

AI 求职直通班



一站掌握 AI 热门领域核心技能，高性价比对接 AI 职场需求

在 AI 求职中，你是否面临以下问题？

【核心技能缺失】

- 缺乏 AI 类职位工作的必备技能，AI 热门领域技能了解；
- 缺少实际 AI 项目经历；
- 有转行想法，但结合自身工作经验和知识结构，不确定什么方向的岗位适合；

【求职准备无从下手】

- 简历改了又改，面试杳无音讯；
- 找不到帮助提升面试成功率的有效方式；
- 苦恼于是否能够通过笔试、技术面；

AI 求职，学这一门课就够了

贴合 AI 岗位核心技能定制系统课程

课程针对想要转行到 AI 领域求职的：**程序员、数据分析师** 设计。内容涵盖 AI 岗位必备核心技能，除机器学习、深度学习外，学习者可在 CV、NLP 方向二选一深入学习。

9 个硅谷实战项目，专业代码审阅，为简历加分

学习者将通过 Learning by doing 的学习方式，挑战硅谷实战项目，获得专业代码审阅。跨越理论到实践的鸿沟，为项目经历加分。

6 大 AI 热门方向，11 个前沿案例，在线演练

除实战项目外，额外包含 11 个前沿案例供学习者在线演练，内容涉及：计算机视觉、自然语言处理、语音识别、金融风控、数据挖掘、计算广告、推荐系统，7 大热门机器学习方向。

竞赛、刷题、求职辅导，求职竞争力 Max

课程将通过：数据科学挑战赛、算法刷题营、简历审阅、行为导向面试、机器学习技术模拟面试，帮助求职者从技能练习与求职准备两个方面，有效提升求职竞争力。
学习者还有机会获得名企实习内推机会，获取更多项目经历。

注：由于技术的快速迭代，中国区教研团队将根据当前热点对毕业项目选题进行实时更新，请以教室中的选题为准。

课程概览

先修要求：掌握 Python 编程，并具备统计学、线性代数、微积分基础

课程有效期：8 个月

学习辅导：群内答疑

第一阶段 机器学习

建议学习时长：9 周

系统学习机器学习基础知识，共4个章节：

- 机器学习基础
- 监督学习
- 非监督学习
- 深度学习基础

实战项目（代码审阅）：

- 【数据挖掘】预测波士顿房价
- 【数据挖掘】为慈善机构寻找捐赠者
- 【数据挖掘】创建客户细分
- 【数据挖掘】预测共享单车使用情况

案例演练：

- 【数据挖掘】泰坦尼克号乘客生存率分析
- 【数据挖掘】预测你的下一道世界料理
- 【数据挖掘】游戏玩家付费金额预测
- 【金融风控】为信贷公司搭建金融风控模型
- 【数据挖掘】电影评分的 k-means 聚类
- 【数据挖掘】分析学生录取数据
- 【计算广告】企业广告点击率预测

