



周志华 著. 机器学习, 北京: 清华大学出版社, 2016年1月.

(ISBN 978-7-302-206853-6) 425页, 62.6万字

[\[前言&目录\]](#) [\[样章\]](#)

[\[出版社网址\]](#) {[网购网址1](#)} {[网购网址2](#)} {[网购网址3](#)}

2016年6月起将为使用本书授课的教师提供PPT (免费; 不提供后续支持)

需要的老师请填写申请表 [\[暂未开通\]](#)

[\[读者敬启\]](#) [\[勘误修订\]](#)

读者敬启:

这只是一本关于机器学习的入门教科书。作者试图通过化繁为简的讲述,使读者能够相对轻松地大致了解机器学习这个学科领域最基本的情况,掌握若干基本概念,理解一些重要的思想。本书的目的,是作为读者初窥机器学习大厦的“初级地形图”。

作为一个学期课程的教材,本书的篇幅进行了仔细考量,16章正文,每章6-7节,一般20-25页。研究生课程若以18周一个学期,则除去习题和答疑时间,基本上每周讲授一章;本科生课程则可进度稍缓,一学期讲授9-10章。囿于此限,作者须对内容材料、以及材料讲述的程度进行取舍。为了尽可能避免读者只见树木、不见森林,作者对许多材料尽可能述其精要、去其细冗,以便让读者能在有限的时间空间中感受更多的、应该接触到的内容;否则若不分巨细,其篇幅可能令读者望而生畏。鉴于上述原因,读者不要指望通过这一本书就能成为机器学习专家。但本书已力图反映出机器学习大厦最基本的面貌,覆盖较为全面,读者若对某方面深感兴趣,不妨按照书中各章给出的文献指引进一步深造,或有助于精进。

初学机器学习易陷入的一个误区,是以为机器学习仅是若干种常用方法的堆积,于是把目光仅聚焦在具体算法的推导和编程实现上;待到实践中拿出所学,却往往发现效果不尽如人意。其实,现实世界问题千变万化,少有照搬书本之法就能奏效之时。一个问题,常可从多种视角抽象;同一种方法,亦常可从多种视角导衍。现实中更多时候,需依据问题条件对已知方法进行改造融通方能有效。须知,方法是“死”的,思想才是“活”的。欲行此道,则务须把握方法背后的思想脉络。无论创新科研还是应用实践,皆以此为登堂入室之始;然而初学者得之甚难。因此,本书在有限的篇幅中更侧重于这个方面,冀望辅助读者莫造进一步学习精进的视野心法,而对一些读者稍下工夫就易自行弄清的冗细则惜墨不赘。

本书重点针对机器学习初学者,但作者仍希望这是一本耐读之书。因此,作者在一些角落融入了多年研习实践的些微心得,虽仅只言片语,但可能不易得之,进阶读者阅之或可莞尔;初入门径者搁后再阅,于原不经意处或能有新得。

读者若仅对某几种具体机器学习方法的推导或实现感兴趣,对其背后的相关思想不甚关心,那么本书可能不太适合。更不要把本书当做机器学习方法的“速查手册”,若仅需那样的材料,不妨直接查看维基百科。

机器学习飞速发展,很多新进展在学界尚无公论之前,作者以为不适于写入入门教科书中。但为了不致与学科前沿脱节,本书也谈及了一些本领域专家有初步共识的相对较新的内容。

作者自认才疏学浅,对机器学习仅略知皮毛,更兼时间和精力所限,书中错谬之处难免,若蒙读者诸君[不吝告知](#),将不胜感激。

勘误修订

(第一版第一次印刷, 2016年1月):

- p. 6, 图1. 2: 图中两处"清脆" --> "浊响"
- p. 28, 第3段倒数第2行: "大量" --> "不少"
- p. 28, 边注: "例如 ……上百亿个参数" --> "机器学习常涉及两类参数: 一类是算法的参数, 亦称"超参数", 数目常在10以内; 另一类是模型的参数, 数目可能很多, 例如……上百亿个参数. 两者调参方式相似, 均是产生多个模型之后基于某种评估方法来进行选择; 不同之处在于前者通常是由人工设定多个参数候选值后产生模型, 后者则是通过学习来产生多个候选模型(例如神经网络在不同轮数停止训练)."
- p. 31, 倒数第3行: "Event" --> "Even"
- p. 256, 第4段: "固定住 $\{\mathbf{\alpha}\}_i$ " --> "以 $\{\mathbf{\alpha}\}_i$ 为初值"
- p. 256, 最后一段第1行: " $\{\mathbf{E}\}_i =$ " --> " $\{\mathbf{E}\}_i = \{\mathbf{X}\} -$ "
- p. 385, 式(16. 25)和(16. 26): 两处" $\mathbf{r}_i$ " --> " $\mathbf{R}_i$ "
- p. 385, 式(16. 25)下一行: "若改用……" --> "其中 $\mathbf{R}_i$ 表示第 $i$ 条轨迹上自状态 $\mathbf{x}$ 至结束的累积奖赏. 若改用……"
- p. 386, 式(16. 28)下一行: "始终为1" --> "对于 $\mathbf{a}_i = \pi(\mathbf{x}_i)$ 始终为1"
- p. 386, 图16. 11, 第4步: 两处 " $\pi(\mathbf{x})$ " --> " $\pi(\mathbf{x}_i)$ "
- p. 386, 图16. 11, 第6步的式子 --> "
$$\mathbf{R} = \frac{1}{T-t} \left( \sum_{i=t+1}^T \mathbf{r}_i \right) \prod_{i=t+1}^{T-1} \frac{\mathbb{I}(\mathbf{a}_i = \pi(\mathbf{x}_i))}{\pi(\mathbf{x}_i)}$$
"
- p. 386, 图16. 11, 边注"计算修正的累积奖赏." --> "计算修正的累积奖赏. 连乘内下标大于上标的项取值为1. "; 去掉边注"重要性采样系数."

(第一版第二次印刷, 2016年2月):

- p. 38, 第6行: " $\epsilon^{m'}$ " --> " $\mathbb{E}_{m \sim \text{choose } m'} \epsilon^m$ "
- p. 103, 最后一行的式子: 求和的" $l$ " --> " $q$ "
- p. 119, 第14行: "318--362" --> "533--536"
- p. 404, 式(B. 3)最后一行的式子 --> " $\lambda g(\mathbf{x}) = 0$ "

(第一版第三次印刷, 2016年3月):

- p. 15, 第5行: "居功" --> "厥功"
- p. 55, 最后一行: 式子括号中的逗号改为分号
- p. 125, 第3行: "减小" --> "增大"
- p. 125, 第4行, 第6行: "减幅" --> "增幅"
- p. 125, 第5行: "减小" --> "增长"

(第一版第四次印刷, 2016年3月):

- p. 59, 式(3. 27)加边注: "考虑  $\mathbf{y}_i \in \{0, 1\}$ "

---

[\[Home\]](#)