



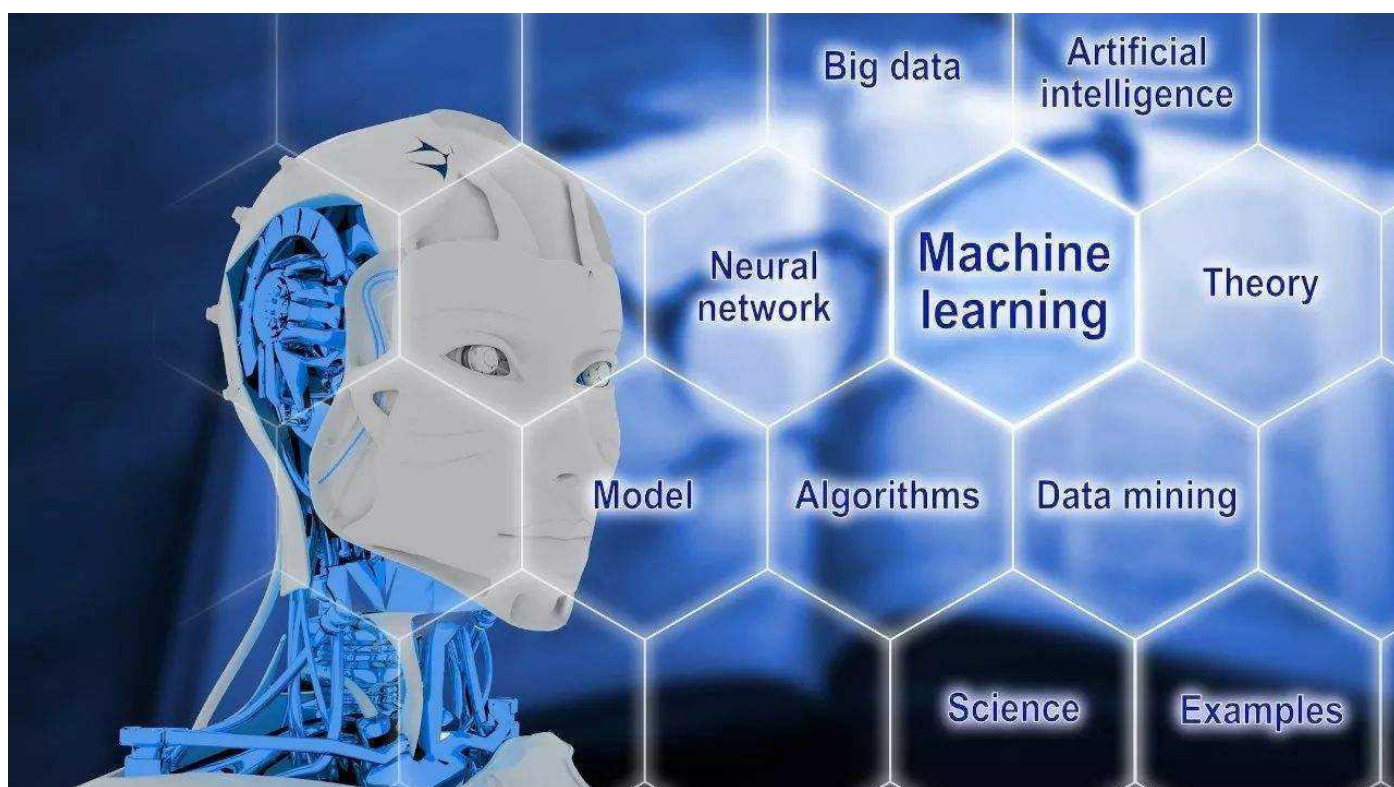
【独家科研项目】

【实地】机器学习实战：从房价预测到疾病预防

科研实训项目

发布时间：2018 年 12 月 10 日

招生周期：暑假开班，同一时间参与科研项目的上限人数为 25 人



机器学习是人工智能的核心

人工智能产业是驱动社会发展的重要力量！

一、项目背景

人工智能是继互联网和移动互联网浪潮之后，又一个驱动社会进步发展的重要力量。人工智能的意义之一，是将人们从重复性，非创造性的脑力劳动中解放出来，提高人们工作效率的同时，帮助人们更好更快的认识我们的世界。

