

基础课第二期大纲

01 课程介绍

课程持续时间为5个月，内容包括：

- 四大模块 循序渐进
 - python入门板块；从入门到精通
 - 数据分析与机器学习板块；深度理解和实践
 - 深度学习与人工智能板块；NLP、CV等方向解决典型问题
 - 工程能力板块；能力培养、习惯养成和工作准备
- 公选课作为补充：
 - 人工智能数学基础、人工智能数学提升、人工智能英语提升、如何阅读AI论文、简历与项目展示指导讲座。

02 课程时间规划

python入门 数据分析与机器学习 深度学习与人工智能 工程认知与能力培养

	第一周	第二周	第三周	第四周
第一个月	Python初识：基础语法 变量类型	修炼基本功：基本数据结构、循环	建造工具是工程的核心思想：初识函数	百宝箱：模块、包、标准库
第二个月	上帝式思维：面向对象	函数进阶：面向函数的编程	爬虫应用：数据采集	应用：搜索引擎
	统计学、数学	快速处理数据: NumPy pandas	直观展示数据特征：绘图可视化	数据预处理
第三个月	更智能的分析：机器学习的背景与原理	除了深度学习为什么需要机器学习的经典模型	用于人工智能的开源神器：TensorFlow教程	
	线性回归与逻辑回归	入门深度学习理解神经网络反向传播	CNN 综述 I：层	CNN 综述 II：实现细节
第四个月	具有记忆功能：循环神经网络	计算机视觉导引：Deep Tesla 无人车案例研究	推荐系统导引：基于TensorFlow的电影推荐	自然语言处理导引：新闻自动摘要系统
		工程化认知：Git与版本控制 代码风格	Linux 环境编程	大数据环境搭建：数据仓库
第五行	快速定位Bug 软件质量保证	前端开发基础	网络编程	并发编程

03 课程大纲

1. Python 入门

1.1 Python初识：基础语法，变量类型

作业：安装好编程环境；根据可视化结果，探索共享单车数据集

1.2 修炼基本功：基本数据结构、循环

作业：记账工具开发：记录一个星期的开支与收入；ATM机器实现；

新闻数据集：词汇表生成及统计

1.3, 建造工具是工程的核心思想：初识函数

作业：实现基本的爬虫流程-函数调用，获取一个月城市空气质量；手动实现print函数

1.4, 百宝箱：模块、包、标准库

作业：安装Numpy sklearn；新闻数据集：预处理与词频统计

1.5, 上帝式思维：面向对象

作业：使用面向对象完善ATM；开发小游戏：群英战吕布

1.6, 函数进阶：面向函数的编程

作业：对城市空气质量进行多条件筛选排序及数据读写

1.7, 应用：数据采集方法

实践项目：中国各城市PM2.5数据收集

关键技术点：爬虫 Request BeautifulSoup re csv

1.8, 应用：搜索引擎知识

实践项目：实现一个基本搜索引擎

关键技术点：Pandas csv lambda tf-idf pagerank

2. 数据分析与机器学习

2.1, 统计学、数学：统计学基础、假设检验、a-b test；

2.2, 快速处理数据：pandas、SQL基础、数据表结构

2.3, 直观展示数据特征：绘图与可视化

基于 点、线、区域的方法；

其他可视化图形：K线图、盒线图、地图、色块图、关系图等

项目：中国 城市PM2.5分析上（可视化报告）

2.4, 数据预处理：数据分析师在数据预处理上花费了80%的时间

作业：处理爬取的数据

2.5, 更智能的分析：机器学习的背景与原理

什么问题可以使用机器学习？如何评估与选择模型？

正则化、交叉验证；特征工程

作业：实现贝叶斯分类器

2.6, 除了深度学习，为什么依然需要机器学习的经典模型？

机器学习模型运转原理：SVM、决策树、贝叶斯、KNN、K-means

作业：新闻的话题分类（SVM、决策树、贝叶斯、预处理）

2.7, 用于人工智能的开源神器：TensorFlow教程

实践项目：中国 城市PM2.5分析下（可视化报告）

3. 深度学习与人工智能

3.1, 线性回归与逻辑回归

损失函数、Normal Equation、过拟合及欠拟合、

梯度下降与推导、优化、调参：步长、学习率、

评价方法：召回率/准确率/精确度、AP/ROC

作业：实现线性回归与逻辑回归

3.2, 入门深度学习, 理解神经网络、反向传播

作业：推导及实现反向传播 (python+numpy)

3.3, CNN 综述 I：层

CNN的基础层、功能层, 各类层的目的 和实现

3.4, CNN综述 II：实现细节

网络参数与优化方式

3.5, 具有记忆功能：循环神经网络

作业：预测未来一年航空公司的客运流量

3.6, 计算机视觉导引

实践项目：Deep Tesla 无人车案例研究, 为无人驾驶汽车寻找最佳避障、加速策略。

关键技术：强化学习、神经网络

3.7, 推荐系统导引

实践项目：基于TensorFlow的电影推荐

关键技术：Wide&Deep模型

3.8, 自然语言处理导引

实践项目：新闻自动摘要系统

关键技术点：循环神经网络 seq2seq LDA主题模型

4. 工程能力

4.1, 工程化认知：Git与版本控制、代码风格

4.2, Linux环境编程

4.3, 大数据环境搭建：数据仓库

4.4, 快速定位Bug的能力与软件质量保证

4.5, 前端开发基础

作业：面试必考——同步与异步的区别

4.6, 网络编程

作业：面试必考——socket建立连接的过程

4.7, 并发编程

作业：面试必考——进程、线程、协程概念及代码实现

04 企业真实项目实训

实践项目：中国各城市PM2.5数据收集与分析

关键技术点：Request beautifulsoup re csv tableau数据可视化

实践项目：实现一个基本搜索引擎

关键技术点：Pandas csv lambda tf-idf pagerank

实践项目：Deep Tesla 无人车案例研究, 为无人驾驶汽车寻找最佳避障、加速策略。

关键技术：Deep Tesla 强化学习、神经网络

实践项目：基于TensorFlow的电影推荐

关键技术：Wide&Deep模型

实践项目：新闻自动摘要系统

关键技术点：循环神经网络 seq2seq LDA主题模型

05 公选课

- **人工智能数学基础**
 - 课程简介：正式课程中会包含一些数学相关术语和公式，针对非理工科学生，和对数学基础有遗忘情况的学员，对大学高等数学，线性代数，概率论等基本概念进行讲解，不涉及对数学领域的深度研究。
- **人工智能数学提升**
 - 课程简介：本课将会加深你对人工智能相关数学理论的理解，包括某些数学理论的深度解析与引导，供有能力的学员进一步探索。
- **人工智能英语提升**
 - 课程简介：算法职位需要一直接触国际前沿信息。针对阅读英语资料有问题的同学。你将提升英语能力，可以自行阅读英文书籍、论文和资料。
- **如何阅读AI论文**
 - 课程简介：AI岗位需要阅读论文。本课讲解阅读论文的方法及技巧，让你能够高效的读完一篇论文，知始至终。
- **简历和项目展示指导讲座**
 - 课程简介：针对简历和展示项目薄弱的学员，提高简历和项目展示的能力，学员可边学课程边制作简历及项目整理。