# Text Processing In Java

*Create your own natural language training corpus for machine learning. Whether you're working with English, Chinese, or any other natural language, this hands-on book guides you through a proven annotation development cycle—the process of adding metadata to your training corpus to help ML algorithms work more efficiently. You don't need any programming or linguistics experience to get started. Using detailed examples at every step, you'll learn how the MATTER Annotation Development Process helps you Model, Annotate, Train, Test, Evaluate, and Revise your training corpus. You also get a complete walkthrough of a real-world annotation project. Define a clear annotation goal before collecting your dataset (corpus) Learn tools for analyzing the linguistic content of your corpus Build a model and specification for your annotation project Examine the different annotation formats, from basic XML to the Linguistic Annotation Framework Create a gold standard corpus that can be used to train and test ML algorithms Select the ML algorithms that will process your annotated data Evaluate the test results and revise your annotation task Learn how to use lightweight software for annotating texts and adjudicating the annotations This book is a perfect companion to O'Reilly's Natural Language Processing with Python.*

*Design, build, and deploy your own machine learning applications by leveraging key Java machine learning librariesAbout This Book- Develop a sound strategy to solve predictive modelling problems using the most popular machine learning Java libraries- Explore a broad variety of data processing, machine learning, and natural language processing through diagrams, source code, and real-world applications- Packed with practical advice and tips to help you get to grips with applied machine learningWho This Book Is ForIf you want to learn how to use Java's machine learning libraries to gain insight from your data, this book is for you. It will get you up and running quickly and provide you with the skills you need to successfully create, customize, and deploy machine learning applications in real life. You should be familiar with Java programming and data mining concepts to make the most of this book, but no prior experience with data mining packages is necessary.What You Will Learn- Understand the basic steps of applied machine learning and how to differentiate among various machine learning approaches- Discover key Java machine learning libraries, what each library brings to the table, and what kind of problems each are able to solve- Learn how to implement classification, regression, and clustering- Develop a sustainable strategy for customer retention by predicting likely churn candidates- Build a scalable recommendation engine with Apache Mahout- Apply machine learning to fraud, anomaly, and outlier detection- Experiment with deep learning concepts, algorithms, and the toolbox for deep learning- Write your own activity recognition model for eHealth applications using mobile sensorsIn DetailAs the amount of data continues to grow at an almost incomprehensible rate, being able to understand and process data is becoming a key differentiator for competitive organizations. Machine learning applications are everywhere, from self-driving cars, spam detection, document search, and trading strategies, to speech recognition. This makes machine learning well-suited to the present-day era of Big Data and Data Science. The main challenge is how to transform data into actionable knowledge.Machine Learning in Java will provide you with the techniques and tools you need to quickly gain insight from complex data. You will start by learning how to apply machine learning methods to a variety of common tasks including classification, prediction, forecasting, market basket analysis, and clustering.Moving on, you will discover how to detect anomalies and fraud, and ways to perform activity recognition, image recognition, and text analysis. By the end of the book, you will explore related web resources and technologies that will help you take your learning to the next level.By applying the most effective machine learning methods to real-world problems, you will gain hands-on experience that will transform the way you think about data.Style and approachThis is a practical tutorial that uses hands-on examples to step through some real-world applications of machine learning. Without shying away from the technical details, you will explore machine learning with Java libraries using clear and practical examples. You will explore how to prepare data for analysis, choose a machine learning method, and measure the success of the process.*

