


# PPO x Family 决策 AI 入门公开课

## 常见问题集 (2023.2.24)

### 目录

1. PPO x FAMILY 决策 AI 入门公开课主要讲什么	2
2. 教学大纲	3
3. 课程时间安排	4
4. 想要获得学习证书，我应该怎么做	4
5. 很想学习这门课，担心我的计算资源跟不上怎么办	6
6. 我可以在哪里学习课程	7
7. 学习中遇到问题，我可以在哪里提问	9
8. 如何加入官方 SLACK 频道	10
9. 课程是中文的吗	11
10. 课程是直播课吗	11
11. 我是初学者，初次学习应该预习哪些知识	11
12. 课程 GITHUB 仓库什么时候更新	12
13. 作业题在哪里可以看到，以及如何提交作业	12
14. 如何获得课程学习资料	13
15. 用手机在 B 站学习课程的时候，字幕被挡住了怎么办	15
16. 课程中会使用到的 OPENDILAB 开源决策 AI 平台是什么？怎么用呢	16
17. 课程会讲解多智能体强化学习吗	16
18. 课程是否提供在线训练环境	16
19. 我有 AI 技术问题想和大家讨论，能否举办一些交流活动	16

# 决策 AI 的探索者，欢迎你加入 PPO × Family 决策 AI 系列课程

 这里是 PPO × Family 决策 AI 入门公开课常见问题集，我们会持续更新这个文档，方便大家更容易了解我们的课程，也许你的疑问会在这个文档里找到哦。

## 1. PPO × Family 决策 AI 入门公开课主要讲什么



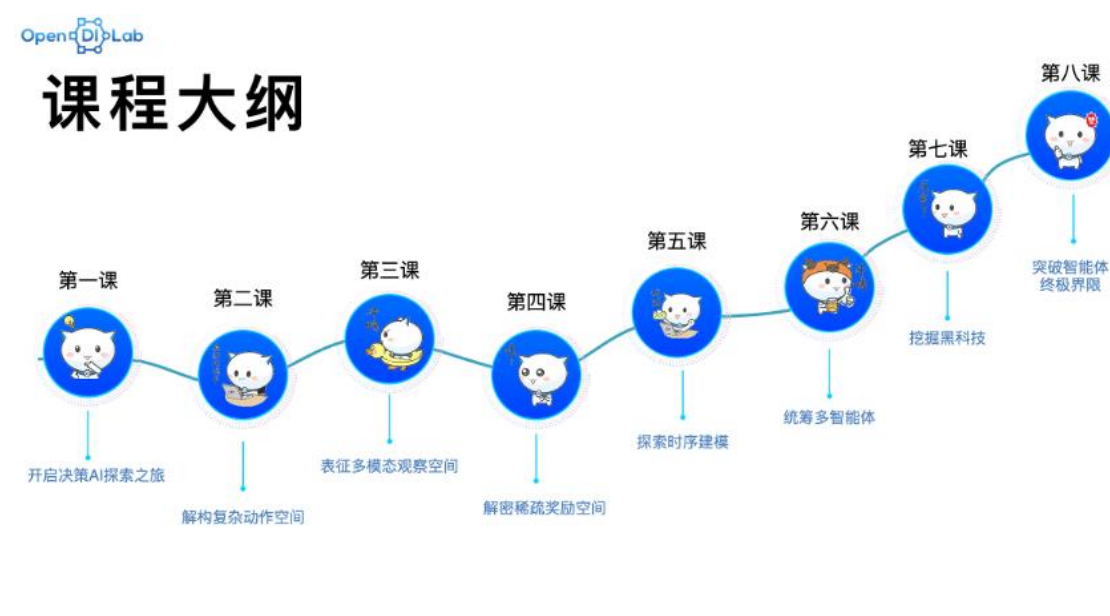
决策智能入门级公开课 PPO × Family 旨在运用一种最经典的**深度强化学习算法 Proximal Policy Optimization (PPO)**，解决各种各样的决策智能问题，帮助一切对于深度强化学习技术有好奇心的人，轻便且高效地制作应用原型，了解和学习最强大最易用的 PPO × Family。在上完这门课程后，如果有进一步兴趣的话，受众可以将所学知识应用到相关领域中，利用更多的计算资源，去追逐和挑战更强大的人工智能。

