

机器学习理论与实践

机器学习理论与实践

前言

课程内容

课程大作业

概述

例子

评分标准

提交方式

开发框架介绍

概述

Python

Scikit-learn 库

VUE 框架

WebSocket

开发工具简介与安装

概述

Visual Studio Code

后记

前言

同学你好，欢迎参加 [机器学习理论与实践](#) 夏季学期课程。

本课程为期 14 天，旨在让你熟练掌握机器学习的经典算法的原理，并通过 Python 和 VUE 等前后端框架实现一个网页可视化界面，方便展示和比较各种算法在经典数据集上的性能，最终产出一份实验报告。

课程内容

1. 机器学习基本概念与理论
2. [机器学习十大算法原理与实现](#)
3. [数据集介绍](#)
4. [算法性能评估与比较](#)
5. [Python 编程基础](#)
6. [Python 多态、注册框架与计算图思想](#)
7. [Scikit-learn 算法包简介](#)
8. [VUE 框架和前端技术概述](#)
9. 网页可视化界面设计与实现
10. 前后端通信与项目实践
11. 团队合作与 git 版本管理工具
12. 大作业汇报展示

课程大作业

概述

课程推荐 3-5 人组成一个学习小组，以组为单位上交作业。

课程大作业由两部分组成：

- 1. 算法原理介绍 + 实验报告
- 2. 算法实现与前后端程序 + 架构文档

大作业允许使用现有的机器学习和前后端包，如 Scikit-learn 、dynamic-links 等，但是会一定程度影响分数，具体详见评分标准。

例子

数据集

☐ iris

☒ watermelon

数据分割比例

☐ 10%

☒ 30%

数据分割器

☒ k-fold

☐ random

模型学习率

☐ low

☒ high

模型

☐ cnn

☒ resnet

评价指标

☒ accuracy

☐ f1

运行

运行结果

[Kfold_splitter]: data type: watermelon

[Kfold_splitter]: ratio: 0.3

[Resnet_model]: training..., learning rate= 0.1

[Resnet_model]: testing...

[Accuracy_judger]: the accuracy is 0.957

评分标准

最终分数参考以下指标

- 模型原理理解的准确性
- 实验报告中包含的模型数量
- 自己实现的模型比例
- 模型性能（可通过调参、tricks等方式提升性能）
- 前后端完成度与原创性
- 前后端架构设计的合理性与可扩展性
- 作业提交格式

推荐“卷”的项目

- 实验报告逻辑性、简洁性
- 经典模型实现数量
- 模型、算法改进
- 高可扩展的架构设计
- 数据可视化（重点是更好地展示信息）

不推荐“卷”的项目

- 实验报告长度
- 代码长度、时空复杂度
- 数据集个数、个性加载
- 页面美观度
- ...

提交方式

请使用邮件，将作业以 [机器学习理论与实践 20xx] 第xx组 为标题提交至 picpic2013@163.com。提交截止日期另行通知。

邮件中仅允许包含 一个 pdf 文件作为附件，其中包含：

- 目录
- 组号 + 组名
- 组员（姓名 + 学号）
- 详细分工：如完成某某模块、架构设计等
- 机器学习总体流程介绍（如数据集载入、分割、模型训练、测试、评估等）
- 算法原理介绍
- 算法改进（选做）
- 算法性能比较与分析
- 架构设计思路
- 前后端代码仓库链接
- 代码运行方式说明
- 代码运行截图
- 每位组员的收获与感悟
- 其它认为重要的内容

注意！附件仅允许包含 一个 pdf 文件，不允许在邮件中提交压缩包、代码、数据集等附件。

代码请上传至 github、gitlab 等助教可以访问到的代码托管平台，然后将链接展示在 pdf 文件中。请不要使用各种网盘等需要会员的渠道，且代码中不要包含数据集本体。

如有标题错误、附件不合规、代码上传错误等问题，将影响小组分数。

开发框架介绍

概述

本课程推荐使用 python 作为主要语言进行算法编写（推荐自行实现算法，但也可使用 Scikit-learn 库），并作为后端服务，为前端网页提供数据；推荐使用 VUE 框架编写前端，进行效果展示；推荐使用 websocket 进行前后端通信。