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

*Explore various approaches to organize and extract useful text from unstructured data using Java Key Features Use deep learning and NLP techniques in Java to discover hidden insights in text Work with popular Java libraries such as CoreNLP, OpenNLP, and Mallet Explore machine translation, identifying parts of speech, and topic modeling Book Description Natural Language Processing (NLP) allows you to take any sentence and identify patterns, special names, company names, and more. The second edition of Natural Language Processing with Java teaches you how to perform language analysis with the help of Java libraries, while constantly gaining insights from the outcomes. You'll start by understanding how NLP and its various concepts work. Having got to grips with the basics, you'll explore important tools and libraries in Java for NLP, such as CoreNLP, OpenNLP, Neuroph, and Mallet. You'll then start performing NLP on different inputs and tasks, such as tokenization, model training, parts-of-speech and parsing trees. You'll learn about statistical machine translation, summarization, dialog systems, complex searches, supervised and unsupervised NLP, and more. By the end of this book, you'll have learned more about NLP, neural networks, and various other trained models in Java for enhancing the performance of NLP applications. What you will learn Understand basic NLP tasks and how they relate to one another Discover and use the available tokenization engines Apply search techniques to find people, as well as things, within a document Construct solutions to identify parts of speech within sentences Use parsers to extract relationships between elements of a document Identify topics in a set of documents Explore topic modeling from a document Who this book is for Natural Language Processing with Java is for you if you are a data analyst, data scientist, or machine learning engineer who wants to extract information from a language using Java. Knowledge of Java programming is needed, while a basic understanding of statistics will be useful but not mandatory.*

*Natural Language Processing and Text Mining not only discusses applications of Natural Language Processing techniques to certain Text Mining tasks, but also the converse, the use of Text Mining to assist NLP. It assembles a diverse views from internationally recognized researchers and emphasizes caveats in the attempt to apply Natural Language Processing to text mining. This state-of-the-art survey is a must-have for advanced students, professionals, and researchers.*

*Much of the data available today is unstructured and text-heavy, making it challenging for analysts to apply their usual data wrangling and visualization tools. With this practical book, you'll explore text-mining techniques with tidytext, a package that authors Julia Silge and David Robinson developed using the tidy principles behind R packages like ggraph and dplyr. You'll learn how tidytext and other tidy tools in R can make text analysis easier and more effective. The authors demonstrate how treating text as data frames enables you to manipulate, summarize, and visualize characteristics of text. You'll also learn how to integrate natural language processing (NLP) into effective workflows. Practical code examples and data explorations will help you generate real insights from literature, news, and social media. Learn how to apply the tidy text format to NLP Use sentiment analysis to mine the emotional content of text Identify a document's most important terms with frequency measurements Explore relationships and connections between words with the ggraph and widyr packages Convert back and forth between R's tidy and non-tidy text formats Use topic modeling to classify document collections into natural groups Examine case studies that compare Twitter archives, dig into NASA metadata, and analyze thousands of Usenet messages*

*This revised and expanded new edition of an internationally successful classic presents an accessible introduction to the key methods in digital image processing for both practitioners and teachers. Emphasis is placed on practical application, presenting precise algorithmic descriptions in an unusually high level of detail, while highlighting direct connections between the mathematical foundations and concrete implementation. The text is supported by practical examples and carefully constructed chapter-ending exercises drawn from the authors' years of teaching experience, including easily adaptable Java code and completely worked out examples. Source code, test images and additional instructor materials are also provided at an associated website. Digital Image Processing is the definitive textbook for students, researchers, and professionals in search of critical analysis and modern implementations of the most important algorithms in the field, and is also eminently suitable for self-study.*

*Our world is being revolutionized by data-driven methods: access to large amounts of data has generated new insights and opened exciting new opportunities in commerce, science, and computing applications. Processing the enormous quantities of data necessary for these advances requires large clusters, making distributed computing paradigms more crucial than ever. MapReduce is a programming model for expressing distributed computations on massive datasets and an execution framework for large-scale data processing on clusters of commodity servers. The programming model provides an easy-to-understand abstraction for designing scalable algorithms, while the execution framework transparently handles many system-level details, ranging from scheduling to synchronization to fault tolerance. This book focuses on MapReduce algorithm design, with an emphasis on text processing algorithms common in natural language processing, information retrieval, and machine learning. We introduce the notion of MapReduce design patterns, which represent general reusable solutions to commonly occurring problems across a variety of problem domains. This book not only intends to help the reader "think in MapReduce", but also discusses limitations of the programming model as well. This volume is a printed version of a work that appears in the Synthesis Digital Library of Engineering and Computer Science. Synthesis Lectures provide concise, original presentations of important research and development topics, published quickly, in digital and print formats. For more information visit www.morganclaypool.com*

