

Udacity AI VIP班挑战赛

说明文档

竞赛目的

1. 学以致用，通过实战竞赛消化课程中学习的模型及数据技能，为简历增加亮点
2. 了解主流的竞赛平台，如Kaggle和天池
3. 熟悉竞赛参与全流程，包括：应用场景、赛题理解、数据分析、特征工程、模型搭建、结果上传、模型优化等

竞赛时间

2018年12月10日 - 2018年12月23日

p.s. 我们建议大家利用这两周时间集中完成竞赛项目，我们也会在这两周内为大家提供经验丰富的竞赛助教提供指导。竞赛时间过后，建议学员回归到正式课程继续学习，以免影响学习进度。学有余力的同学也可以继续在竞赛平台刷分，提高竞赛名次。

竞赛平台

kaggle

竞赛方式

1. 自行组队（最多不超过3人）并填写[申请表](#)（每支参赛队填写一次申请表即可）
2. 前往Kaggle网站注册并参赛：[竞赛地址](#)

竞赛内容

《通过对新闻文本的分析来实现股票预测》

现如今，越来越多的投资者使用无处不在的数据来制定投资决策。制定投资决策的挑战在于如何提取以及如何解释数据。[Two Sigma](#)对这一挑战充满热情，并很高兴与Kaggle社区分享。

作为一个科学驱动的投资管理者，Two Sigma已经将技术和数据科学应用于财务预测超过17年。他们在大数据，人工智能和机器学习方面的开创性进步推动了投资行业的发展。现在，他们渴望与Kagglers一起继续追求创新。

此次的任务是通过分析新闻数据来预测股票价格，数据说明可以点击[这里](#)查看。

本次比赛的数据来自以下来源：

- 市场数据由Intrinio提供。
- 路透社提供新闻数据。

Kaggle 竞赛时间：

- 开始时间: 9/25/2018
- 组队截止时间: 1/2/2019
- 提交截止时间: 1/8/2019
- 结束时间: 7/15/2019

竞赛辅导日历

	12.3 Mon	12.4 Tue	12.5 Wed	12.6 Thu	12.7 Fri	12.8 Sat
				公开赛题； 自由组队（1-4人）并报名； 平台注册； 理解赛题并整理疑问；		
12.9 Sun	12.10 Mon	12.11 Tue	12.12 Wed	12.13 Thu	12.14 Fri	12.15 Sat
*第一次直播	竞赛开始		第一次集中在线答疑 (20:30-21:30)			
12.16 Sun	12.17 Mon	12.18 Tue	12.19 Wed	12.20 Thu	12.21 Fri	12.22 Sat
第二次集中在线答疑 (20:30-21:30)				第三次集中在线答疑 (20:30-21:30)		
12.23 Sun						
*第二次直播	竞赛辅导结束，学生可选择继续刷竞赛排名					

*竞赛辅导期间共有两次直播：

- 第一次直播讲解包含应用场景、赛题理解、数据分析、特征工程、模型搭建、结果上传、模型优化等内容；
- 第二次直播每班报名一个代表队展示自己的项目思路，由竞赛助教点评分析（报名队伍可额外获得Udacity周边礼物一份！）

*竞赛辅导期间共有三次集中答疑时间，每次1小时，助教将在微信群内通过文字、语音、图片或链接等方式集中解决大家的共性或个性问题；为保证高效的答疑效果，建议学员提前提交自己的问题。

*在竞赛助教集中答疑时间之外，学员如遇到安装环境、编程报错等问题，可在竞赛群内自主讨论；如遇到较为复杂的问题，建议经过个人思考过后提交至问题池，由竞赛助教集中comment或在下一次集中答疑时间内统一在群内解答。

竞赛指导助教简介

钱乾

- 毕业于佐治亚理工学院，目前于上大岂网络科技有限公司担任数据科学家，负责设计研发实现公司的AI产品，包括但不限于推荐引擎，知识图谱，对话系统等。
- 擅长自然语言处理与推荐法，对分类回归等传统机器学习有充足的经验，对于任何机器学习类场景的数据集都能快速上手并应用合适的方法解决问题。现专攻基于深度学习实现对话系统与知识图谱结合的应用。
- Kaggle Grand Master，Kaggle全球最高排名46，华人排名top10。
- **获奖经历：**
 - Kaggle Porto Seguro 2nd price
 - Bytecup 2016 2nd price (leader)
 - Tianchi IJCAI-17 Koubai 10th price
 - Kaggle Santa's uncertain bags 3rd place
 - Data Castle Job prediction 9th place
 - Kaggle Home Depot Product Search Relevance 15th