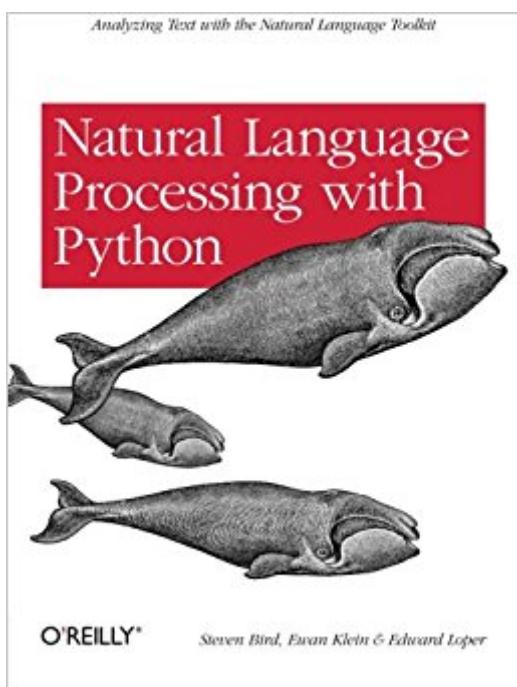


The book was found

Natural Language Processing With Python: Analyzing Text With The Natural Language Toolkit



Synopsis

This book offers a highly accessible introduction to natural language processing, the field that supports a variety of language technologies, from predictive text and email filtering to automatic summarization and translation. With it, you'll learn how to write Python programs that work with large collections of unstructured text. You'll access richly annotated datasets using a comprehensive range of linguistic data structures, and you'll understand the main algorithms for analyzing the content and structure of written communication. Packed with examples and exercises, Natural Language Processing with Python will help you:

- Extract information from unstructured text, either to guess the topic or identify "named entities"
- Analyze linguistic structure in text, including parsing and semantic analysis
- Access popular linguistic databases, including WordNet and treebanks
- Integrate techniques drawn from fields as diverse as linguistics and artificial intelligence

This book will help you gain practical skills in natural language processing using the Python programming language and the Natural Language Toolkit (NLTK) open source library. If you're interested in developing web applications, analyzing multilingual news sources, or documenting endangered languages -- or if you're simply curious to have a programmer's perspective on how human language works -- you'll find Natural Language Processing with Python both fascinating and immensely useful.

Book Information

Paperback: 504 pages

Publisher: O'Reilly Media; 1 edition (July 10, 2009)

Language: English

ISBN-10: 0596516495

ISBN-13: 978-0596516499

Product Dimensions: 7 x 1.2 x 9.2 inches

Shipping Weight: 1.5 pounds (View shipping rates and policies)

Average Customer Review: 4.1 out of 5 stars 46 customer reviews

Best Sellers Rank: #39,658 in Books (See Top 100 in Books) #7 in Books > Computers & Technology > Computer Science > AI & Machine Learning > Natural Language Processing #17 in Books > Computers & Technology > Programming > Web Programming > JavaScript #26 in Books > Computers & Technology > Programming > Algorithms

Customer Reviews

Analyzing Text with the Natural Language Toolkit --This text refers to an out of print or unavailable

edition of this title.

Steven Bird is Associate Professor in the Department of Computer Science and Software Engineering at the University of Melbourne, and Senior Research Associate in the Linguistic Data Consortium at the University of Pennsylvania. He completed a PhD on computational phonology at the University of Edinburgh in 1990, supervised by Ewan Klein. He later moved to Cameroon to conduct linguistic fieldwork on the Grassfields Bantu languages under the auspices of the Summer Institute of Linguistics. More recently, he spent several years as Associate Director of the Linguistic Data Consortium where he led an R&D team to create models and tools for large databases of annotated text. At Melbourne University, he established a language technology research group and has taught at all levels of the undergraduate computer science curriculum. In 2009, Steven is President of the Association for Computational Linguistics. Ewan Klein is Professor of Language Technology in the School of Informatics at the University of Edinburgh. He completed a PhD on formal semantics at the University of Cambridge in 1978. After some years working at the Universities of Sussex and Newcastle upon Tyne, Ewan took up a teaching position at Edinburgh. He was involved in the establishment of Edinburgh's Language Technology Group in 1993, and has been closely associated with it ever since. From 2000-2002, he took leave from the University to act as Research Manager for the Edinburgh-based Natural Language Research Group of Edify Corporation, Santa Clara, and was responsible for spoken dialogue processing. Ewan is a past President of the European Chapter of the Association for Computational Linguistics and was a founding member and Coordinator of the European Network of Excellence in Human Language Technologies (ELSNET). Edward Loper has recently completed a PhD on machine learning for natural language processing at the the University of Pennsylvania. Edward was a student in Steven's graduate course on computational linguistics in the fall of 2000, and went on to be a TA and share in the development of NLTK. In addition to NLTK, he has helped develop two packages for documenting and testing Python software, epydoc, and doctest.