*Natural Language Processing with Java and LingPipe Cookbook*
*Train neural networks for classification, NLP, and reinforcement learning using Deeplearning4j*
*Mastering Text Mining with R*
*Digital Image Processing*
*Natural Language Processing and Text Mining*
*Data-intensive Text Processing with MapReduce*
*Text Processing in Python*
*Develop Deep Learning Models for your Natural Language Problems*
*Processing*
*Text Analytics with Python*
*Over 70 recipes to create linguistic and language translation applications using Java libraries*
*How to Think Like a Computer Scientist*

*The java New I/O (NIO) packages in J2SE 1.4 introduce many new, indispensable features previously unavailable to Java programmers. These include APIs for high-performance I/O operations, regular expression processing, and character set coding. These new libraries are a treasure trove for java developers. The NIO APIs are especially valuable where high-performance I/O is a requirement, but they can also be useful in a wide range of scenarios. The new APIs let you work directly with I/O buffers, multiplex nonblocking streams, do scattering reads and gathering writes, do channel-to-channel transfers, work with memory-mapped files, manage file locks, and much more. The new high-performance Regular Expression Library provides sophisticated, Perl-like regex-processing features such as pattern matching, search and replace, capture groups, look ahead assertions, and many others. The Charset API gives you complete control over character set encoding and decoding, which are vital for properly managing the exchange of documents on the Web, for localization, or for other purposes. You can also create and install your own custom character sets. Staying current with the latent java technology is never easy. NIO, new in Java 1.4, is quite possibly the most important new java feature since Swing. Understanding it thoroughly is essential for any serious Java developer. NIO closes the gap between java and natively compiled languages and enables java applications to achieve maximum I/O performance by effectively leveraging operating-system services in a portable way. Java NIO is a comprehensive guide to the java New I/O facilities. It lots you take full advantage of NIO features and shows you how they work, what they can do for you, and when you should use them. This book brings you up to speed on NIO and shows you how to bring your I/O-bound Java applications up to speed as well. Java NIO is an essential part of any Java professional's library.*

*Learning Processing, Second Edition, is a friendly start-up guide to Processing, a free, open-source alternative to expensive software and daunting programming languages. Requiring no previous experience, this book is for the true programming beginner. It teaches the basic building blocks of programming needed to create cutting-edge graphics applications including interactive art, live video processing, and data visualization. Step-by-step examples, thorough explanations, hands-on exercises, and sample code, supports your learning curve. A unique lab-style manual, the book gives graphic and web designers, artists, and illustrators of all stripes a jumpstart on working with the Processing programming environment by providing instruction on the basic principles of the language, followed by careful explanations of select advanced techniques. The book has been developed with a supportive learning experience at its core. From algorithms and data mining to rendering and debugging, it teaches object-oriented programming from the ground up within the fascinating context of interactive visual media. This book is ideal for graphic designers and visual artists without programming background who want to learn programming. It will also appeal to students taking college and graduate courses in interactive media or visual computing, and for self-study. A friendly start-up guide to Processing, a free, open-source alternative to expensive software and daunting programming languages No previous experience required—this book is for the true programming beginner! Step-by-step examples, thorough explanations, hands-on exercises, and sample code supports your learning curve*

*Includes the proceedings of the 11th International Conference on Computational Linguistics and Intelligent Text Processing, held in IaAui, Romania, in March 2010. This title covers topics such as: lexical resources, syntax and parsing, word sense disambiguation and named entity recognition, semantics and dialog, and humor and emotions.*

*Specialised linguistic research needs can no longer be met by available software. This book enables the researcher to write programs for text and corpus processing, using the popular and easy to learn Java language.*

