

李宏毅 (Hung-yi Lee) · HYLEE | Machine Learning (2021)

HYLEE(2021) · 课程资料包 @ShowMeAI



视频

中英双语字幕



课件

一键打包下载



笔记

官方笔记翻译



代码

作业项目解析



视频 · B 站 [扫码或点击链接]

<https://www.bilibili.com/video/BV1fM4y137M4>



课件 & 代码 · 博客 [扫码或点击链接]

<http://blog.showmeai.tech/ntu-hylee-ml>

机器学习
深度学习

Auto-encoder
卷积神经网络

生成式对抗网络
GAN 自监督

学习率
自注意力机制

批次标准化

神经网络压缩

强化学习

元学习

Transformer

Awesome AI Courses Notes Cheatsheets 是 [ShowMeAI](#) 资料库的分支系列，覆盖最具知名度的 **TOP50+** 门 AI 课程，旨在为读者和学习者提供一整套高品质中文学习笔记和速查表。

点击课程名称，跳转至课程**资料包**页面，**一键下载**课程全部资料！

机器学习	深度学习	自然语言处理	计算机视觉
Stanford · CS229	Stanford · CS230	Stanford · CS224n	Stanford · CS231n
# Awesome AI Courses Notes Cheatsheets · 持续更新中			
知识图谱	图机器学习	深度强化学习	自动驾驶
Stanford · CS520	Stanford · CS224W	UCBerkeley · CS285	MIT · 6.S094



微信公众号

资料下载方式 2: 扫码点击**底部菜单栏**
称为 **AI 内容创作者**? 回复 [添砖加瓦]

Machine Learning HW12

ML TAs

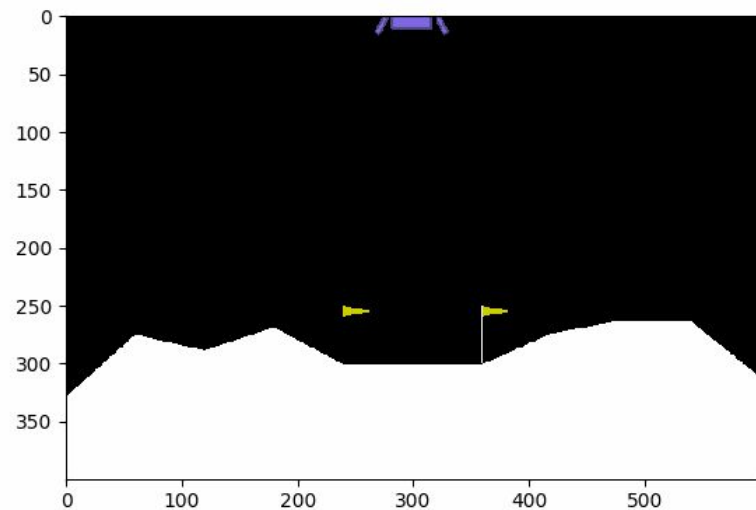
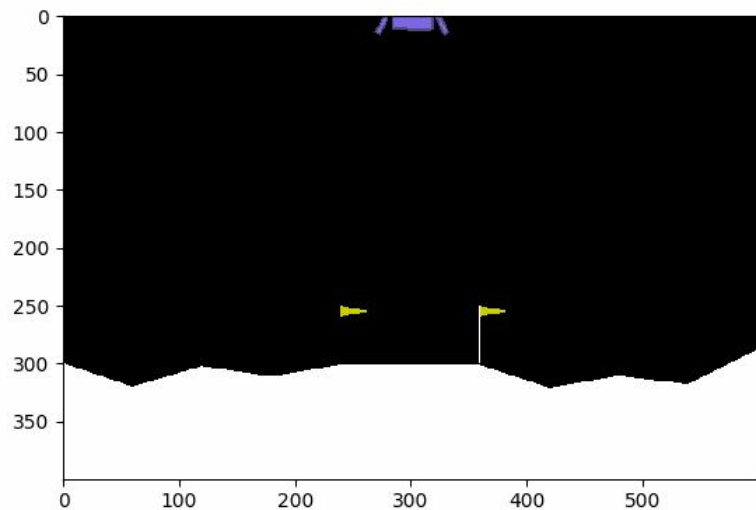
ntu-ml-2021spring-ta@googlegroups.com