在商业，科学，医疗，教育等众多领域，人工智能展现出极其激动人心的应用前景，或者已经在这些领域中展现出卓有成效的价值。

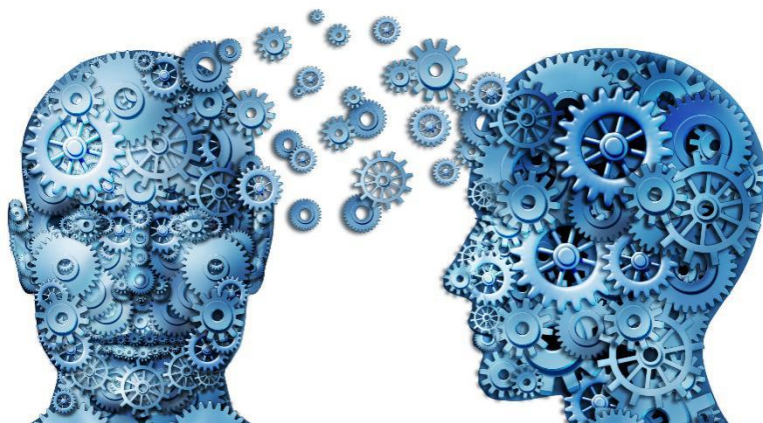
在众多的商业案例中，各大公司通过基于人工智能技术的推荐算法，利用用户产生的海量浏览数据，为用户带来无痛的产品推荐体验，同时，也实实在在提高了公司的营收。

在最近的人工智能应用中，他帮助医生诊断病了病情，预防疾病的发生。在科学界，他帮助科学家设计新药的结构，推断蛋白质的 3D 结构。

随着人工智能技术的发展，以及和生活各方各面的深度结合，人工智能的技术势必迸发出更大的活力。

机器学习本质上是一种智能算法的大类，他是人工智能这个重大话题的一个重要组成部分，你甚至可以在狭义上将目前的人工智能应用和机器学习的应用划上等号。诸如卷积神经网络和循环神经网络的深度学习技术，已经被反复验证，超越了传统的语义分析和图像识别的算法。

机器学习的算法将帮助人们发掘数据中的隐含信息，帮助人们预测那部分希望了解的结果，除此之外，机器学习中的强化学习算法，可以让智能体自动的和环境交互，这种更接近人们通常了解的人工智能技术，已经被应用于机器人控制和游戏 AI 的设计。



最近诸如 GANs 生成对抗网络在图像生成领域也展现出机器激动人心的成果。

二、师资介绍

李教授【个人介绍】

南京理工大学电子工程与光电技术学院教授，博导，国家特聘专家。曾主持多项国家级科研项目，在通信领域

有着突出贡献。

三、科研课题介绍

机器学习实战：从房价预测到疾病预防

(1) 项目背景

此次项目，沿着机器学习的一个**经典案例**—房价预测的数据处理到系统部署的全过程，带领大家切实感受机器学习在现实生活中的应用，利用这个项目，使得大家掌握处理数据，分析数据，利用机器学习算法的技能，为大家今后进一步研究机器学习算法做准备。通过房价预测这一案例的训练，我们将带领大家进行机器学习**进阶使用案例**，即利用大家日常采集的身体特征，利用机器学习的算法进行疾病预防系统的开发。