*Examine the techniques and Java tools supporting the growing field of data science About This Book Your entry ticket to the world of data science with the stability and power of Java Explore, analyse, and visualize your data effectively using easy-to-follow examples Make your Java applications more capable using machine learning Who This Book Is For This book is for Java developers who are comfortable developing applications in Java. Those who now want to enter the world of data science or wish to build intelligent applications will find this book ideal. Aspiring data scientists will also find this book very helpful. What You Will Learn Understand the nature and key concepts used in the field of data science Grasp how data is collected, cleaned, and processed Become comfortable with key data analysis techniques See specialized analysis techniques centered on machine learning Master the effective visualization of your data Work with the Java APIs and techniques used to perform data analysis In Detail Data science is concerned with extracting knowledge and insights from a wide variety of data sources to analyse patterns or predict future behaviour. It draws from a wide array of disciplines including statistics, computer science, mathematics, machine learning, and data mining. In this book, we cover the important data science concepts and how they are supported by Java, as well as the often statistically challenging techniques, to provide you with an understanding of their purpose and application. The book starts with an introduction of data science, followed by the basic data science tasks of data collection, data cleaning, data analysis, and data visualization. This is followed by a discussion of statistical techniques and more advanced topics including machine learning, neural networks, and deep learning. The next section examines the major categories of data analysis including text, visual, and audio data, followed by a discussion of resources that support parallel implementation. The final chapter illustrates an in-depth data science problem and provides a comprehensive, Java-based solution. Due to the nature of the topic, simple examples of techniques are presented early followed by a more detailed treatment later in the book. This permits a more natural introduction to the techniques and concepts presented in the book. Style and approach This book follows a tutorial approach, providing examples of each of the major concepts covered. With a step-by-step instructional style, this book covers various facets of data science and will get you up and running quickly.*

*Become an advanced practitioner with this progressive set of master classes on application-oriented machine learning About This Book Comprehensive coverage of key topics in machine learning with an emphasis on both the theoretical and practical aspects More than 15 open source Java tools in a wide range of techniques, with code and practical usage. More than 10 real-world case studies in machine learning highlighting techniques ranging from data ingestion up to analyzing the results of experiments, all preparing the user for the practical, real-world use of tools and data analysis. Who This Book Is For This book will appeal to anyone with a serious interest in topics in Data Science or those already working in related areas: ideally, intermediate-level data analysts and data scientists with experience in Java. Preferably, you will have experience with the fundamentals of machine learning and now have a desire to explore the area further, are up to grappling with the mathematical complexities of its algorithms, and you wish to learn the complete ins and outs of practical machine learning. What You Will Learn Master key Java machine learning libraries, and what kind of problem each can solve, with theory and practical guidance. Explore powerful techniques in each major category of machine learning such as classification, clustering, anomaly detection, graph modeling, and text mining. Apply machine learning to real-world data with methodologies, processes, applications, and analysis. Techniques and experiments developed around the latest specializations in machine learning, such as deep learning, stream data mining, and active and semi-supervised learning. Build high-performing, real-time, adaptive predictive models for batch- and stream-based big data learning using the latest tools and methodologies. Get a deeper understanding of technologies leading towards a more powerful AI applicable in various domains such as Security, Financial Crime, Internet of Things, social networking, and so on. In Detail Java is one of the main languages used by practicing data scientists; much of the Hadoop ecosystem is Java-based, and it is certainly the language that most production systems in Data Science are written in. If you know Java, Mastering Machine Learning with Java is your next step on the path to becoming an advanced practitioner in Data Science. This book aims to introduce you to an array of advanced techniques in machine learning, including classification, clustering, anomaly detection, stream learning, active learning, semi-supervised learning, probabilistic graph modeling, text mining, deep learning, and big data batch and stream machine learning. Accompanying each chapter are illustrative examples and real-world case studies that show how to apply the newly learned techniques using sound methodologies and the best Java-based tools available today. On completing this book, you will have an understanding of the tools and techniques for building powerful machine learning models to solve data science problems in just about any domain. Style and approach A practical guide to help you explore machine learning—and an array of Java-based tools and frameworks—with the help of practical examples and real-world use cases.*