作業內容

在本次作業當中，你們將可以實做幾項 Deep Reinforcement Learning 方法：

- Policy Gradient
- Actor-Critic
- 作業的實做環境為 OpenAI 的 gym 當中的 [Lunar Lander](#)。其餘實做細節請參考助教提供的範例程式。

範例展示



Policy Gradient 方法(to get 8 points)

Algorithm 1 Policy Gradient

function REINFORCE

 Initialize policy parameters θ

for each episode $\{s_1, a_1, r_1, \dots, s_T, a_T, r_T\} \sim \pi_\theta$ **do**

for $t = 1$ to T **do**

 Calculate discounted reward $R_t = \sum_{i=t}^T \gamma^{i-t} r_i$

$\theta \leftarrow \theta + \alpha \nabla_\theta \log \pi_\theta(a_t | s_t) R_t$

end for

end for

return θ

end function

Actor-Critic 方法(to get 10 points)

Algorithm 2 Actor-Critic

function REINFORCE WITH BASELINE

 Initialize policy parameters θ

 Initialize baseline function parameters ϕ

for each episode $\{s_1, a_1, r_1, \dots, s_T, a_T, r_T\} \sim \pi_\theta$ **do**

for $t = 1$ to T **do**

 Calculate discounted reward $R_t = \sum_{i=t}^T \gamma^{i-t} r_i$

 Estimate advantage $A_t = R_t - b_\phi(s_t)$

 Re-fit the baseline by minimizing $\|b_\phi(s_t) - R_t\|^2$

$\theta \leftarrow \theta + \alpha \nabla_\theta \log \pi_\theta(a_t | s_t) A_t$

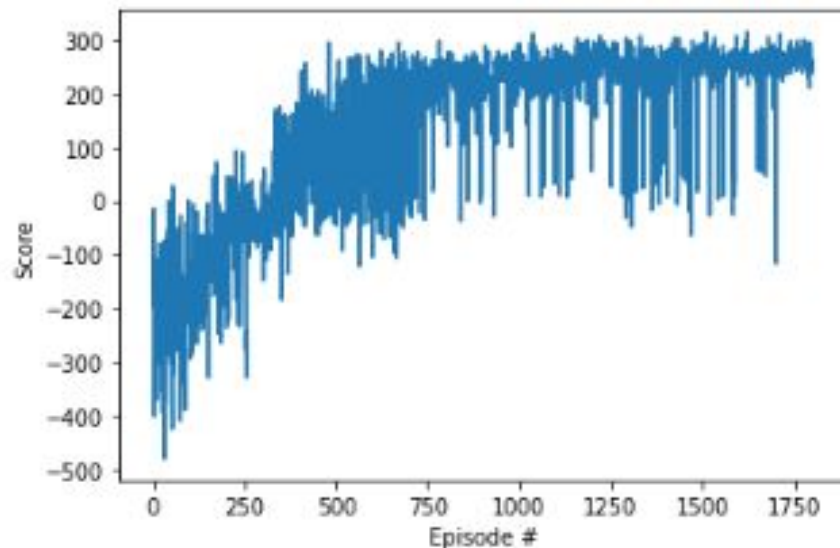
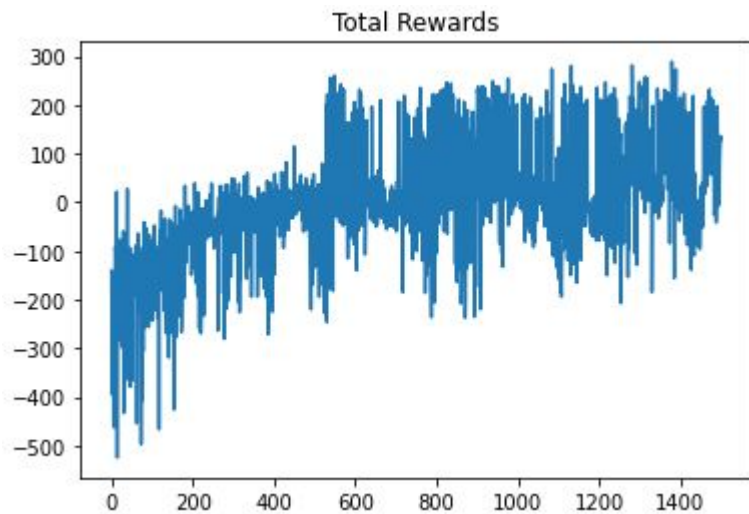
end for