房价预测：

Kaggle 成立于 2010 年，是主要为开发商和数据科学家提供举办机器学习竞赛、托管数据库、编写和分享代码的一个进行数据发掘和预测竞赛的在线平台，已经吸引了全球超过 80 万名数据科学家的关注。Kaggle 官方每年会举办一次大规模的竞赛，奖金高达一百万美金，吸引了广大的数据科学爱好者参与其中。Kaggle 一直致力于解决业界难题，因此也创造了一种全新的劳动力市场——不再以学历和工作经验作为唯一的人才评判标准，而是着眼于个人技能，为顶尖人才和公司之间搭建了一座桥梁。



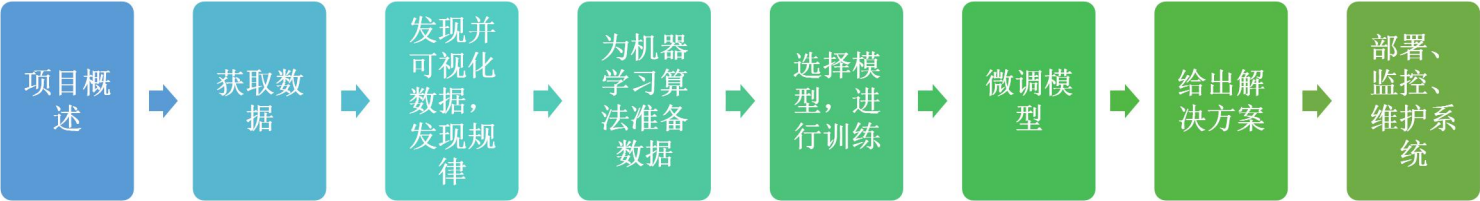
Kaggle 房价预测竞赛作为 kaggle 的入门项目，主要是利用回归分析的算法来预测房价。从 2016 年 8 月开始，到 2017 年 2 月结束。这段时间内，超过 2000 多人参与比赛，选手采用高级回归技术，基于我们给出的 79 个特征，对房屋的售价进行了准确的预测。通过这个项目，可以全面的带领大家了解机器学习系统的部署全过程。

疾病预防：

作为未来国家战略和全球关注的焦点，中国预防医学的黄金时代已然来临。中国双高人群众多，原因复杂，是脑卒中、冠心病等严重疾病发生的主要原因。中国国家医疗系统每年生产和管理着数以千万计的家庭。如何推动健康数据和 AI 的融合，实现双高早期筛查并及时干预，对



延缓疾病显性化和防止慢病重症化的有着重要意义。



(2) 课程思路

(3) 项目安排

课程安排时间	课程内容	备注
前 2 周	自主学习	预习或学习微积分、概率论、线性代数
第一天课程安排	从零开始机器学习	通过例程展示，培养兴趣
	- 展示机器学习趣味例程 - Python 初步编程 - 常用 python 库	
第二天课程安排	机器学习入门	初步讲解机器学习概念
	- 机器学习环境设置 - 数据及其处理 - 分类问题 - 回归问题	
第三天课程安排	机器学习进阶	对常用算法进行讲解
	- SVM - 决策树 - 集成算法等	
第四至五天课程安排	机器学习实战演练—房价预测	从数据清洗，到提出解决方案，通过完整的案例来深入对机器学习的理解和使用
	一套完整的机器学习项目，利用数据集进行房价预测	
第六天课程安排	神经网络详解	机器学习过渡到深度学习，并详细解释概念
	- 机器学习与深度学习联系 - 全连接神经网络 - 卷积神经网络 - 循环神经网络	
第七天课程安排	深度学习实战演练	学习对神经网络的使用 从数据处理到系统部署
	基于卷积神经网络的手写数字识别	

第八天课程安排	实战检验—疾病预防	
	● 数据处理到系统部署产生 baseline ● 问题解答	利用先前的房价预测案例中实践过的方法，进行疾病预防案例的系统部署，并产生 baseline
第九天课程安排	分组对抗—疾病预防	
	● 分出若干组 ● 一天的时间内，在 baseline 的基础上提出各自方案 ● 实践方案并调整 ● 上线并评分	