*This book shows how UNIX can be used effectively in the preparation of written documents, especially in the process of producing book-length documents, i.e. typesetting. As this book will demonstrate the tools available in the UNIX environment, it is also valuable to examine text processing in terms of problems and solutions: the problems faced by a writer undertaking a large writing project and the solutions offered by using the resources and power of a computer system.*

*Text is everywhere. Web pages, databases, the contents of files--for almost any programming task you perform, you need to process text. Cut even the most complex text-based tasks down to size and learn how to master regular expressions, scrape information from Web pages, develop reusable utilities to process text in pipelines, and more. Most information in the world is in text format, and programmers often find themselves needing to make sense of the data hiding within. It might be to convert it from one format to another, or to find out information about the text as a whole, or to extract information fromit. But how do you do this efficiently, avoiding labor-intensive, manual work? Text Processing with Ruby takes a practical approach. You'll learn how to get text into your Ruby programs from the file system and from user input. You'll process delimited files such as CSVs, and write utilities that interact with other programs in text-processing pipelines. Decipher character encoding mysteries, and avoid the pain of jumbled characters and malformed output. You'll learn to use regular expressions to match, extract, and replace patterns in text. You'll write a parser and learn how to process Web pages to pull out information from even the messiest of HTML. Before long you'll be able to tackle even the most enormous and entangled text with ease, scything through gigabytes of data and effortlessly extracting the bits that matter. What You Need: This book requires a passing familiarity with the Ruby programming language, and assumes that you already have Ruby installed on your computer.*

*A Beginner's Guide to Programming Images, Animation, and Interaction*

*Regular Expressions and High-Performance I/O*
*Groovy 2 Cookbook*
*Natural Language Processing with Java*
*Build innovative deep neural network architectures for NLP with Python, PyTorch, TensorFlow, BERT, RoBERTa, and more*
*Java Deep Learning Cookbook*
*Deep Learning for Natural Language Processing*
*The Bulgarian C# Book*
*Java NIO*
*An Algorithmic Introduction Using Java*
*Mastering Java for Data Science*
*A Practitioner's Guide to Natural Language Processing*

A problem-solution guide to encounter various NLP tasks utilizing Java open source libraries and cloud-based solutions Key Features Perform simple-to-complex NLP text processing tasks using modern Java libraries Extract relationships between different text comp using a problem-solution approach Utilize cloud-based APIs to perform machine translation operations Book Description Natural Language Processing (NLP) has become one of the prime technologies for processing very large amounts of unstructured data from dis information sources. This book includes a wide set of recipes and quick methods that solve challenges in text syntax, semantics, and speech tasks. At the beginning of the book, you'll learn important NLP techniques, such as identifying parts of speech, tagging word analyzing word semantics. You will learn how to perform lexical analysis and use machine learning techniques to speed up NLP operations. With independent recipes, you will explore techniques for customizing your existing NLP engines/models using Java libraries OpenNLP and the Stanford NLP library. You will also learn how to use NLP processing features from cloud-based sources, including Google and Amazon's AWS. You will master core tasks, such as stemming, lemmatization, part-of-speech tagging, and named entit recognition. You will also learn about sentiment analysis, semantic text similarity, language identification, machine translation, and text summarization. By the end of this book, you will be ready to become a professional NLP expert using a problem-solution approach analyze any sort of text, sentences, or semantic words. What you will learn Explore how to use tokenizers in NLP processing Implement NLP techniques in machine learning and deep learning applications Identify sentences within the text and learn how to train spec NER models Learn how to classify documents and perform sentiment analysis Find semantic similarities between text elements and extract text from a variety of sources Preprocess text from a variety of data sources Learn how to identify and translate languages W book is for This book is for data scientists, NLP engineers, and machine learning developers who want to perform their work on linguistic applications faster with the use of popular libraries on JVM machines. This book will help you build real-world NLP applications u recipe-based approach. Prior knowledge of Natural Language Processing basics and Java programming is expected.

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 prog 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 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 popu 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 th 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 la works -- you'll find Natural Language Processing with Python both fascinating and immensely useful.