end for

return θ

end function

範例結果



繳交項目及評分標準

1. Python 程式碼 (Submit on NTU COOL) 佔4分
2. Action List (Submit on JudgeBoi, 沒有**private set**, 自動選擇最高分)
3. 給分標準: (Your submission must be valid)

15min~20min

Avg_Reward		
< 0		2
0~99		3
100~199		4
200~240		5
241~		6

30min



繳交項目及評分標準

More on a "valid submission ":

agent在action list最後一個動作輸入之後，應該要輸出done。長度過長或過短的action list都會被系統reject。

Action list 的長相



```
1 print("Action list looks like ", action_list)
2 print("Action list's shape looks like ", np.shape(action_list))
```

```
Action list looks like  [[3, 3, 3, 3, 3, 3, 3, 3, 3, 3, 3, 3, 3, 2, 3, 2, 3, 2, 2, 2, 3, 2, 2, 2, 2, 2, 2, 2, 2, 2, 3, 2, 2, 2, 3
Action list's shape looks like  (5,)
```



Bonus

- If you successfully get 10 pts:
 - Your code will be made public to students.
 - You can submit a report in **PDF** format briefly describing what you have done (in English, less than 100 words) for **extra 0.5 pts**.
 - Reports will also be made public to students.
 - Notice, we do not have private score, so omit it in the report.
- [Report template](#)

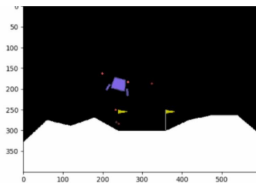
注意事項

- You should finish your homework on your own.
- You should NOT modify your prediction files manually.
- Do NOT share codes or prediction files with any living creatures.
- Do NOT use any approaches to submit your results more than 5 times a day.
- **Do NOT search or use additional data or pre-trained models.**
- Your **final grade x 0.9** if you violate any of the above rules.
- Prof. Lee & TAs preserve the rights to change the rules & grades.

注意事項

- 所有作業相關問題請在 NTU COOL詢問(推薦)或是寄信至助教信箱，並於信件主題處註明：**[HW12]**

當你試著想讓
火箭降落在旗子內
好完成HW12



Submit Deadline: 6/04 - 6/25 (23:59)

你家的狗：
那是什麼？
可以吃嗎？



李宏毅 (Hung-yi Lee) · HYLEE | Machine Learning (2021)

HYLEE(2021) · 课程资料包 @ShowMeAI



视频

中英双语字幕



课件

一键打包下载



笔记

官方笔记翻译



代码

作业项目解析



视频 · B 站 [扫码或点击链接]

<https://www.bilibili.com/video/BV1fM4y137M4>



课件 & 代码 · 博客 [扫码或点击链接]

<http://blog.showmeai.tech/ntu-hylee-ml>

机器学习
深度学习

批次标准化

Auto-encoder

卷积神经网络

生成式对抗网络

GAN

自监督

强化学习

元学习

学习率

自注意力机制

Transformer

Awesome AI Courses Notes Cheatsheets 是 [ShowMeAI](#) 资料库的分支系列，覆盖最具知名度的 **TOP50+** 门 AI 课程，旨在为读者和学习者提供一整套高品质中文学习笔记和速查表。

点击课程名称，跳转至课程**资料包**页面，**一键下载**课程全部资料！

机器学习	深度学习	自然语言处理	计算机视觉
Stanford · CS229	Stanford · CS230	Stanford · CS224n	Stanford · CS231n
# Awesome AI Courses Notes Cheatsheets · 持续更新中			
知识图谱	图机器学习	深度强化学习	自动驾驶
Stanford · CS520	Stanford · CS224W	UCBerkeley · CS285	MIT · 6.S094



微信公众号

资料下载方式 2: 扫码点击**底部菜单栏**
称为 **AI 内容创作者**? 回复 [添砖加瓦]