第二阶段 深度学习

建议学习时长：4 周

学生将在CV与NLP中二选一，深入学习该领域的专项知识，两个方向各2个章节：

注：可以选择都学

- 计算机视觉基础
- 计算机视觉进阶
- 自然语言处理基础
- 自然语言处理进阶

实战项目（代码审阅）：

- 【计算机视觉】面部关键点检测
- 【计算机视觉】图像标记
- 【自然语言处理】词性标注

注：由于技术的快速迭代，中国区教研团队将根据当前热点对毕业项目选题进行实时更新，请以教室中的选题为准。

	<ul style="list-style-type: none"> ● 【自然语言处理】机器翻译 ● 机器学习毕业项目 <p>案例演练：</p> <ul style="list-style-type: none"> ● 【计算机视觉】医学影像检测皮肤癌 ● 【自然语言处理】情感分析 ● 【自然语言处理】构建垃圾邮件分类器 ● 【时间序列分析】预测股票价格
第三阶段 实践拔高 建议学习时长：3 周	<p>数据科学挑战赛</p> <p>时长 2 周，帮助学生参与到正在进行中的竞赛中（如天池、Kaggle等），通过导师的指导，完成自己的第一次提交，并在不停优化中获得不错的竞赛成绩，丰富简历。</p> <p>算法刷题营</p> <p>时长 1 周，帮助学生短周期内备战算法面试，在训练营中能够掌握面试高频算法，练习真题。</p>
第四阶段 求职辅导 建议学习时长：3 周	<p>求职辅导</p> <ul style="list-style-type: none"> ● 简历审阅 ● 行为导向面试 ● 机器学习技术模拟面试 <p>*【福利】名企实习内推机会</p>
热门技能包（选修） 注：无答疑服务	<p>该模块中，补充了与核心课程相关的部分知识内容，常常在JD中出现，不作为学习要求，供学员根据求职目标和技术背景等自行安排，共5个章节，包含：</p> <ul style="list-style-type: none"> ● 语音识别 ● 深度学习框架：Keras、Tensorflow、PyTorch ● C++ ● GitHub使用 ● Shell Workshop <p>案例演练：</p> <ul style="list-style-type: none"> ● 【语音识别】搭建DNN语音识别器 ● 【算法优化】高性能粒子滤波器

注：由于技术的快速迭代，中国区教研团队将根据当前热点对毕业项目选题进行实时更新，请以教室中的选题为准。

第一阶段：机器学习（第1-9周）

第0章：欢迎学习 AI 求职直通班！

课程名称	学习目标
课程简介	→ 学前准备，了解课程架构，做好准备！
什么是机器学习？	→ 初步探索机器学习概念，并尝试用机器学习解决实际问题

案例演练：泰坦尼克号乘客生存率分析

你将创建决策函数，并根据1912年泰坦尼克号海难的乘客特征，如：性别、年龄等，对乘客生还结果进行预测。你将从一个简单的算法入手，然后逐渐增加该算法的复杂度，直至你至少能精确地预测出所提供数据中 80% 的乘客的生还结果。通过该项目，你可以在正式开始学习本纳米学位前，了解机器学习的一些概念。

应用领域：数据分析与预测

第1章：机器学习基础

课程名称	学习目标
训练与测试模型	→ 使用 Pandas 读取数据集，并使用 scikit-learn 训练与测试模型。
评估指标	→ 了解用于评估模型性能的指标，如准确率、精度、召回率和 ROC 得分。
模型选择	→ 学习如何进行交叉验证，通过学习曲线判断过/欠拟合，并学习如何使用网格搜索来训练模型。
自我评估：NumPy 与 pandas	→ 测试你的 NumPy 和 pandas 技能

注：由于技术的快速迭代，中国区教研团队将根据当前热点对毕业项目选题进行实时更新，请以教室中的选题为准。

自我评估：模型评估与验证

→ 测试你对模型评估与验证的理解

案例演练：预测你的下一道世界料理

你将利用一个涵盖了中国菜、越南菜、法国菜等信息的菜谱数据集来训练一个监督学习模型，并对模型的性能和预测能力进行测试。通过该数据训练后的好的模型可以根据一道菜的原料来预测这道菜所属的菜系。

应用领域：建模预测/监督学习/数据挖掘

实战项目：预测波士顿房价

波士顿地区的房地产市场竞争激烈，而你目标成为该地区最好的房地产经纪人。为了使自己更具有竞争力，你决定使用一些机器学习概念来帮助自己和客户制定适合该地区房屋的最佳销售价格。幸运的是，你找到了波士顿房价数据集，该数据集包含有关大波士顿社区房屋各类属性的汇总数据，包括每个区域的房屋价值中位数。你的任务是基于可用工具进行统计分析，并构建最佳模型。这个模型将被用来预测你的客户想要购买的房屋的最合适的销售价格。