Use Java to create a diverse range of Data Science applications and bring Data Science into production About This Book An overview of modern Data Science and Machine Learning libraries available in Java Coverage of a broad set of topics, going from the basics Machine Learning to Deep Learning and Big Data frameworks. Easy-to-follow illustrations and the running example of building a search engine. Who This Book Is For This book is intended for software engineers who are comfortable with developing Java application are familiar with the basic concepts of data science. Additionally, it will also be useful for data scientists who do not yet know Java but want or need to learn it. If you are willing to build efficient data science applications and bring them in the enterprise environment w changing the existing stack, this book is for you! What You Will Learn Get a solid understanding of the data processing toolbox available in Java Explore the data science ecosystem available in Java Find out how to approach different machine learning problems with Process unstructured information such as natural language text or images Create your own search engine Get state-of-the-art performance with XGBoost Learn how to build deep neural networks with DeepLearning4j Build applications that scale and process large a of data Deploy data science models to production and evaluate their performance In Detail Java is the most popular programming language, according to the TIOBE index, and it is a typical choice for running production systems in many companies, both in the startu and among large enterprises. Not surprisingly, it is also a common choice for creating data science applications: it is fast and has a great set of data processing tools, both built-in and external. What is more, choosing Java for data science allows you to easily integra solutions with existing software, and bring data science into production with less effort. This book will teach you how to create data science applications with Java. First, we will revise the most important things when starting a data science application, and then brush basics of Java and machine learning before diving into more advanced topics. We start by going over the existing libraries for data processing and libraries with machine learning algorithms. After that, we cover topics such as classification and regression, dimensiona reduction and clustering, information retrieval and natural language processing, and deep learning and big data. Finally, we finish the book by talking about the ways to deploy the model and evaluate it in production settings. Style and approach This is a practical gui all the important concepts such as classification, regression, and dimensionality reduction are explained with the help of examples.

This book follows a Cookbook style and is packed with intermediate and advanced level recipes.This book is for Java developers who have an interest in discovering new ways to quickly get the job done using a new language that shares many similarities with Java. recipes start simple, therefore no previous Groovy experience is required to understand the code and the explanations accompanying the examples.

Leverage Natural Language Processing (NLP) in Python and learn how to set up your own robust environment for performing text analytics. This second edition has gone through a major revamp and introduces several significant changes and new topics based on th trends in NLP. You'll see how to use the latest state-of-the-art frameworks in NLP, coupled with machine learning and deep learning models for supervised sentiment analysis powered by Python to solve actual case studies. Start by reviewing Python for NLP fundam strings and text data and move on to engineering representation methods for text data, including both traditional statistical models and newer deep learning-based embedding models. Improved techniques and new methods around parsing and processing text are di well. Text summarization and topic models have been overhauled so the book showcases how to build, tune, and interpret topic models in the context of an interest dataset on NIPS conference papers. Additionally, the book covers text similarity techniques with a rea example of movie recommenders, along with sentiment analysis using supervised and unsupervised techniques. There is also a chapter dedicated to semantic analysis where you'll see how to build your own named entity recognition (NER) system from scratch. Whi overall structure of the book remains the same, the entire code base, modules, and chapters has been updated to the latest Python 3.x release. What You'll Learn • Understand NLP and text syntax, semantics and structure• Discover text cleaning and feature engine Review text classification and text clustering • Assess text summarization and topic models• Study deep learning for NLP Who This Book Is For IT professionals, data analysts, developers, linguistic experts, data scientists and engineers and basically anyone with a k interest in linguistics, analytics and generating insights from textual data.