Python

Python 是一种广泛使用的高级编程语言，以其简洁的语法、易于阅读的代码和丰富的标准库而受到开发者喜爱。Python 适用于多种应用领域，如 Web 开发、数据分析、人工智能等。Python 拥有庞大的社区支持，因而有许多第三方库可供选择，大大丰富了 Python 的功能。

Scikit-learn 库

Scikit-learn（简称sklearn）是一个基于 Python 的开源机器学习库。它提供了一系列简单且高效的工具，用于数据挖掘、数据分析和机器学习任务。Scikit-learn 内置了许多常用的机器学习算法，如分类、回归、聚类、降维等。同时，它还提供了数据预处理、模型评估和参数调优等实用工具，使得机器学习开发更加便捷。

VUE 框架

Vue.js（简称 Vue）是一款用于构建用户界面的渐进式 JavaScript 框架。Vue 的核心库主要关注视图层，易于与其他库或现有项目集成。Vue 通过简洁的模板语法和组件化的开发方式，让前端开发更加高效和灵活。Vue.js 还提供了许多高级功能，如响应式数据绑定、计算属性、指令、过滤器等，有助于构建可维护和可扩展的前端应用。

Vue.js 文档及官网参见：<https://vuejs.org>

Vue 学习路径推荐：

1. HTML + css:
 1. 学习网址: <https://www.runoob.com/html/html-tutorial.html>
 2. B站: BV1XJ411X7Ud
 3. 重点学习语法、DOM、BOM等
2. javascript:
 1. 学习网址: <https://www.runoob.com/js/js-tutorial.html>
 2. 重点学习语法
3. nodejs & npm:
 1. B站: BV1a34y167AZ
 2. 重点理解 nodejs 与浏览器中js执行器的异同
 3. 重点学习 webpack 的 import、export 等
 4. 学习 npm 的作用 & package.json 中常用关键字的含义
4. ts & VUE3
 1. B站: BV1ra4y1H7ih

2. 学习 component、router、数据绑定等
 3. 注意结合 webpack 理解
 5. elementui
1. 官网: <https://element-plus.gitee.io>

WebSocket

WebSocket 是基于TCP协议的通信方式，是一种双向通信协议，允许在客户端和服务端之间进行全双工通信。WebSocket通过一次握手建立持久连接，之后所有的数据交换都在这个连接上进行，减少了通信开销。WebSocket支持文本和二进制数据的传输，适用于实时交互和高性能的应用场景。WebSocket的缺点是其相对较新，不是所有浏览器和服务端都完全支持。

开发工具简介与安装

概述

本课程推荐使用 Visual Studio Code 进行代码编写。也可根据个人习惯使用 pycharm、webstorm 等工具，但在此不做教学。

Visual Studio Code

Visual Studio Code（简称 VSCode）是一款免费、开源的代码编辑器，由微软开发。它支持多种编程语言和平台，具有强大的代码编辑、调试、版本控制和扩展插件功能。VSCode 提供了许多内置功能，如语法高亮、代码折叠、智能代码补全等，同时还能通过安装插件来扩展其功能。

对于本课程中的前后端开发，我们需要安装以下插件：

1. Python插件：Python 插件为 VSCode 提供了Python开发所需的功能和支持，如代码补全、代码导航、调试、单元测试、代码重构等。安装 Python 插件后，VSCode 将能够更好地支持 Python 语言的开发，提高后端代码的编写效率。
2. Vetur - Vue.js 插件：Vetur 是一个专为 Vue.js 开发者提供的 VSCode 插件。它提供了诸如语法高亮、代码段补全、错误检查、格式化等功能，以提高Vue.js项目的开发效率。安装 Vetur 后，VSCode 将能够更好地识别 Vue.js 项目的结构和语法。

安装插件的步骤如下：

1. 打开VSCode。
2. 点击左侧边栏中的扩展图标（Extensions，一个方形的图标）。
3. 在搜索框中输入插件名称（如 Vetur 或 Python），然后从搜索结果中找到对应的插件。
4. 点击插件旁边的“Install”按钮进行安装。
5. 安装完成后，按照插件的文档进行配置和使用。

通过安装这些插件，VSCode 将能够更好地支持本课程中的前后端开发，帮助学生更高效地完成项目实践。

后记

祝同学们能有所收获！