第2章：监督学习

课程名称

学习目标

线性回归

→ 使用 Pandas 读取数据集，并使用 scikit-learn 训练与测试模型。

感知器算法

→ 了解用于评估模型性能的指标，如准确率、精度、召回率和 ROC 得分。

决策树

→ 学习如何进行交叉验证，通过学习曲线判断过/欠拟合，并学习如何使用网格搜索来训练模型。

朴素贝叶斯

→ 测试你的 NumPy 和 pandas 技能

支持向量机

→ 测试你对模型评估与验证的理解

集成方法

→ 通过 boosting 提升传统方法；Adaboost

自我评估：监督学习

→ 监督学习相关的测试题

注：由于技术的快速迭代，中国区教研团队将根据当前热点对毕业项目选题进行实时更新，请以教室中的选题为准。

案例演练：追踪部落联盟中的高氪玩家——游戏玩家付费金额预测

某游戏开发商希望能够准确预测每一位玩家的价值，有助于带来更好的游戏广告投放策略和高效的运营活动（如精准的促销活动和礼包推荐）。你将通过玩家在游戏内前7日的行为数据，预测他们每个人在45日内的付费总金额。

应用领域：建模预测/监督学习/数据挖掘

案例演练：企业广告点击率预测

广告点击率预估是很多广告算法工程师喜爱的战场，可以帮助广告主和广告平台更好地做决策。你将搭建机器学习模型帮助企业预测广告是否会被点击，辅助公司决策，为公司带来增量收入。

应用领域：计算广告/建模预测

案例演练：为信贷公司搭建风控模型

信贷平台中，如何判断申请人的信用非常重要。你将通过机器学习算法对每一位申请人的基本信息、交易信息等数据进行甄别，减少信贷公司的坏账损失。

应用领域：金融风控

实战项目：寻找慈善机构的捐助者

CharityML 需要你帮助他们建立一种模型，来识别潜在的最可能捐助的人并降低发送邮件的费用。你的目标是评估和优化几个不同的监督学习模型，以确定哪种算法能够获得最高的捐赠收益率，同时减少发送信件中的字数。

第 3 章：非监督学习

课程名称	学习目标
聚类	→ 学习如何聚类算法，并尝试使用 k-means 对数据进行聚类
聚类迷你项目	→ 使用 k-means 对电影评分进行聚类
层次聚类法与密度聚类	→ 学习单连接聚类法和层次聚类法，DBSCAN
高斯混合模型与聚类验证	→ 学习高斯混合模型及相关示例
特征缩放	→ 通过案例学习特征缩放
PCA（主成分分析）	→ 了解降维的作用，并学习 PCA 的原理和使用场景
PCA 迷你项目	→ 使用特征脸方法和 SVM 进行脸部识别
随机投影与 ICA	→ 学习随机投影与独立成分分析，并通过 Lab 学习如何应用这些方法
非监督学习自我评估	→ 非监督学习相关的测试题

案例演练: 电影评分的 *k-means* 聚类

假设你是 Netflix 的一名数据分析师，你想要根据用户对不同电影的评分研究用户在电影品位上的相似和不同之处，了解这些评分对用户电影推荐系统是否有帮助。你将探索研究用户电影评分数据集，并学习推荐系统的经典算法——协同过滤。

应用领域：聚类/推荐系统/协同过滤

案例演练: 使用特征脸方法和 SVM 进行脸部识别

你将学习如何将主成分分析运用到面部识别。首先，你将脸部数据集（当做无标签数据集）计算 PCA（特征脸），即无监督式特征提取/降维。然后，你将训练SVM分类模型，并通过测试集评估模型质量。

应用领域：SVM/PCA/特征脸

注：由于技术的快速迭代，中国区教研团队将根据当前热点对毕业项目选题进行实时更新，请以教室中的选题为准。

实战项目: 创建客户细分

你将针对葡萄牙里斯本批发商的客户收集的支出数据应用非监督学习技术，以确定隐藏在数据中的客户群。你需要首先选择一个小样本子集来探索数据，并确定是否有任何产品类别彼此高度相关。之后，你将通过缩放每个产品类别来预处理数据，然后移除不需要的异常值。借助良好、干净的客户支出数据，您将对数据应用PCA转换，并使用聚类算法以对转换后的客户数据进行细分。最后，您将比较发现的细分和附加标签，并考虑这些信息是否可以帮助批发商进行未来服务变更。