Haskell in Depth unlocks a new level of skill with this challenging language. Going beyond the basics of syntax and structure, this book opens up critical topics like advanced types, concurrency, and data processing. Summary Turn the corner from "Haskell student" t "Haskell developer." Haskell in Depth explores the important language features and programming skills you'll need to build production-quality software using Haskell. And along the way, you'll pick up some interesting insights into why Haskell looks and works the wa does. Get ready to go deep! Purchase of the print book includes a free eBook in PDF, Kindle, and ePub formats from Manning Publications. About the technology Software for high-precision tasks like financial transactions, defense systems, and scientific research m absolutely, provably correct. As a purely functional programming language, Haskell enforces a mathematically rigorous approach that can lead to concise, efficient, and bug-free code. To write such code you'll need deep understanding. You can get it from this book! the book Haskell in Depth unlocks a new level of skill with this challenging language. Going beyond the basics of syntax and structure, this book opens up critical topics like advanced types, concurrency, and data processing. You'll discover key parts of the Haskell e and master core design patterns that will transform how you write software. What's inside Building applications, web services, and networking apps Using sophisticated libraries like lens, singletons, and servant Organizing projects with Cabal and Stack Error-handlin testing Pure parallelism for multicore processors About the reader For developers familiar with Haskell basics. About the author Vitaly Bragilevsky has been teaching Haskell and functional programming since 2008. He is a member of the GHC Steering Committee. T Contents PART 1 CORE HASKELL 1 Functions and types 2 Type classes 3 Developing an application: Stock quotes PART 2 INTRODUCTION TO APPLICATION DESIGN 4 Haskell development with modules, packages, and projects 5 Monads as practical function providers 6 Structuring programs with monad transformers PART 3 QUALITY ASSURANCE 7 Error handling and logging 8 Writing tests 9 Haskell data and code at run time 10 Benchmarking and profiling PART 4 ADVANCED HASKELL 11 Type system advances 1 Metaprogramming in Haskell 13 More about types PART 5 HASKELL TOOLKIT 14 Data-processing pipelines 15 Working with relational databases 16 Concurrency

Providing an introduction to the ideas of computer programming within the context of the visual arts, this thorough book targets an audience of computer-savvy individuals who are interested in creating interactive and visual work through writing software but have littl prior experience. --

Text mining is a new and exciting area of computer science research that tries to solve the crisis of information overload by combining techniques from data mining, machine learning, natural language processing, information retrieval, and knowledge management. S link detection – a rapidly evolving approach to the analysis of text that shares and builds upon many of the key elements of text mining – also provides new tools for people to better leverage their burgeoning textual data resources. The Text Mining Handbook presen comprehensive discussion of the state-of-the-art in text mining and link detection. In addition to providing an in-depth examination of core text mining and link detection algorithms and operations, the book examines advanced pre-processing techniques, knowledge representation considerations, and visualization approaches. Finally, the book explores current real-world, mission-critical applications of text mining and link detection in such varied fields as M&A business intelligence, genomics research and counter-terrorism activ

Java Cookbook
How to Do Text Analysis with Java
Mastering Java Machine Learning
11th International Conference, CICLing 2010, Iasi, Romania, March 21-27, 2010, Proceedings
Learning Processing
Extract Value from the Data That Surrounds You
Haskell in Depth
Natural Language Processing with Python
Natural Language Processing with Java Cookbook
A Tidy Approach
Text Mining with R
Java Data Science Cookbook

Use Java and Deeplearning4j to build robust, scalable, and highly accurate AI models from scratch Key Features Install and configure Deeplearning4j to implement deep learning models from scratch Explore recipes for developing, training, and fine-tuning your neural network models in Java Model neural netw datasets containing images, text, and time-series data Book Description Java is one of the most widely used programming languages in the world. With this book, you will see how to perform deep learning using Deeplearning4j (DL4J) – the most popular Java library for training neural networks efficiently. This showing you how to install and configure Java and DL4J on your system. You will then gain insights into deep learning basics and use your knowledge to create a deep neural network for binary classification from scratch. As you progress, you will discover how to build a convolutional neural network (CNN) in understand how to construct numeric vectors from text. This deep learning book will also guide you through performing anomaly detection on unsupervised data and help you set up neural networks in distributed systems effectively. In addition to this, you will learn how to import models from Keras and change in a pre-trained DL4J model. Finally, you will explore benchmarking in DL4J and optimize neural networks for optimal results. By the end of this book, you will have a clear understanding of how you can use DL4J to build robust deep learning applications in Java. What you will learn Perform data normalization using DL4J Build deep neural networks using DL4J Implement CNNs to solve image classification problems Train autoencoders to solve anomaly detection problems using DL4J Perform benchmarking and optimization to improve your model's performance Implement reinforcement learning for real-world use Leverage the capabilities of DL4J in distributed systems Who this book is for If you are a data scientist, machine learning developer, or a deep learning enthusiast who wants to implement deep learning models in Java, this book is for you. Basic understanding of Java programming as well as some experience learning and neural networks is required to get the most out of this book.