Edward Loper's book is an introduction to the Natural Language Toolkit (NLTK) for the Python programming language. Its target audience is a narrow one. It assumes a working familiarity with Python. It's true that an experienced programmer could learn Python along the way, but getting the most from the code examples and walkthrough explanations requires enough familiarity to "think" in Python. The book also assumes sufficient familiarity with Natural Language Processing (NLP) to understand why one would want to build lexicons, grammars, and parsers. The book has several

strengths. It is tightly integrated with Python and NLTK code. There are numerous examples throughout and the author walks through and modifies them to clarify how the NLTK works. The sizeable reference sections at the end of each chapter are also valuable. These sections include both introductory and advanced sources. And a lot of them. There is also useful integration with the NLTK web site which provides and points to additional resources. Not to be missed are the end-of-chapter questions. Readers have come to expect little from these learning aids; they usually invite us to parrot back a small number of key concepts or try a few calculations or code segments. This book's questions go far beyond the norm. They introduce new concepts, encourage writing and comparing several versions of a program, and otherwise extend each chapter's contents. Even readers who don't plan to complete these exercises should read them closely. Weaknesses are few. As noted, the book may assume too much Python and NLP background for some users. It does have a narrow focus and is not organized the right way to be used as a reference book. Readers who want something a little more modular and reference-like might prefer Jacob Perkins' *Python 3 Text Processing with NLTK 3 Cookbook*. David Mertz's *Text Processing in Python* is an older source, but still useful as well.

Buy this book only if you: 1. Know the basics of natural language processing (NLP) or linguistics; 2. Know the Python programming language or you're willing to learn it; 3. Are using the NLTK library or plan to do so. NLTK is a Python library that offers many standard NLP tools (tokenizers, POS taggers, parsers, chunkers and others). It comes with samples of several dozens of text corpora typically used in NLP applications, as well as with interfaces to dictionary-like resources such as WordNet and VerbNet. No FrameNet, though. NLTK is well documented, so you might not need this book initially. However, it definitely helps to have it on your desk if you are serious about using NLTK. The first chapters are a bit messy, as they attempt to introduce all three themes (NLP, NLTK and Python) together. Beginners may have some difficulty sorting things out. By the time you reach the WordNet section, you either got lost in the forest, realize that you would never understand this topic without the book, or both. However, if you are a bit patient and try out all simple code examples, you'll make it eventually. In my opinion, NLTK remains the simplest, most elegant and well rounded library of its kind.

NLTK was originally designed for teaching NLP, but because NLTK is so comprehensive, it is also quite vast, and you will need a guide to get you started on using it effectively. This book is that guide. Its organized into 4 sections. Chapters 1-4 cover the basics; 5-7 covers language processing,

tagging, classification and information extraction; Chapters 8-10 covers sentence parsing, syntax, structure and representations of meaning, and Chapter 11 covers managing linguistics data. If you are looking to get an overview of NLP, as well as reasonable proficiency in manipulating text and extracting information from it, this book may be for you. People who are not NLP specialists but need to use NLP techniques at work will find this book particularly helpful. People who plan on specializing in NLP will probably find the book useful as a stepping stone into the field. My interest in NLP (and the book) is limited to being able to apply machine learning techniques to solve NLP problems, so I found the first two sections really useful. However, the entire book (including the exercises) is a great source of ideas on what you can accomplish in NLP with NLTK.