该课程由中国计算机学会主办，上海人工智能实验室承办，全球高校人工智能学术联盟、北京大学人工智能研究院多智能体中心、浙江大学上海高等研究院、商汤科技协办，知乎、机器之心、智海平台、Paperweekly 支持。顺利结课的学员还可获得由中国计算机学会颁发的课程学习认证证书。

详情可点击：

[https://www.bilibili.com/video/BV1sK411R7JP/?spm\\_id\\_from=333.337.search-card.all.click&vd\\_source=3af305df8f20566dc8518dd039e4c9a8](https://www.bilibili.com/video/BV1sK411R7JP/?spm_id_from=333.337.search-card.all.click&vd_source=3af305df8f20566dc8518dd039e4c9a8)

## 2. 教学大纲



本系列课程将分为 8 节线上公开课，每节课 40-50 分钟。

第一课：开启决策 AI 探索之旅

第二课：解构复杂动作空间

第三课：表征多模态观察空间

第四课：解密稀疏奖励空间

第五课：探索时序建模

第六课：统筹多智能体

第七课：挖掘黑科技

第八课：突破智能体终极界限

在 PPO x Family 八节课的过程中，受众会接触到各种各样的决策智能应用，最终结课时，就相当于完成了自动驾驶、量化交易、机器人控制、游戏 AI 等多个领域相关决策任务的入门实践，从而真正做到集中一点，演化出无限可能。



### 3. 课程时间安排

具体课程时间安排可以定期关注[官方 GitHub 主页](#)和[微信交流群公告](#)（还没有加入课程交流群的朋友，可添加课程小助手微信号：OpenDILab）

我们的课程制作会根据大家的体验反馈进行课程内容和呈现方式的微调，第五节课预计在 2023 年 3 月中旬更新。

### 4. 想要获得学习证书，我应该怎么做

★如何获得[普通等级](#)的学习认证证书：

- 提交 8 节课布置的算法理论推导/实践小作业
- 积极参与课程讨论

完成如上任务，即可获得由中国计算机学会颁发的“通过学习”认证证书。

### ✨如何获得**优秀等级**的学习认证证书：

- 提交 **8** 节课布置的**算法理论推导/实践**小作业
- **积极参与课程讨论**
- 完成期末实训大作业，**搭建出独具创意、有趣的决策智能体**

完成如上任务，即可获得由**中国计算机学会**颁发的“优秀”认证证书。

### ✨关于有效作业提交：

- 我们希望能够启发同学们对算法本质和原理的思考，所以对于大家的作业并不要求全对，尽力就好
- 如果有一道题实在不会，可以写下自己的思路，**提交适当的思路和解题过程即可视为合格**
- 有效作业提交必须是在截止时间前发送邮件/发起 Pull Request
- 可不能因为学习要求放宽松了就偷懒哦，还是要好好写作业，好好思考👍

### ✨关于补交作业：

- 我们会在课程期末给大家两次补交作业的机会，最终 **8** 次作业都有相应的有效提交记录，才可以获取证书
- 每次只可提交一个课时的作业

### ✨关于后续作业中的**数学题和代码题**：

- 从第二节课开始，我们给大家布置的作业，一般会有两种题型，**数学题和代码题**
- 大家可以根据自己的兴趣，选择做数学题还是代码题，**二选一提交即可**（当然两类都做是最受欢迎的）
- 选择做数学题，则需要通过邮箱提交作业
- 选择代码题，需要在官方 **GitHub** 发起 Pull Request，提交作业

