



研究生课程教学大纲

课程名称：	材料基因与数据科学：工程应用		
	Materials Gene and Data Science: Engineering Application		
课程编号：	ZX14206M		
开课单位：	材料科学与工程学院	开课学期：	春季
课内学时：	32	学分：	2
适用学科专业及层次：	材料、物理或化学化工类研究生及高年级本科生		
授课语言：	双语		
先修课程：	普通物理（或物理化学）、量子力学（或结构化学）、程序设计		
负责人：	任浩	团队成员：	赵文、朱后禹

一、课程简介

(300-500字，为宋体小四号字。简要介绍的性质、主要教学内容、课程学习目标等。)

1) 课程性质

材料基因融合了材料科学、信息科学、先进实验方法等学科，采用数值计算、数据挖掘、机器学习等技术研究材料的微观结构、性能和工艺过程对性能的调控机制，引导并支撑实体材料的研发和应用。然而，目前研究生的知识和技术储备距实际科研活动中用到的材料基因技术有较大的缺口。针对这一点，本课程将以材料物理或材料化学低年级研究生的知识储备为基础，面向材料、物理、或化学专业研究生或高年级本科生开设，力图通过较短时间的课程学习，使学生领略到人工智能和机器学习技术的魅力，了解材料设计的基本原理和应用范围，为从事理论计算方向和实验研究方向的研究生补充技术储备，进行新型材料理性设计。

2) 主要教学内容

本课程主要内容为新材料研发和设计过程中需要的材料基因技术和机器学习算法。结合选课学生的背景知识，课程首先简单介绍 Python 程序设计，在此基础上，讲授常见机器学习算法和框架、爬虫与数据采集、监督和无监督学习、数据可视化等内容，课程最后通过电极材料高通量筛选实践熟悉实际操作。通过本课程学习，使学生了解机器学习技术的基本原理并能够应用于自己的课题研究和学业学习，熟悉材料设计的流程，充实同学们的科研工具库，开拓思路，在遇到相关问题时能够想到借助不同工具克服困难、达成目标。

3) 课程学习目标

本课程的主要教学目标为引导同学们了解大数据时代基于人工智能技术的材料设计模式，介绍机器学习这一蓬勃发展的强力工具及其在化学、材料、生物、能源等领域的应用。通过本课程的学习，同学们将掌握机器学习技术和材料设计的基本思路，了解常见的机器学习算法，迅速积累知识和技巧；通过上机实践，能够结合自己的课题研究进行机器学习技术的应用；搭建起材料基因组技术的知识框架，在日后的科研和学习中能够快速定位自身知识结构的问题，借助数据库和网络资源，对科研课题进行理性设计和规划。

二、课程大纲

(一) 课程目标

课程目标	培养目标、基本要求指标点
目标 1：掌握 Python 语句语法，能够阅读已有代码并按需修改。	2 (知识结构要求)：牢固掌握材料基因工程领域专业技术；
目标 2：了解机器学习的基本概念。	2 (知识结构要求)：掌握材料基因工程领域相关培养方向坚实的基础理论和系统专业知识。
目标 3：了解常见的机器学习算法和材料基因组技术流程。	3(专业能力要求)：能够运用科学原理，掌握较强的解决本领域工程实际问题的能力；具备一定的国际视野和跨文化交流能力。
目标 4：掌握 scikit-learn、TensorFlow 和 PyTorch 框架的基本操作和应用。	3(专业能力要求)：能够运用科学原理，掌握较强的解决本领域工程实际问题的能力；具备一定的国际视野和跨文化交流能力。