This is one of the most interesting programming books I have ever read, and it's so easy to jump right in and play with the NLTK. I have devoured this book. A lot of the code examples in the book no longer work as written, as the library has changed over the last eight years, but it is easy enough to Google the new way, and it is educational trying to figure out how to get things working that aren't well documented. Great book.

My subject area background when I picked up this book was NLP 101, Python 201. Halfway through the book, I can say I am at NLP 201, Python 301. I really like the example-driven approach. It motivates me to write Python modules operating in a similar fashion to the examples but with heuristic changes (wherever applicable). It is a good journey of discovery and reinforcement of NLP concepts via a DIY approach. I do find that relatively complex concepts sneak up on you as you get deeper in the book. But you can always google for literature on areas, which do not have exhaustive coverage. And there is a decent amount of literature easily and/or freely available. Tip - it really helps to know your probability and statistics theory. To summarize, this book has sustained and grown my interest in NLP - an area with amazing opportunities and possibilities. And Python is a really elegant language. Thank you Steven Bird and team!

Although this text is available for free online through NLTK, it is an incredible resource for anybody trying to get started with NLP in Python. With only basic knowledge of Python and a week with this book, I was able to write a work application to identify key themes in survey data utilizing part of speech tagging and a custom built Regex parser. I highly recommend this text.

[Download to continue reading...](#)

Python: Programming: Your Step By Step Guide To Easily Learn Python in 7 Days (Python for

Beginners, Python Programming for Beginners, Learn Python, Python Language) Natural Language Processing with Python: Analyzing Text with the Natural Language Toolkit Python: The Complete Python Quickstart Guide (For Beginner's) (Python, Python Programming, Python for Dummies, Python for Beginners) Hacking with Python: Beginner's Guide to Ethical Hacking, Basic Security, Penetration Testing, and Python Hacking (Python Programming, Hacking, Python Coding, Python and Hacking Book 3) PYTHON: PYTHON'S COMPANION, A STEP BY STEP GUIDE FOR BEGINNERS TO START CODING TODAY! (INCLUDES A 6 PAGE PRINTABLE CHEAT SHEET)(PYTHON FOR BEGINNERS, PYTHON FOR DUMMIES, PYTHON PROGRAMMING) Python Programming: Python Programming for Beginners, Python Programming for Intermediates, Python Programming for Advanced PYTHON: LEARN PYTHON in A Day and MASTER IT WELL. The Only Essential Book You Need To Start Programming in Python Now. Hands On Challenges INCLUDED! (Programming for Beginners, Python) Python: Learn Python in a Day and Master It Well: The Only Essential Book You Need to Start Programming in Python Now Maya Python for Games and Film: A Complete Reference for Maya Python and the Maya Python API Python Programming: An In-Depth Guide Into The Essentials Of Python Programming (Included: 30+ Exercises To Master Python in No Time!) Python: The Fundamentals Of Python Programming: A Complete Beginners Guide To Python Mastery. CLEP Analyzing & Interpreting Literature with CD-ROM (REA): The Best Test Prep for the CLEP Analyzing and Interpreting Literature Exam with REA's TESTware (Test Preps) The Production Manager's Toolkit: Successful Production Management in Theatre and Performing Arts (The Focal Press Toolkit Series) The Technical Director's Toolkit: Process, Forms, and Philosophies for Successful Technical Direction (The Focal Press Toolkit Series) The Assistant Lighting Designer's Toolkit (The Focal Press Toolkit Series) The Don't Get Me Started! Toolkit - Workbook and Teacher Answer Key: Strategies for a Culturally-Challenged World (The Don't Get Me Started! Toolkit - Workbook and Teacher Key) (Volume 1) Python Programming Advanced: A Complete Guide on Python Programming for Advanced Users Python: Python Programming for Intermediates How to Code 2.0: Pushing Your Skills Further with Python: Learn how to code with Python and Pygame in 10 Easy Lessons (Super Skills) Python Programming Guide + SQL Guide - Learn to be an EXPERT in a DAY!: Box Set Guide (Python Programming, SQL)

[Contact Us](#)

[DMCA](#)

[Privacy](#)

FAQ & Help