

Information Retrieval

信息检索

潘微科

教师信息

- 潘微科
- Office: 沧海校区致腾楼（计算机与软件学院）724室
- Email: panweike@szu.edu.cn
- Office hour: 周一下午14:00-16:00
- 课程QQ群: 261821110
- 助教: 黄炳森
- 更多信息: <http://csse.szu.edu.cn/staff/panwk/>



群名称: 202102-信息检索
群 号: 261821110

课程信息(1/2)

- 基本信息：
 - 课程号：1502940001
 - 课程名称：信息检索
 - 课程类别：学科专业选修课
 - 学分：2.5学分（理科学分）
 - 学时：54学时（36课堂+18实验）
 - 主选班级（上限30人）：2019信息与计算科学02; 2019软件工程01,02; 2019计算机科学与技术01,02,03,04,06
 - 非主选班级（上限20人）
 - 理论课：每周五7-8节, 致理楼L3-418教室
 - 实验课：双周五9-10节, 致腾楼240实验室

课程信息(2/2)

- 先修课程: Java程序设计/面向对象程序设计, 数据结构和算法
- 开发环境: JDK, Eclipse等常用集成开发环境

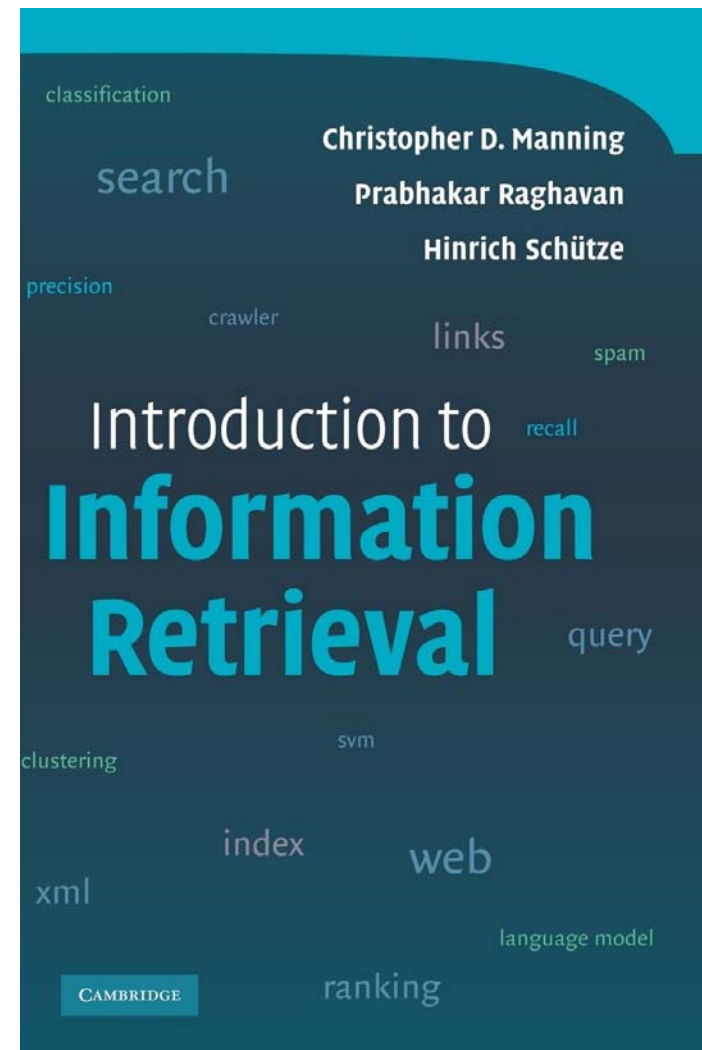
教材

- Christopher D. Manning, Prabhakar Raghavan and Hinrich Schütze. **Introduction to Information Retrieval**. Cambridge University Press, 2008. ISBN:0521865719.
- 注: 作者分享了 **电子版教材** 等资源 <https://nlp.stanford.edu/IR-book/>