(二) 课程内容

第1章 Python 基础（2学时）

本章重点难点：Python语法、科学计算与可视化

1. Python简介
2. Python语法基础
3. Python科学计算和可视化应用

第2章 机器学习简介（2学时）

本章重点难点：机器学习基本概念及现代应用

1. 机器学习是什么
2. 机器学习应用、挑战与发展历程
3. 机器学习主要算法介绍
4. Scikit-learn安装与配置

第3章 机器学习项目实例（4学时）

本章重点难点：数据爬取、数据清洗、选择模型、调参

1. 实例一：手写数字识别
2. 实例二：鸢尾花分类

第4章 分类（2学时）

本章重点难点：各类模型评价指标

1. 二分类器
2. 模型评价指标
3. 多分类器
4. 错误分析

第5章 回归（4学时）

本章重点难点：矢量运算、线性回归、逻辑回归

1. 线性回归
2. 梯度下降
3. 多项式回归
4. 学习曲线
5. 正则线性回归
6. 逻辑回归

第6章 常用机器学习算法（6学时）

本章重点难点：决策树、随机森林、主成分分析、K-Means

1. 支持向量机
2. 决策树
3. 集成学习与随机森林
4. 数据降维
5. 无监督学习

第7章 神经网络与深度学习（6学时）

本章重点难点：神经网络种类及优缺点

1. 神经网络种类
2. 神经网络框架：TensorFlow 和 PyTorch
3. 训练神经网络

第8章 电极材料高通量筛选与设计实验（6学时）

应用机器学习算法和技术，爬取数据，清洗数据，选择合适机器学习算法，训练模型，筛选合适的电极材料。

三、教学安排及要求

内容	课内学时	教学方式	课外学时	课外环节	课程目标
1. 1	1. 5	理论讲授	0. 2	文献阅读	目标 1
1. 2	0. 5	理论讲授	0. 5	文献阅读	目标 1
1. 3	0. 5	理论讲授/实验实践	0. 5	线上学习	目标 1
2. 1	0. 5	理论讲授	0. 2	文献阅读	目标 2
2. 2	0. 5	理论讲授	0. 2	文献阅读	目标 2
2. 3	0. 5	理论讲授	0. 5	文献阅读	目标 2
2. 4	0. 5	理论讲授/实验实践	0. 5	线上学习	目标 2
3. 1	2. 0	理论讲授/实验实践	0. 5	线上学习	目标 3
3. 2	2. 0	理论讲授/实验实践	0. 5	线上学习	目标 3
4. 1	0. 5	理论讲授/案例研讨	0. 5	线上学习	目标 3
4. 2	0. 5	理论讲授/案例研讨	0. 5	线上学习	目标 3
4. 3	0. 5	理论讲授/案例研讨	0. 5	线上学习	目标 3
4. 4	0. 5	理论讲授/案例研讨	0. 5	线上学习	目标 3
5. 1	1. 0	理论讲授/案例研讨	0. 5	线上学习	目标 3
5. 2	0. 5	理论讲授/案例研讨	0. 5	线上学习	目标 3
5. 3	0. 5	理论讲授/案例研讨	0. 5	线上学习	目标 3
5. 4	0. 5	理论讲授/案例研讨		线上学习	目标 3
5. 5	0. 5	理论讲授/案例研讨		线上学习	目标 3
5. 6	1. 0	理论讲授/案例研讨		线上学习	目标 3
6. 1	1. 0	理论讲授/案例研讨	0. 5	线上学习	目标 3
6. 2	1. 0	理论讲授/案例研讨	0. 5	线上学习	目标 3
6. 3	1. 0	理论讲授/案例研讨	0. 5	线上学习	目标 3
6. 4	1. 0	理论讲授/案例研讨		线上学习	目标 3
6. 5	2. 0	理论讲授/案例研讨		线上学习	目标 3
7. 1	2. 0	理论讲授	0. 5	线上学习	目标 3
7. 2	2. 0	理论讲授/实验实践	0. 5	线上学习	目标 3
7. 3	2. 0	理论讲授/案例研讨	0. 5	线上学习	目标 3
8. 1	6. 0	实验	1	撰写实验报告	目标 4

四、考核内容、方式及评分标准

(一) 考核环节

考核环节		总成绩占比	支撑课程目标
平时作业	1. 共布置若干道题目，平均每周 3 道小题或 1 道大题。 2. 成绩采用百分制，根据作业完成准确性、是否按时上交、是否独立完成评分。 3. 考核学生对基本知识的掌握能力，综合运用所学知识分析问题、解决问题的能力，题型主要是数据分析、代码编写和画图。	35%	目标 1-4
课堂表现	1. 成绩采用百分制，根据课堂提问、互动等评分。	5%	目标 1-4
实验	本课程有一次 8 个学时实验。成绩采用百分制，根据实验完成情况评分。	60%	目标 1-4

(二) 评分标准

考核环节	<60	60-75	75-90	90-100
平时作业	不能按照作业要求完成作业	基本概念不清晰	基本按照作业要求并及时完成，基本概念基本清晰	严格按照作业要求并及时完成，基本概念清晰
课堂表现	旷课	不能回答老师提出的问题	基本能回答老师提出的问题	积极沟通、积极提问、能回答老师提出的问题
实验	不能按照作业要求完成实验报告	基本概念不清晰，解决问题的方案基本不正确、不合理	基本按照实验要求并及时完成，思路基本清晰，解决问题的方案基本正确、基本合理	严格按照实验要求并及时完成，思路清晰，解决问题的方案正确、合理，能提出不同的解决问题方案

(三) 考核环节

考核环节	目标 1	目标 2	目标 3	目标 4
平时作业	35%	35%	35%	35%
课堂表现	30%	30%	30%	30%
实验	35%	35%	35%	35%
总权重	100%	100%	100%	100%

五、教材与参考资料

(正文为宋体小四号字。正式出版教材要求注明教材名称、作者姓名、出版社、是否自编教材；自编教材要求注明是否成册、编写者姓名、编写者职称、字数等。

(一) 教材

1. 周志华, 机器学习, 清华大学出版社, 2016.

六、其它说明

大纲执笔人：

审核人（学位点负责人）：

分管院长签字：

教学大纲编写负责人（材料化学系）

课程性质	课程编码	课程名称	学时	课程负责人
专业基础课		能源材料工程前沿进展	32	邢伟
专业选修课	7144101	功能复合膜材料	32	张晓云
	7143104	高性能科学计算	32	燕友果
	6145001	新能源化学与材料	32	陈艳丽
	6144002	固体量子化学	48	任浩
	6144003	新能源材料创新设计与评价	32	邢伟
	6143006	材料基因与数据科学：工程应用	32	任浩
	6143007	石油工程领域的材料设计与模拟	32	燕友果