### P.S.

- 统一给大家颁发证书的时候，记得给 DI 小助手（微信 ID：OpenDILab）私信自己在微信/Slack/GitHub 的用户名
- 提交作业，要保持每次提交邮箱/Github 账号和姓名统一，否则会影响证书发放哦

## 5. 很想学习这门课，担心我的计算资源跟不上怎么办

OpenDILab 也特地优化了每节课的实践任务的计算开销，整个 PPO × Family 系列课程并不需要大量计算资源，一定配置的个人电脑完全可以应对各种挑战。

如果个人电脑不足以支持课程所需，课程官方也提供了免费的计算资源。大家可以使用 [autodl](#) 平台获取。具体的使用说明如下：

- 登陆 [autodl](#) 平台。课程官方创建了一个公共账号，账号名称为：18016297479 对应密码为：Opendilab1234567。
- 点击左上角算力市场，选择合适的算力资源。温馨提示：由于课程的 demo 需要算力不大，经过测试，即便是最低配置的一张 2080Ti 都可以顺利完成所有任务。为了大家都可以方便地使用，希望同学们可以珍惜计算资源，不要浪费选择过高配置。



- 创建实例。在最下方的镜像处点击“我的镜像”，下拉框选择 **di-engine**。这个镜像 是课程官方为方便大家进行尝试，配置好了相关的安装依赖。如果大家想要重新进行 配置，可以选择合适版本的基础镜像，并在此基础上安装相关的依赖。



- 测试代码。如果配置无误，在点开生成的机器之后，在 **terminal** 中输入：

```
Shell
xvfb-run --auto-servernum --server-num=1 --server-args="-screen 1
1024x768x24" python rocket_landing_demo.py
```

这是一段运行 **rocket** 环境的简单代码，执行半分钟后如果正常退出，就证明整个流程 无误。

- 样例代码。课程官方还提供了样例代码，存放在了镜像中的 **demo.ipynb** 里。需要 说明的是，由于镜像已经提供了所有的安装内容，因此有关安装的部分均无需执行。
- **注意事项**：关于 **autodl** 平台的使用问题，大家可以多参考其[文档](#)；希望大家节约 资源，不要租用过贵的计算资源；跑完代码之后，及时关机，避免持续扣费；尽量不 要用公共资源做无关课程的事情；不要自己更改密码。

在上完这门课程后，如果有进一步兴趣的话，受众可以将所学知识应用到相关领域中， 利用更多的计算资源，去追逐和挑战更强大的人工智能。

## 6. 我可以在哪里学习课程

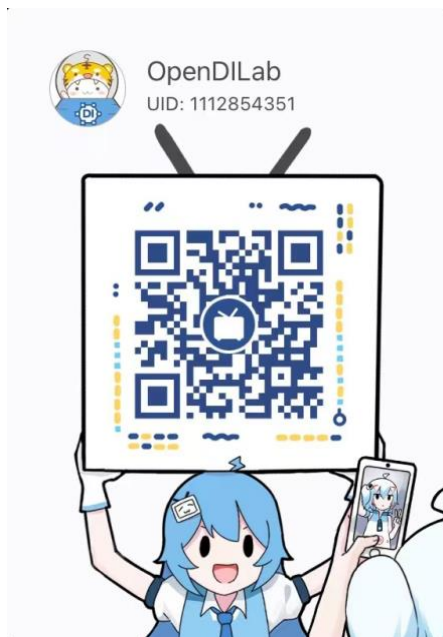
目前，OpenDILab 已经建立了 课程 [GitHub](#) 相关仓库，每一节课程的课件、文字 讲稿、代码、答疑等教学素材都会全部公开在这个仓库中。而每节课的课程录像，将 会在 B 站（账号：OpenDILab）、公众号平台（账号：OpenDILab 决策 AI）、知乎平 台（账号：OpenDILab 浦策）同步发布，欢迎保持关注。