【保密要求】

由于所有项目中的科研导师所研究课题均为国家专项课题，均具有保密条款，学员需签署相关保密协议。

四、项目优势

顶级师资团队授课，国家级重点实验室学习

南京理工大学电子工程与光电技术学院教授，博士生导师领衔指导科研，理论与实践相结合，深入国家级重点实验室进行科研学习。

独家课程定制，多学科理论知识相融合

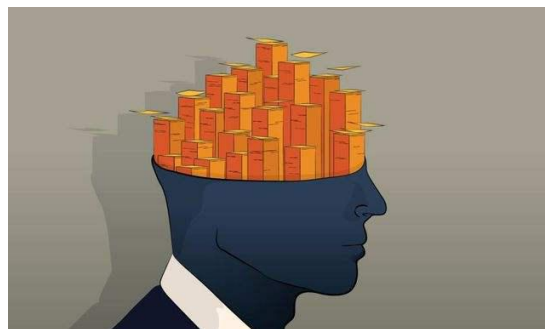
最前沿的人工智能领域基础课程，涉及到通信工程、电子信息工程、计算机技术、人工智能等相关专业，在讲解理论的同时，配合多个项目实战，使学生学习更加透彻。

科研级计算机仿真实验室

学生将深入南京理工大学雾连智能移动实验室进行科研研究，该实验室拥有多台服务器级别仿真计算机，可进行强化学习相关的代码仿真调试。

深入了解前沿科技的实际应用

了解人工智能领域在实际生活中的应用，并尝试利用机器学习解决实际生活中的真实案例，将理工科知识转换为真正社会价值。



五、项目成果

- 1、**实习证明**：南京理工大学电子工程与光电技术学院雾连智能移动实验室课程证书；
- 2、**高含金量的推荐信**：南京理工大学教授**个性化推荐信**（20%优秀学生可获得）；
- 3、**科研成果**：合作完成多个实训项目，项目组之间经验交流，完成科研报告；
- 4、**论文发表**：根据阶段性的科研经历和成果，特别优异的优秀成果可推荐在**国内外学术期刊或国际会议**上发表。
- 5、**国际会议展示机会**：优秀学员可推荐参加国内、国际学术峰会交流展示；
- 6、**多项能力的提高**：提高在课题背景下的分析、归纳、总结能力与科研逻辑思维，尤其是科研创新中的基本态度、技能、习惯；提高和强化学员的资料搜集和处理能力、书面表达能力、动手操作能力、人际交往等多方面能力。
- 7、**掌握前沿学科**：掌握强化学习有关知识，在人工智能领域迈出一大步；
- 8、**顶级人脉关系的搭建**：与专业领域专家面对面的沟通和学习，有可能获得更多的学习深造机会；与志同道合的科研团队成员合作，为将来的学业、职业发展储备优质的人脉资源。
- 9、**长期科研助理**：表现特别优异的，亦可有机会长期进入科研团队担任**长期科研助理**。

六、科研基本信息

- 1、**项目时间**：9天，2019年7月
- 2、**项目地点**：中国-南京 南京理工大学
- 3、**招生人数**：15-25人；
- 4、**招生截止时间**：2019年6月31日
- 5、**招生对象**：大学生及部分优秀高中生
 - 1) 适用于申请通信工程、电子信息工程、计算机技术、人工智能等相关专业的理工科学生；
 - 2) 适用于对人工智能方向感兴趣，有一定理工科基础，尤其是逻辑思维较强的学生；
 - 3) 有一定科研背景、需要提升科研能力的学生优先考虑。
- 6、**项目服务费用**：¥13800元，食宿费¥2500元（**自选**）
 - ✧ 费用包含：项目期间课程的实验费用、教学器材费（不包含器材损坏赔偿费）、保险费，以及远程指导费用。
 - ✧ 费用不包含：学员从自己所在城市往返调研地点的交通费，食宿费（自选），及因个人原因产生的额外费用。
- 7、**报名方式**：学生需填写报名表并参与面试，面试通过即可参加项目。