第二阶段：深度学习（第10-15周）

第 4 章：深度学习基础

课程名称	学习目标
神经网络简介	→ 学习深度学习与神经网络的基础知识。你也将在教室里亲手用 Python 实现梯度下降法与反向传播。
实现梯度下降	→ 学习另一个误差函数，并带领你使用 NumPy 矩阵乘法实现梯度下降。
训练神经网络	→ 你将学习如何训练神经网络，以提升其训练效果。

案例演练：分析学生录取数据

你将把神经网络用于加州大学洛杉矶分校的学生录取数据，并预测学生的录取情况。你也将执行神经网络训练的一些步骤，包括One-hot 编码数据、缩放数据、编写反向传播步骤等。

应用领域：神经网络/建模预测

实战项目：预测共享单车使用情况

你将从零开始搭建并训练一个神经网络，并用该网络预测每日自行车租客人数，为某一共享单车预测某一天内需要的使用量，帮助他们作出管理自行车的决策。

第 5 章：计算机视觉基础

课程名称	学习目标
欢迎学习计算机视觉课程！	了解课程教学内容，并了解CV的行业应用！
图像表示法和分类	了解如何以数字方式表示图像并实施图像处理技巧，例如颜色遮罩和二元分类。
卷积过滤器和边缘检测	学习图像中的频率并实现你自己的图像过滤器，用于检测图像中的边缘和形状。使用计算机视觉进行脸部检测。
特征类型和图像分割	开发角点检测器，并学习 k 均值聚类等技巧以将图像分割为独特的部分。
特征向量	学习如何使用特征向量描述对象和图像。
CNN 层级和特征可视化	定义并训练你自己的服饰识别卷积神经网络。使用特征可视化技巧了解网络学习的规律。

案例演练：皮肤科医生的人工智能——通过深度学习进行皮肤癌筛选检测

该案例是论文实现，通过医学影像的处理来进一步学习模型评估的ROC指标，并深入理解敏感性与特异性的分类标准。跟随 Google 无人车之父 Sebastian 与讲师一起完成整个医学分类模型的全流程，搭建从图像数据训练、到验证并可视化，并与真实医生检测结果对比，进而辅助皮肤科医生对皮肤癌进行筛选和检测。

实战项目：脸部关键点检测

你会将计算机视觉技术和深度学习架构知识相结合，构建一个脸部关键点检测系统。脸部关键点包括眼睛、鼻子和嘴部周围的点。脸部关键点用在了很多应用中，包括：脸部跟踪、脸部姿势识别、脸部过滤器和情感识别。你完成的代码应该能够查看任何图像后检测脸部，并预测每个脸部的脸部关键点位置。

第 6 章：计算机视觉进阶

课程名称	学习目标
高级 CNN 架构	了解高级 CNN 架构并了解基于区域的 CNN（例如 Faster R-CNN）如何快速定位并识别图像中的对象。
YOLO	了解 YOLO (You Only Look Once) 多对象检测模型，并实现 YOLO 算法。
RNN	学习如何使用递归神经网络 (RNN) 使深度学习网络具有记忆。了解 RNN 如何从有序数据中学习规律并生成有序数据。
LSTM	学习长短期记忆网络 (LSTM) 以及有助于保留长期记忆的类似架构。
超参数	了解用于定义和训练深度学习模型的不同超参数，课程中将讨论调整每个超参数的起始值。
深度学习注意力机制 Attention	注意力是近期深度学习的最重要创新技术之一，你将学习注意力模型的原理并完成基本的代码实现过程。