**PPO x Family** 课程 [GitHub](#) 链接：

<https://github.com/opencv/PPOxFamily>



✦官方 B 站二维码：



✦官方知乎二维码：





🌟官方公众号二维码:



另外，由于剪辑和制作需要，同时响应受众的需求不断优化课程内容，我们将会以每1-2周一节课的速度逐渐更新课程内容，也欢迎大家积极参与贡献，希望能够真正将知识传达给有需要的人，共建知识共享的开源社区。

## 7. 学习中遇到问题，我可以在哪里提问

- 在 **PPO × Family** 官方 **GitHub Issues** 专区提问：

<https://github.com/pendilab/PPOxFamily/issues>

- 在官方课程微信交流群提出课程相关的非技术问题：

添加官方课程小助手微信，小助手将邀请您加入官方课程微信交流群，我们会在微信群中进行简单的课程相关的问题解答。



扫一扫上面的二维码图案，加我为朋友。

官方课程小助手微信二维码

- 在官方课程小助手发布的调研问卷/官方 **slack** 频道中提出课程相关的技术问题：

每周结课后，小助手都会在群内发布问卷，来收集大家在学习过程中遇到的疑难问题，会通过 Slack 私信回复大家，另外还会不定时举办研讨会，方便大家理解课程内容。

所以中高阶的学习问题，请一定要填写在问卷里哦，我们会根据问卷结果确定研讨会的主题。P.S.一定要在问卷中填写自己的 Slack ID 哦。

## 8. 如何加入官方 Slack 频道

点击下方链接即可加入：

[https://opendilab.slack.com/join/shared\\_invite/zt-v9tmv4fp-nUBAQEH1\\_Kuyu\\_q4pIBssQ#/shared-invite/email](https://opendilab.slack.com/join/shared_invite/zt-v9tmv4fp-nUBAQEH1_Kuyu_q4pIBssQ#/shared-invite/email)



我们建议使用你的工作电子邮件帐户。



## 9. 课程是中文的吗

课程是中文授课的哦，课程 PPT 也是中文的～

## 10. 课程是直播课吗

我们采用边播“边复盘边优化”的课程迭代模式，会实时结合大家的反馈及需求，去优化和调整课程内容和模式，希望能够给大家持续带来良好的上课体验，同时也希望大家持续关注我们的课程并及时向我们反馈您的建议。

## 11. 我是初学者，初次学习应该预习哪些知识

学习这门课，需要了解一定的机器学习、强化学习基础概念，另外还需要提前熟悉 **PyTorch**。

✨学习这门课程，需要了解哪些机器学习基础概念：

了解数据集、训练/测试、监督学习、欠拟合和过拟合等等这些基本名词即可

大家可以从李宏毅老师的机器学习视频课中提前了解：

[https://www.bilibili.com/video/BV1m3411p7wD/?spm\\_id\\_from=333.337.search-card.all.click&vd\\_source=3af305df8f20566dc8518dd039e4c9a8](https://www.bilibili.com/video/BV1m3411p7wD/?spm_id_from=333.337.search-card.all.click&vd_source=3af305df8f20566dc8518dd039e4c9a8)

✨学习这门课程，需要了解哪些强化学习基础概念：

了解马尔科夫决策过程，尝试思考下想用强化学习解决什么样的决策问题

参考资料：

[https://di-engine-docs.readthedocs.io/zh\\_CN/latest/10\\_concepts/basic\\_rl\\_zh.html](https://di-engine-docs.readthedocs.io/zh_CN/latest/10_concepts/basic_rl_zh.html)

