Natural Language Processing With Java And LingPipe Cookbook
**Synopsis**

Over 60 effective recipes to develop your Natural Language Processing (NLP) skills quickly and effectively

About This Book

Build effective natural language processing applications

Transit from ad-hoc methods to advanced machine learning techniques

Use advanced techniques such as logistic regression, conditional random fields, and latent Dirichlet allocation

Who This Book Is For

This book is for experienced Java developers with NLP needs, whether academics, industrialists, or hobbyists. A basic knowledge of NLP terminology will be beneficial.

What You Will Learn

Master a broad range of classification techniques for text data

Track people, concepts, and things in data, within and across documents

Understand the importance of evaluation in creation of NLP applications and how to do it

Yield best practices for common text-analytics problems

Tune systems for high performance and trade off various aspects of the performance curve

Become a master in customizing NLP systems at all levels

Build systems for non-tokenized languages such as Chinese and Japanese

In Detail

NLP is at the core of web search, intelligent personal assistants, marketing, and much more, and LingPipe is a toolkit for processing text using computational linguistics.

This book starts with the foundational but powerful techniques of language identification, sentiment classifiers, and evaluation frameworks. It goes on to detail how to build a robust framework to solve common NLP problems, before ending with advanced techniques for complex heterogeneous NLP systems.

This is a recipe and tutorial book for experienced Java developers with NLP needs. A basic knowledge of NLP terminology will be beneficial. This book will guide you through the process of how to build NLP apps with minimal fuss and maximal impact.

**Book Information**

File Size: 3729 KB

Print Length: 312 pages

Publisher: Packt Publishing (November 28, 2014)

Publication Date: November 28, 2014

Sold by: Digital Services LLC

Language: English

ASIN: B00QFBNXGQ

Text-to-Speech: Enabled

X-Ray: Not Enabled

Word Wise: Not Enabled

Lending: Not Enabled
LingPipe is a Natural Language Processing (NLP) library that is released under a dual commercial and an open-source AGPL license, and the basis for a NLP consulting company (Alias-I) that one of the authors (Breck Baldwin) founded. In fact, the preface of the book states that some of the recipes in this book come from Breck’s private repository. This book is the first one devoted exclusively to LingPipe. While LingPipe provides comprehensive Javadocs and tutorials on its website, but it is fairly dense material (NLP is hard!) - the book is an easier, gentler way to understand it. One other reason LingPipe’s API is so dense (even compared to other Java NLP libraries) is because it is written for performance, making heavy use of encapsulation to wrap common tasks and the visitor pattern to consume data in streaming mode. The book does a good job explaining the latter pattern in some depth, and deconstructing the code examples so the former becomes more obvious. The book also covers class hierarchies of common NLP components (and their functionality) such as tokenizers and classifiers, which also helps in understanding the API. LingPipe is a mature library and has components to do just about anything - notably tokenizers, classifiers (Naive Bayes, MaxEnt), Chunkers (HMM, CRF, rule based), clusterers (Single Link, Complete Link, LDA), etc. The book covers the use of these components through novel recipes, through which the authors share their NLP insights and experience with the reader. If you are serious about using LingPipe to solve your NLP problems and you don't have access to their consulting services (or even if you do), this book can help you get there.

*Download to continue reading...*