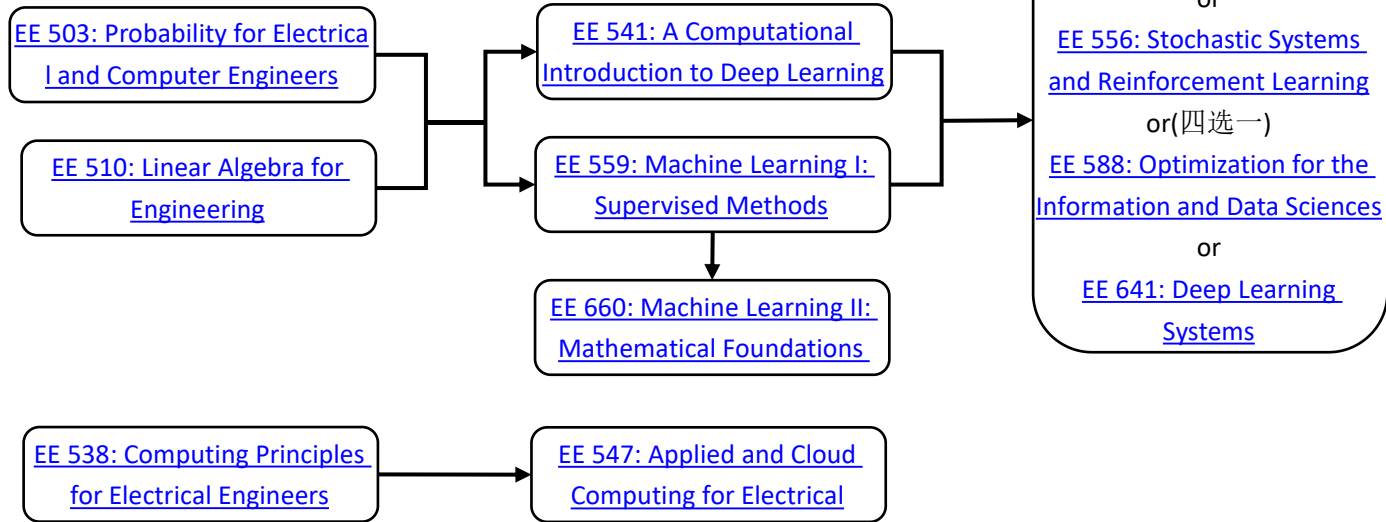


Machine Learning and Data Science

基础+必修



以下中一共再选 2~3 门：

理论：①算法 ②数据库 ③云计算 ④AI 理论 ⑤数学 ⑥信息论 ⑦信号处理

应用：①语言处理 ②多媒体处理 ③医学 ④图像处理 ⑤视觉 ⑥设计课

本文主要基于 2023Fall, 2023Summer 和 2024Spring，具体课程请浏览 <https://classes.usc.edu/>，具体要求请浏览 https://catalogue.usc.edu/preview_program.php?catoid=18&poid=25983&returnto=7300 或联系 advisor。如果本身录取的就是这个方向的话可以选择在开学前联系教授参加 [EE483](#) 和 EE538 的考试来省去学分。

ML(至少 32 学分)

1. 必选

①基础+先修要求(22):

EE 483: Introduction to Digital Signal Processing (4.0 units): 数字信号处理基础包括:离散时间线性系统, 量化, 采样, z 变换, 傅里叶变换, fft 和滤波器设计。

EE 503: Probability for Electrical and Computer Engineers (4.0 units) 概率论

EE 510: Linear Algebra for Engineering (4.0 units) 线代

EE 538: Computing Principles for Electrical Engineers (2.0 units): 软件设计的计算原理和实践概览:面向对象和非面向对象、测试、模板库、时空复杂性、数据结构、算法和动态规划。

EE 541: A Computational Introduction to Deep Learning (2.0 units): 用于训练多层感知器、卷积和循环神经网络的深度学习框架;Python, 虚拟环境, linux 和 shell 脚本, 云计算, 包括 gpu 和数据集。

