

大语言模型：基础、实践与挑战

课题大纲

01

课程背景及简介

适合人群：本课程属于通识课程，适合对（计算机、软件、人工智能专业方向）感兴趣的同学。

可持续发展目标 9：产业、创新和基础设施

近年来，伴随着深度学习技术和计算机软硬件性能突飞猛进的发展，诞生了大语言模型如 ChatGPT，GPT-4 等。大模型的出现现在各行各业掀起了一股人工智能应用的热潮。国内外涌现出各种围绕大模型的技术，形成了一片蓬勃发展的局面。中国头部科技企业（阿里、百度、腾讯、华为、字节等）、新兴创业公司（百川智能、MiniMax 等）、传统 AI 企业（科大讯飞、商汤科技等）以及高校研究院（复旦大学、中科院等）亦加速大模型领域投入，推出文心一言、通义千问、百川、盘古大模型等。对于大语言模型的原理和应用，人们表现出浓厚的好奇心，对大模型所带来的未来变革更是充满了憧憬。然而，与此同时，大模型也为我们的生产何生活方式带来了全新的挑战。如何应对这些挑战成为当前亟需关注的焦点。

本课程将带领同学们深入理解大模型的概念、应用、意义以及其在技术上的重要性。通过系统的课程基础知识讲解，学生将了解大模型背后的机器学习原理，学习使用编程工具进行大模型的微调，并了解在特定领域中应用大模型的实践方法，例如机器翻译、智能客服、程序生成等。本课程将助力学生建立对大模型的全面认识，为他们未来参与新的科技革命做好充分的准备，也为推动科技进步贡献自己的力量。希望同学们能够在学习过程中深化对大模型的理解，为未来科技的发展注入新的活力。

02

学习目标

本课程将收获：

1. 深入理解大语言模型的基本原理
2. 将大模型技术应用到实际工程问题中
3. 运用编程框架微调大模型

03

导师信息

顾教授

上海交通大学软件学院副教授，研究方向大语言模型、智能软件工程、自然语言处理，为软件代码开发高效的机器学习算法，研究课题包括代码大模型、程序自动生成、程序翻译等。研究成果发表在 ICSE、ICLR、FSE、AAAI 等国际重要期刊和会议上。谷歌学术引用 1800 余次，主持过 CCF-腾讯犀牛鸟基金项目、国家自然科学基金项目等。

04

课程设置

模块 1：数学、机器学习基础

学习目标：拆解自动水果贩卖机背后的机器学习原理

简介：自动水果贩卖机为例介绍机器学习系统的构建过程

结合人类学习的类比介绍机器学习的基本概念、基本要素（数据、模型、损失函数、优化算法）。最后，介绍课程内容及所需的数学基础知识，例如概率分布、期望、方差、贝叶斯规则等。

模块 2：神经网络

学习目标：了解神经网络基本原理和发展现状

简介：介绍神经网络原理和训练过程。从人脑的工作机制出发，了解神经元的基本功能，引出人工神经元——感知机。通过异或函数讲解多层感知机的构造及其功能。神经网络目前的发展现状和未来可能存在的重要技术突破。

模块 3：语言模型

学习目标：学习语言模型的基本原理和技术未来发展的可能性

简介：计算机如何理解单词和句子。由此引出语言模型的概念、计算。从序列理解的角度讲解循环神经网络的产生过程和基本原理。通过机器翻译为例介绍序列生成问题的基本思路和方案。语言模型这项技术当下的发展阶段和对生活的深远影响。

模块 4：大语言模型原理及应用实践

学习目标：大语言模型原理及应用实践

简介：介绍 Transformer、BERT、GPT 等的基本原理和训练方法。在此基础上介绍大语言模型(大模型)，以 ChatGPT 为例，讲述大模型的原理和典型应用场景，如机器翻译、智慧医疗、智能客服、程序生成等。最后，以自动驾驶为例，介绍常见的 AI 应用背后的技术(如视觉神经网络、强化学习等)。

模块 5：大模型技术挑战与治理

学习目标：大模型技术挑战与治理

简介：介绍大模型应用中遇到的挑战难题如计算资源开销、幻觉、评估指标、版权保护、数据泄露、内容风控、以及对传统行业的颠覆等，引导学生就特定问题提出治理方案。

05

延伸阅读

1. Jean Kaddour et al. Challenges and Applications of Large Language Models [M]. arXiv preprint arXiv:2307.10169. 2023
2. 车万翔，郭江，崔一鸣. 自然语言处理：基于预训练模型的方法. 电子工业出版社. 2021 年