← → ↻ nlp.stanford.edu/IR-book/

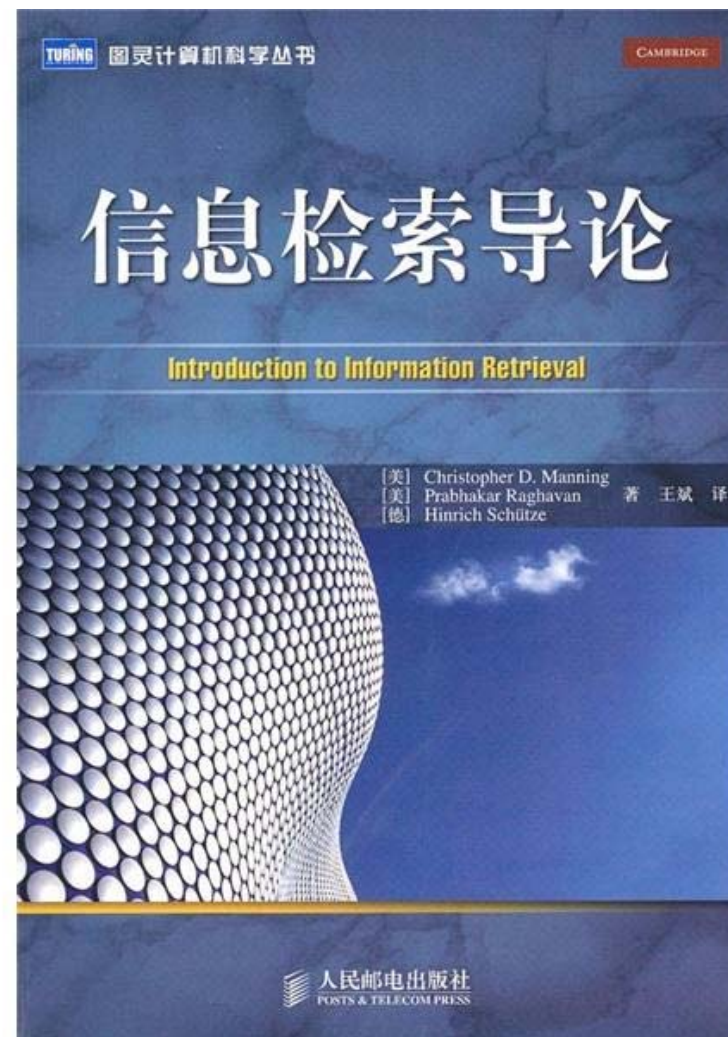
Introduction to Information Retrieval: Table of Contents

	chapter
Front matter (incl. table of notations)	pdf
01 Boolean retrieval	pdf html
02 The term vocabulary & postings lists	pdf html
03 Dictionaries and tolerant retrieval	pdf html
04 Index construction	pdf html
05 Index compression	pdf html
06 Scoring, term weighting & the vector space model	pdf html
07 Computing scores in a complete search system	pdf html
08 Evaluation in information retrieval	pdf html
09 Relevance feedback & query expansion	pdf html
10 XML retrieval	pdf html
11 Probabilistic information retrieval	pdf html
12 Language models for information retrieval	pdf html
13 Text classification & Naive Bayes	pdf html
14 Vector space classification	pdf html
15 Support vector machines & machine learning on documents	pdf html
16 Flat clustering	pdf html
17 Hierarchical clustering	pdf html
18 Matrix decompositions & latent semantic indexing	pdf html
19 Web search basics	pdf html
20 Web crawling and indexes	pdf html
21 Link analysis	pdf html
Bibliography & Index	pdf
bibtex file	bib



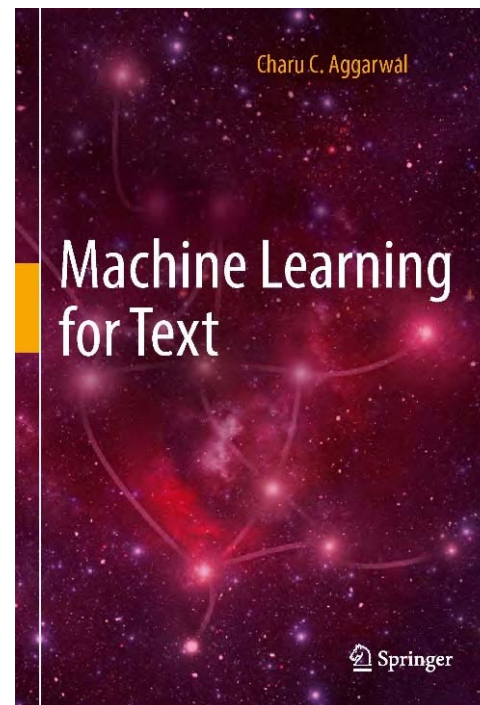
参考教材(1/2)

- 王斌(译). 信息检索导论. 人民邮电出版社. 2010.9.
ISBN:9787115234247.
- 注: 是英文教材的中译本.