EE 547: Applied and Cloud Computing for Electrical Engineers (2.0 units): 云应用介绍;云平台内的软件开发和部署;交互式用户界面;API 开发;数据库、虚拟化、异步执行、测试和可伸缩性。

EE 559: Machine Learning I: Supervised Methods (4.0 units): 无分布和概率的监督分类和回归方法学习算法;优化技术;特征空间转换;参数和非参数方法;贝叶斯决策理论;人工神经网络。

EE 660: Machine Learning II: Mathematical Foundations and Methods (4.0 units): 有监督、半监督和无监督机器学习;领域适应与迁移学习;人类的可解释性。学习的可行性、模型的复杂性和对不可见数据的性能(误差)。

②四选一(2~4):

EE 546: Mathematics of High-Dimensional Data (4.0 units): 数据表示/分析、非凸优化、高维概率、素描、聚类、低秩模型、逆问题、成像应用、机器学习、计算机视觉和神经科学等方面的现代发展。

EE 556: Stochastic Systems and Reinforcement Learning (4.0 units): 随机系统模型、动态规划、线性二次控制、卡尔曼滤波与估计、系统辨识、近似动态规划方法、自适应控制、强化和在线学习。

EE 588: Optimization for the Information and Data Sciences (4.0 units): 凸集, 函数和优化问题。基本的凸分析和凸规划理论。新颖、高效的一阶算法。在信息和数据科学中的应用。

EE 641: Deep Learning Systems (2.0 units): 神经网络用于非线性回归、分类、强化学习;多层感知器、卷积和循环网络的反向传播学习;在音频处理、视觉和自主性方面的应用。

以下内容中选 2~3 门:

2. 理论

①**算法: CSCI 570: Analysis of Algorithms (4.0 units):** 探索基本技术, 如递归, 傅里叶变换排序, 动态规划有效的算法构造。Examples include arithmetic, algebraic, graph, pattern matching, sorting, searching algorithms.

②**数据库: CSCI 585: Database Systems (4.0 units):** 数据库系统架构;概念数据库模型;语义数

数据库、面向对象数据库、基于逻辑数据库和关系数据库;用户和程序接口;数据库系统实现;完整性、安全性、并发性和恢复。推荐准备:需要关系型数据库、SQL、关系型代数和物理数据库设计知识。

③**云计算: EE 542: Internet and Cloud Computing (4.0 units):** 互联网及其协议的理论, 架构, 硬件/软件和编程, 云计算, 云计算网络, 它们如何与物联网交互。

④**AI 理论?: EE 561: Foundations of Artificial Intelligence (4.0 units):** 符号智能系统的基础, 搜索, 逻辑, 知识表示, 计划, 学习。

EE 689: Computational Intelligence and Neural Learning (4.0 units): 计算智能技术包括神经系统、深度学习、自适应模式分类、模糊函数近似、模拟退火和进化计算以及混合系统。

⑤**数学: EE 562: Random Processes in Engineering (4.0 units):** 概率公理。随机向量和过程, 它们的收敛性和极限定理。平稳过程, 马尔可夫链, 鞅, 高斯过程。随机过程的基本演算。维纳滤波。

EE 563: Inference and Estimation: Theory and Algorithms (4.0 units): 贝叶斯和非贝叶斯假设检验和估计, 充分统计, 指数族, 期望最大化, 近似, 卡尔曼滤波和平滑, 递归贝叶斯跟踪。

EE 592: Computational Methods for Inverse Problems (4.0 units): 对向量空间和功能分析概念和工具的严格描述, 这些概念和工具对解决实际应用中的反问题很有用。

⑥**信息论: EE 565: Information Theory and Its Application to (Big) Data Science (4.0 units):** 熵和互信息。可变和固定长度, 无损和有损压缩。通用压缩。文本和多媒体压缩。信道容量。纠错码。擦除和高斯通道。

