验

1. 实验目的：

通过对鸢尾花数据集的分类，掌握机器学习导论课程的基础知识理论和实践应用能力，熟悉常用分类算法。

1. 实验要求：
2. 掌握数据集的获取、预处理及训练集、测试集划分方法。
3. 使用kNN、SVM或者决策树方法对数据集进行分类。
4. 掌握所使用分类方法的原理和过程。
5. 分类精度达到90%以上。
6. 实验数据：

Iris 鸢尾花数据集是一个经典数据集，在统计学习和机器学习领域都经常被用作示例。数据集内包含 3 类共 150 条记录，每类各 50 个数据，每条记录都有 4 项特征：花萼长度、花萼宽度、花瓣长度、花瓣宽度，可以通过这4个特征预测鸢尾花卉属于（iris-setosa, iris-versicolour, iris-virginica）中的哪一品种。下载链接：https://pan.baidu.com/s/1ReA5RjAUvph0BYyYIHBlYg提取码：2grr（或使用sklearn获取）。

实验二 人工神经网络实验

1. 实验名称：

人工神经网络实验

1. 实验目的：

通过对人工神经网络的实践应用，掌握人工神经网络的基础知识理论，熟悉人工神经网络的构建和训练过程。

1. 实验要求：
2. 掌握一种人工神经网络训练环境的搭建（如tensorflow、keras、matlab、pytorch等）。
3. 掌握人工神经网络的损失函数定义和训练优化过程。
4. 尝试使用调参的方法对网络训练（例如改变网络层数、连接方式、优化器等）。
5. 分类精度实现95%以上。
6. 实验数据：

MINST数据库是由Yann提供的手写数字数据库文件，数据库的里的图像都是28\*28大小的灰度图像，每个像素的是一个八位字节（0~255）。主要包含了60000张的训练图像和10000张的测试图像。数据集可以通过网址<http://yann.lecun.com/exdb/mnist/>进行下载。

大数据分析课程实验报告

1. 程序说明、软硬件环境。
2. 软件环境
3. 编程语言版本：
4. 编译器：
5. 用到的数据库：
6. 其他：
7. 硬件环境
8. CPU：
9. GPU：
10. 内存：
11. 其他：
12. 数据预处理。
13. 数据预处理方法介绍
14. 数据预处理代码
15. 数据预处理代码解释
16. 采用的分类算法介绍、分类算法代码、代码解释（逐行或逐块）。
17. 分类算法介绍
18. 分类算法代码
19. 分类算法代码解释
20. 实验结果。
21. 结果截图：
22. 实验结果分析：