参考教材(2/2)

- Charu C. Aggarwal. **Machine Learning for Text**. Springer, 2018. ISBN: 9783319735306.
- 注:
 - Chapter 9: **Information Retrieval and Search Engines**
 - 黎琳,潘微科,明仲 (译). **文本机器学习**. 机械工业出版社. 2020.5. ISBN:9787111648055.

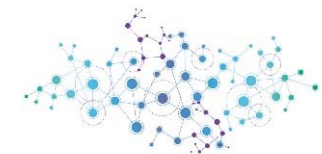


 Springer

文本机器学习

Machine Learning for Text

[美] 查鲁·C.阿加瓦尔 (Charu C. Aggarwal) 著
黎琳 潘微科 明仲 译



 机械工业出版社
CHINA MACHINE PRESS

成绩评定(1/3)

- 课堂参与： 5%
 - 理论课参与(互动)： 3%
 - 实验课参与(演示)： 2%
- 课堂测试： 5%
- 实验作业： 60% ($6 \times 10\% = 60\%$)
- 期末考试(一般为“期末大作业”)： 30%

成绩评定(2/3)

- 实验作业：
 - 实验1: 布尔检索和倒排索引的实验（教材第1章）
 - 实验2: 词典、倒排记录表和容错式检索的实验(教材第2-3章)
 - 实验3: 索引构建和压缩的实验（教材第4-5章）
 - 实验4: 文档评分和概率检索模型的实验（教材第6-7,11-12章）
 - 实验5: 分类和聚类的实验（教材第13,18章）
 - 实验6: 链接分析的实验（教材第21章）

成绩评定(3/3)

- 期末大作业
 - 201602学期：面向校园新闻的信息检索系统开发
 - 201702学期：面向DBLP学术论文的信息检索系统开发
 - 201802学期：面向学院网站的信息检索系统
 - 201902学期：面向法律文本的信息检索系统
 - 202002学期：面向垂直领域的多语言文本信息检索系统

FAQ-0

Q: 潘微科老师的联系方式有哪些?

A: 深大邮箱: panweike@szu.edu.cn; QQ: 1926768129; 微信/手机: 13632865400

- 注: 特殊情况补交作业, 请发到深大邮箱 (不要发QQ邮箱)。
紧急情况请拨打手机。

FAQ-1

Q: 潘微科老师的《信息检索》课程使用哪个线上教学平台？

A: 腾讯课堂

<https://ke.qq.com/webcourse/4359319/104523939#from=800021724&lite=1&live=1>



FAQ-2

Q: 潘微科老师的《信息检索》课程何时上课?

A: 理论课: 每周五7-8节; 实验课: 双周五9-10节

第7节 14:15-14:55

第8节 15:00-15:40

大课间 20分钟

第9节 16:00-16:40

第10节 16:45-17:25

FAQ-3

Q: 潘微科老师的《信息检索》课程有没有[课程QQ群](#)或课程微信群？

A: 课程QQ群: 261821110

课程QQ群二维码:



群名称: 202102-信息检索
群 号: 261821110

FAQ-4

Q: 在潘微科老师的《信息检索》课程中，QQ群的主要用途是什么？

A: 发布实时通知（重要通知以Blackboard为准）

上传课件、电子版教材等资料

实验课/平时问答等

FAQ-5

Q: 潘微科老师的《信息检索》课程使用什么智慧学习支撑平台/系统?

A: Blackboard (简称BB) : <https://elearning.szu.edu.cn/>

FAQ-6

Q: 在潘微科老师的《信息检索》课程中，BB的主要用途是什么？

A: 发布重要通知

上传课件、电子版教材等资料

布置&批改作业