⑦**信号处理: EE 596 Wavelets and Graphs for Signal Processing and Machine Learning(4.0 units):** 多速率信号处理, 小波和滤波器组。过完备和稀疏信号表示。图形信号处理。应用于压缩, 传感, 时频分析和机器学习。

3.应用

①**语言处理:**

CSCL 544: Applied Natural Language Processing (4.0 units): 介绍人类语言技术的关键组成部分, 包括:信息提取、情感分析、问答、机器翻译。

②**多媒体处理:**

EE 519: Speech Recognition and Processing for Multimedia (3.0 units): 语音产生、声学、感知、合成、压缩、识别、传输。语音、音乐和 cd 质量的编码。特征提取。回波消除。音频、视觉同步。多媒体, 互联网使用。

EE 669: Multimedia Data Compression (4.0 units): 无损压缩、矢量量化、音频/语音编码、JPEG 和 JPEG-2000、MPEG 和 H.26x 视频压缩标准、图像/视频质量评估和新兴多媒体压缩技术及应用。

③**医学:**

EE 638: Applications of Machine Learning for Medical Data (4.0 units): 机器学习模型和算法在医学应用中的应用, 从数据和疾病分类中学习。概述健康数据, 收集与传感器, 身体区域网络, 脑图像数据和其他公开可用的医疗应用数据。

EE 675: Data Analysis and Control Techniques for Neurotechnology Des (4.0 units): 数据分析, 机器学习和控制理论数学工具用于研究大脑和设计神经技术, 如脑机接口。

④图像处理:

EE 569: Introduction to Digital Image Processing (4.0 units): 图像采样、二维图像变换、图像增强、几何图像修改、形态学处理、边缘检测、纹理分析、图像滤波与恢复。

⑤视觉:

EE 676(CSCI677): Advanced Computer Vision (4.0 units): 计算机视觉的基本问题:理论、算法和应用。图像形成, 图像分割, 三维推理和测量, 运动分析, 目标和活动识别。

⑥大设计(具体课程随缘, 这是 2022 年 Fall 的 ML 相关选课):

EE 599: Special Topics (1.0-8.0 units, max 9): Systems for Machine Learning(4.0 units): 本课程着重于探索各种各样的模型和用于执行这些模型的异构硬件加速器。本课程进一步探讨了在机器学习系统中处理大量训练数据所面临的挑战。学生将通过近数据处理学习有效的数据处理技术, 包括解析、预处理和为 ML 计算管道格式化训练数据。本课程涵盖了促进近数据处理的新型内存和存储范式。最后, 本课程深入探讨了机器学习系统安全的关键方面。学生将研究对安全漏洞的日益关注以及数据隐私和完整性的重要性。该模块致力于隐私和安全方面, 为学生提供知识和工具, 以确保值得信赖的机器学习系统, 特别是在医疗和交通等关键应用领域。

内容:

- 全面了解机器学习算法的硬件系统, 包括芯片多处理器, 多线程, GPU 和并行软件。
- 具备将机器学习算法有效映射到 GPU 和大规模 CMPs 上的知识和技能, 以优化机器学习性能。
- 了解功率效率和机器学习系统中其他限制的重要性, 使他们能够设计和实施最大化能源效率的 ML 加速器。
- 深入了解各种 ML 模型的不同计算需求, 从推荐系统到基于 transformer 的自然语言处理模型, 并能够相应地选择合适的硬件加速器。
- 学习 ML 系统中高效数据处理的各種技术, 包括近数据处理方法, 能够处理和预处理大量训练数据。
- 了解与机器学习系统相关的隐私和安全挑战, 并配备策略和工具, 以确保数据隐私、完整性和安全的微架构。
- 培养对机器学习硬件系统的新兴趋势和最新发展的批判性意识, 使他们能够保持更新并适应该领域的未来发展。