实战项目：图像说明

在此项目中，你将创建一个神经网络架构来自动生成图像说明。使用 Microsoft 上下文常见对象 (MS COCO) 数据集训练网络后，你将用新的图像测试网络！

第 7 章：自然语言处理基础

课程名称	学习目标
自然语言处理简介	→ 学习如何构建 NLP 模型管道,
文本处理	→ 学习通过以下方式整理来自不同来源的文本以进一步处理这些文本：清理、标准化并拆分为单个单词或标记
利用朴素贝叶斯的垃圾邮件分类	→ 你将学习如何使用朴素贝叶斯算法构建垃圾邮件分类器。
利用 HMM 进行词性标注	→ 你将学习多个词性标注方法，并深入讲解隐马尔可夫模型。
书虫	→ 学习如何使用 IBM Watson 构建简单的问答代理。

案例演练：构建垃圾邮件分类

在此案例中，你将学习如何使用朴素贝叶斯算法构建垃圾邮件分类器。垃圾邮件检测是机器学习在现今互联网领域的主要应用之一。几乎所有大型电子邮箱服务提供商都内置了垃圾邮件检测系统，能够自动将此类邮件分类为“垃圾邮件”。

实战项目：词性标注

你将使用 Pomegranate 库构建隐马尔可夫模型，并使用通用标签集进行词性标注。在使用更大型的标签集对实际文本语料库进行标注时，隐马尔可夫模型的准确率达到 96% 以上。隐马尔可夫模型还用于语音识别和语音生成、机器翻译、生物信息学基因识别和计算机视觉人类手势识别等等。

第 8 章：自然语言处理进阶

课程名称	学习目标
特征提取和嵌入	→ 利用词袋、TF-IDF、Word2Vec 和 GloVe 等方法转换文本，以提取可以用于机器学习模型的特征。

注：由于技术的快速迭代，中国区教研团队将根据当前热点对毕业项目选题进行实时更新，请以教室中的选题为准。

主题建模	→ 学习如何使用潜在狄利克雷分析 (LDA) 将文档集合划分为主题。在 Lab 中，你会将此模型应用到一个新闻数据库上。
情感分析	→ 学习如何使用几种机器学习分类器（包括RNN）预测文字情感，并将所学知识运用到影评数据库中。
Seq2seq	→ 学习一种特定的 RNN 架构，用于从一种序列生成另一种序列。
注意力机制 Attention	→ 注意力是近期最重要的深度学习创新技术之一。在此部分，你将学习注意力，并在 Lab 中了解注意力的基本实现过程。

案例演练：情感分析

你将学习如何使用几种机器学习分类器（包括递归神经网络）预测文字情感，并将所学知识运用到影评数据库中。

实战项目：机器翻译

机器翻译是一个热门研究话题，每年都有大量相关论文发表出来。在这些年的研究过程中，出现了各种不同的方法，例如基于规则的机器翻译、统计学机器翻译和基于样本的机器翻译。但是，这依然是一个未解的问题。不过，神经网络使机器翻译出现了重大发展。在此项目中，你将构建一个用在端到端机器翻译管道中的深度神经网络。你完成的管道将接受英文输入，并返回法语翻译。

实战项目：毕业项目（多选一）

在毕业项目中，你将运用你在此纳米学位中学到的机器学习算法和方法，选择一个你感兴趣的问题来解决。

方向一：自然语言处理方向——句子相似度匹配

[Quora Question Pairs数据集](#)是Quora于2017年公开的句子匹配数据集，其通过给定两个句子的一致性标签标注，从而来判断句子是否一致。

方向二：数据挖掘方向——预测 Rossmann 未来的销售额

Rossmann 是欧洲的一家连锁药店。在这个源自Kaggle比赛 [Rossmann Store Sales](#)，你需要根据 Rossmann 药妆店的信息（比如促销，竞争对手，节假日）以及过去的销售情况，来预测它未来的销售额。

方向三：计算机视觉方向——猫狗大战

使用深度学习方法识别一张图片是猫还是狗。

方向四：计算机视觉方向——盐块识别挑战

通过图像识别预测地表下是否为盐块。