✨另外还需提前熟悉的：

PyTorch，包括 Tensor、自动微分机制等等，后续课程相关代码全部使用 PyTorch 哦

参考资料：

<https://pytorch.org/tutorials/beginner/basics/intro.html>

## 12. 课程 GitHub 仓库什么时候更新

第 1-4 节课的课件已更新在 [GitHub](#)，课程的相关材料都会在每节课发布后更新到 GitHub 仓库中，谢谢关注！

如有同学登陆 GitHub 遇到困难，可以在 [Gitee](#) 浏览学习资料。

❤️欢迎大家向我们提出建议，此文档会持续更新下去哦，感谢大家的支持与鼓励！

## 13. 作业题在哪里可以看到，以及如何提交作业

- 作业题我们会放在官方 GitHub 仓库的文件夹中，同时也会在内容导航区为大家贴出对应链接

### 内容导航

章节（视频课）	算法理论资料	补充资料	习题	代码样例	应用样例
第一章：开启决策AI探索之旅	课程PPT 课程文字稿	微课视频 策略梯度 A2C TRPO 符号表 QA总结	习题 习题题解	PG算法示例 A2C算法示例 PPO算法示例	应用混剪
第二章：解构复杂动作空间	课程PPT	重参数化 PPO&DDPG HyAR QA总结	习题 习题题解	离散动作示例 连续动作示例 混合动作示例 应用训练代码	火箭回收等
第三章：表征多模态动作空间	课程PPT	表征学习 PPG 不变性 QA总结	习题 习题题解	编码方法示例 Wrapper示例 计算图示例 应用训练代码	软体机器人等
第四章：解密稀疏奖励空间	课程PPT	逆强化学习 行为克隆BC	习题	[RND示例] [Pop-Art示例] 应用训练代码	自动驾驶等

- 当您提交理论题作业，请发送邮件至 [opendilab@pjlabor.org.cn](mailto:opendilab@pjlabor.org.cn)

请同学们一定按照下方格式命名邮箱主题：

【PPO × Family】+学生名 + vol.1 (第几节课) +作业提交日期

示例：【PPO × Family】+喵小 DI + vol.1 +20221207

- 当您提交代码题的时候，需要在 [PPO × Family](#) 官方 GitHub 上发起 [Pull Request](#)

提交地址：详情见每次作业说明部分

[PR 示例见此超链接](#)

请同学们一定按照下方格式命名 Pull Request：

hw\_submission(学生名称): add hwn (第几节课) +作业提交日期

示例：[hw\\_submission\(nyz\): add hw2\\_20230104](#)

- 提交作业，要保持每次提交邮箱和姓名统一，方便课程小助手统计作业信息

## 14. 如何获得课程学习资料

课程学习资料我们都会第一时间在官方 GitHub 主页放出，您可以选择对应的 chapter 文件夹获取您所需要的课程资料。

PaParaZz1

style(nyz): fix table typo

cef7d74 18 hours ago 74 commits

assets	style(nyz): add compressed application img	2 months ago
chapter1_overview	polish(nyz): polish ch1 code example	last week
chapter2_action	feature(nyz): add chapter2 qa and hw solution	last month
chapter3_obs	feature(nyz): add chapter3 qa	4 days ago
chapter4_reward	feature(nyz): add ch4 application demo code	18 hours ago
chapter5_time	style(nyz): update ch4 pre info	3 days ago
common	polish(nyz): update hw3 info and notation table	2 weeks ago
.gitignore	feature(nyz): add chapter2 application demo	last month
LICENSE	Initial commit	3 months ago
README.md	style(nyz): fix table typo	18 hours ago

您也可以在内容导航区找到对应学习资料。

内容导航

章节（视频课）	算法理论资料	补充资料	习题	代码样例	应用样例
第一章：开启决策AI探索之旅	课程PPT 课程文字稿	微课视频 策略梯度 A2C TRPO 符号表 QA总结	习题 习题题解	PG算法示例 A2C算法示例 PPO算法示例	应用混剪
第二章：解构复杂动作空间	课程PPT	重参数化 PPO&DDPG HyAR QA总结	习题 习题题解	离散动作示例 连续动作示例 混合动作示例 应用训练代码	火箭回收等
第三章：表征多模态动作空间	课程PPT	表征学习 PPG 不变性 QA总结	习题 习题题解	编码方法示例 Wrapper示例 计算图示例 应用训练代码	软体机器人等
第四章：解密稀疏奖励空间	课程PPT	逆强化学习 行为克隆BC	习题	[RND示例] [Pop-Art示例] 应用训练代码	自动驾驶等

