

# Artificial Intelligence



## 开栏：智能决策系列



冯伟  
Hulu推荐算法

已关注

wepon 等 22 人赞同了该文章

站在2017年末回首这一年的心路历程，一方面，我为能见证这个AI时代而感到兴奋，另一方面，又为跟不上大把大把的最新成果而感到焦虑。如果你也有同样的想法，那么欢迎来到这个专栏 - “零基础机器学习”。

思考了很久从哪个角度写才能恰如其分的既不脱离传统的机器学习，又能反应当前的AI进展，于是有了下面的选题（选题意义可以参见[监督学习越来越准，我为什么要写bandit问题](#)）

假设我们开了一家叫Surprise Me的饭馆，客人来了不用点餐，由算法来决定改做哪道菜：

- 问题：有N道菜，推荐哪个用户才会喜欢？
  - 核心方法：Multi-Armed Bandit
  - 核心问题：优化点击率
  - [第一讲: eps-greedy](#)
  - [第二讲: UCB](#)
  - [第三讲: Thompson Sampling](#)
- 拓展1：给定时间（早中晚餐）、客户的年龄、性别，推荐哪个用户才喜欢？
  - 核心方法：Contextual Bandit
  - 核心问题：充分利用周边信息优化点击率
  - [第一讲: LinUCB](#)
  - [第二讲: Thompson Sampling](#)
  - [Contextual Bandit 工业实践: 微软Azure Decision Service](#)
- 拓展2：如何不仅考虑用户单次的就餐，还考虑如何将餐馆打造成一个可以长期用餐的地方？  
无论一道菜多好吃，没有用户喜欢每天都吃同样的东西，我们用一系列的推荐满足用户长期的就餐需求。
  - 核心方法：(Deep) Reinforcement Learning
  - 核心问题：利用多次推荐来优化长期的目标
  - [第一讲: MDP基础概念](#)
  - [第二讲: 价值迭代 \(Value Iteration\)](#)
  - [第三讲: 策略迭代 \(Policy Iteration\)](#)

我会以每周一次以上的更新速度来完成这个系列，欢迎各位斧正！

欢迎订阅微信公众号“零基础机器学习”，搜索微信号：ml-explained

编辑于 2018-02-25

「真诚赞赏，手留余香」

赞赏

3 人已赞赏

[机器学习](#)[深度学习 \(Deep Learning\)](#)[强化学习 \(Reinforcement Learning\)](#)

文章被以下专栏收录



零基础机器学习

已关注

推荐阅读

### 强化学习进阶 第十讲 逆向强化学习（第三节 基于最大熵的方法）

强化学习进阶 第十讲 逆向强化学习（第三节 基于最大熵的方法）在第十讲的概述部分，我们已经介绍过，基于最大边际的方法往往会产生歧义，比如或许存在很多不同的回报函数导致相同的专家策...

天津包子馅儿

### 反向强化学习3——求解线性规划

本文收录在无痛的机器学习第二季目录中。线性规划求解过程在上一节我们介绍了反向强化学习的公式形式，接下来我们就要完成求解了。这个问题本质上是一个线性规划的问题，但是目前的公式形式...

冯超



### DeepMind深度学习高级课程，视频已全部放出

量子位

2

陈

还没有评论

写下你的评论...



▲ 赞同 22 ▼

● 添加评论

➦ 分享

★ 收藏

...

