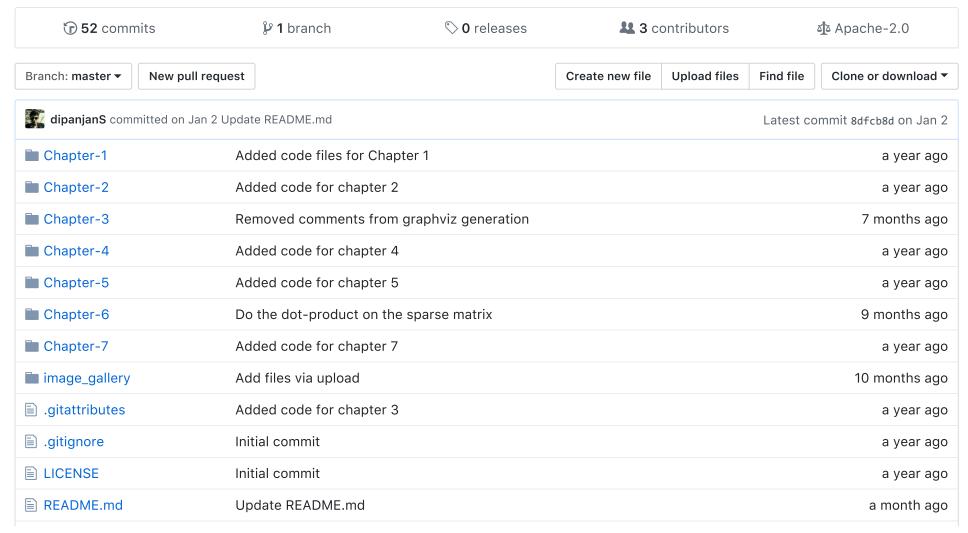# 📖 dipanjanS / **text-analytics-with-python**

Learn how to process, classify, cluster, summarize, understand syntax, semantics and sentiment of text data with the power of Python! This repository contains code and datasets used in my book, "Text Analytics with Python" published by Apress/Springer.

#text-analytics #text-summarization #text-classification #python #natural-language #natural-language-processing #clustering #sentiment #semantic #sentiment-analysis #nltk #stanford-nlp #spacy #pattern #scikit-learn #gensim

⏲ **52** commits	⑂ **1** branch	🏷 **0** releases	👥 **3** contributors	⚖ Apache-2.0

Branch: master ▼	New pull request		Create new file	Upload files	Find file	Clone or download ▼

dipanjanS committed on Jan 2 Update README.md   Latest commit `8dfcb8d` on Jan 2

📁 Chapter-1	Added code files for Chapter 1	a year ago
📁 Chapter-2	Added code for chapter 2	a year ago
📁 Chapter-3	Removed comments from graphviz generation	7 months ago
📁 Chapter-4	Added code for chapter 4	a year ago
📁 Chapter-5	Added code for chapter 5	a year ago
📁 Chapter-6	Do the dot-product on the sparse matrix	9 months ago
📁 Chapter-7	Added code for chapter 7	a year ago
📁 image_gallery	Add files via upload	10 months ago
📄 .gitattributes	Added code for chapter 3	a year ago
📄 .gitignore	Initial commit	a year ago
📄 LICENSE	Initial commit	a year ago
📄 README.md	Update README.md	a month ago

📄 _config.yml	Set theme jekyll-theme-merlot	a year ago

📖 **README.md**

Text Analytics with Python

A Practical Real-World Approach to Gaining Actionable Insights from your Data

Text analytics can be a bit overwhelming and frustrating at times with the unstructured and noisy nature of textual data and the vast amount of information available. "Text Analytics with Python" is a book packed with 385 pages of useful information based on techniques, algorithms, experiences and various lessons learnt over time in analyzing text data. This repository contains datasets and code used in this book. I will also be adding various notebooks and bonus content here from time to time. Keep watching this space!

Work is underway for a revised edition of this book including all code in python 3 with several new examples too!
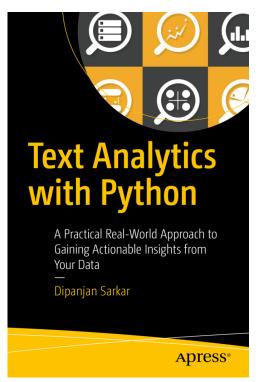
Get the book

About the book

Derive useful insights from your data using Python. Learn the techniques related to natural language processing and text analytics, and gain the skills to know which technique is best suited to solve a particular problem.

Text Analytics with Python teaches you both basic and advanced concepts, including text and language syntax, structure, semantics. You will focus on algorithms and techniques, such as text classification, clustering, topic modeling, and text summarization

A structured and comprehensive approach is followed in this book so that readers with little or no experience do not find themselves overwhelmed. You will start with the basics of natural language and Python and move on to advanced analytical and machine learning concepts. You will look at each technique and algorithm with both a bird's eye view to understand how it can be used as well as with a microscopic view to understand the mathematical concepts and to implement them to solve your own problems.

Edition: 1st   Pages: 385   Language: English

Book Title: Text Analytics with Python   Publisher: Apress (a part of Springer)   Copyright: Dipanjan Sarkar

Print ISBN: 978-1-4842-2387-1   Online ISBN: 978-1-4842-2388-8   DOI: 10.1007/978-1-4842-2388-8

This book:

- Provides complete coverage of the major concepts and techniques of natural language processing (NLP) and text analytics

- Includes practical real-world examples of techniques for implementation, such as building a text classification system to categorize news articles, analyzing app or game reviews using topic modeling and text summarization, and clustering popular movie synopses and analyzing the sentiment of movie reviews

- Shows implementations based on Python and several popular open source libraries in NLP and text analytics, such as the natural language toolkit (`nltk`), `gensim` , `scikit-learn` , `spaCy` and `pattern`

Contents

- Chapter 1: Natural Language Basics
- Chapter 2: Python Refresher
- Chapter 3: Processing and Understanding Text
- Chapter 4: Text Classification
- Chapter 5: Text Summarization
- Chapter 6: Text Similarity and Clustering
- Chapter 7: Semantic and Sentiment Analysis