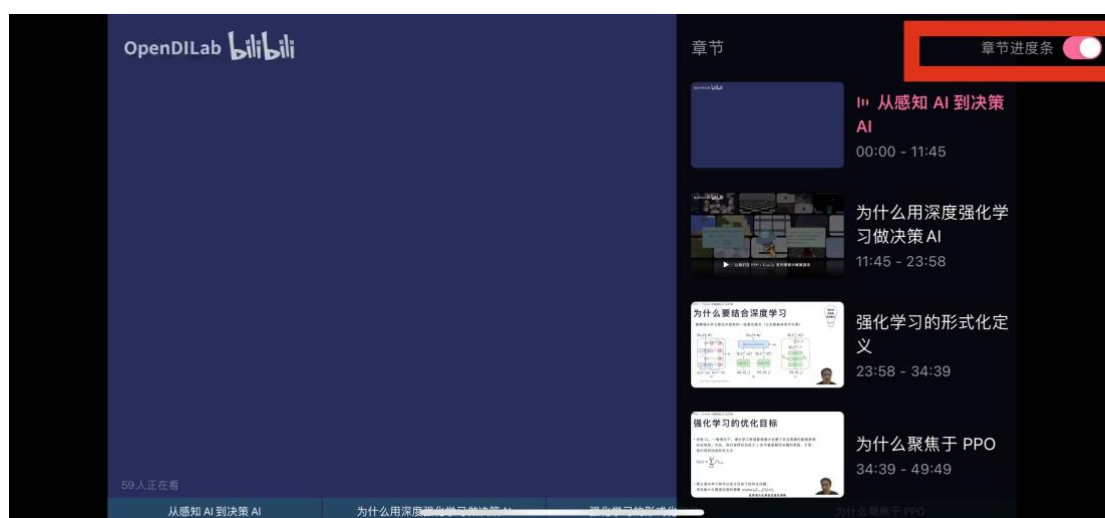
## 15. 用手机在 B 站学习课程的时候，字幕被挡住了怎么办

遇到字幕被章节进度条挡住的情况，不用担心，小助手来帮你啦～

### 步骤一：点击图示位置：“章节”



### 步骤二：关闭“章节进度条”





## 16. 课程中会使用到的 OpenDILab 开源决策 AI 平台是什么？怎么用呢

OpenDILab 开源决策 AI 平台是上海人工智能实验室开源平台体系 OpenXLab 的重要组成部分，是国际上首个覆盖最全学术界算法和工业级规模的决策 AI 平台，具备一套完备的决策 AI 训练与部署框架，自顶而下涵盖了包括应用生态层，算法抽象层，异步分布式管理层和最底层的分布式执行层，适配从单机到万 CPU/GPU 联合训练的全尺度调度系统优化。

PPO x Family 课程的具体代码样例都会配合 OpenDILab 的相关工具库讲解使用哦~

## 17. 课程会讲解多智能体强化学习吗

PPO x Family 课程的第 6 讲就是多智能体专题，MAPPO 和 HAPPO 都会涉及哦，敬请期待。

## 18. 课程是否提供在线训练环境

第一节课是整体概述课，从第二节课起会有实际训练的例子，后续的应用 demo 都做过一些优化，大家在个人 pc 上训练一两个小时就能得到效果不错的智能体。

## 19. 我有 AI 技术问题想和大家讨论，能否举办一些交流活动

我们会举办前沿技术交流分享会 Insight time 以及 DI 极客说用户交流会等，可以在这些活动上交流哦~