Summary Taming Text, winner of the 2013 Jolt Awards for Productivity, is a hands-on, example-driven guide to working with unstructured text in the context of real-world applications. This book explores how to automatically organize text using approaches such as full-text search, proper name recognition, clus information extraction, and summarization. The book guides you through examples illustrating each of these topics, as well as the foundations upon which they are built. About this Book There is so much text in our lives, we are practically drowningin it. Fortunately, there are innovative tools and techniquesfor u unstructured information that can throw thesmart developer a much-needed lifeline. You'll find them in thisbook. Taming Text is a practical, example-driven guide to working withtext in real applications. This book introduces you to useful techniques like full-text search, proper name recognition,clustering, taggin extraction, and summarization.You'll explore real use cases as you systematically absorb thefoundations upon which they are built.Written in a clear and concise style, this book avoids jargon, explainingthe subject in terms you can understand without a backgroundin statistics or natural language processing. E Java, but the concepts can be applied in any language. Written for Java developers, the book requires no prior knowledge of GWT. Purchase of the print book comes with an offer of a free PDF, ePub, and Kindle eBook from Manning. Also available is all code from the book. Winner of 2013 Jolt Awards: The B five notable books every serious programmer should read. What's Inside When to use text-taming techniques Important open-source libraries like Solr and Mahout How to build text-processing applications About the Authors Grant Ingersoll is an engineer, speaker, and trainer, a Lucenecommitter, and a cofoun machine-learning project. Thomas Morton is the primary developer of OpenNLP and Maximum Entropy. Drew Farris is a technology consultant, software developer, and contributor to Mahout,Lucene, and Solr. "Takes the mystery out of verycomplex processes."—From the Foreword by Liz Liddy, Dean, iSchoo University Table of Contents Getting started taming text Foundations of taming text Searching Fuzzy string matching Identifying people, places, and things Clustering text Classification, categorization, and tagging Building an example question answering system Untamed text: exploring the next frontier

This book is intended for Python programmers interested in learning how to do natural language processing. Maybe you've learned the limits of regular expressions the hard way, or you've realized that human language cannot be deterministically parsed like a computer language. Perhaps you have more text what to do with, and need automated ways to analyze and structure that text. This Cookbook will show you how to train and use statistical language models to process text in ways that are practically impossible with standard programming tools. A basic knowledge of Python and the basic text processing conce Some experience with regular expressions will also be helpful.

This book teaches you how to master the subtle art of multilingual text processing and prevent text data corruption. It provides an introduction to natural language processing using Lucene and Solr. It gives you tools and techniques to manage large collections of text data, whether they come from news feeds, documents. Each chapter contains executable programs that can also be used for text data forensics. Topics covered: Unicode code points Character encodings from ASCII and Big5 to UTF-8 and UTF-32LE Character normalization using International Components for Unicode (ICU) Java I/O, including workin zip, gzip, and tar files Regular expressions in Java Transporting text data via HTTP Parsing and generating XML, HTML, and JSON Using Lucene 4 for natural language search and text classification Search, spelling correction, and clustering with Solr 4 Other books on text processing presuppose much of the in this book. They gloss over the details of transforming text from one format to another and assume perfect input data. The messy reality of raw text will have you reaching for this book again and again.

Master text-taming techniques and build effective text-processing applications with R About This Book Develop all the relevant skills for building text-mining apps with R with this easy-to-follow guide Gain in-depth understanding of the text mining process with lucid implementation in the R language Example-ri you gain high-quality information from text data Who This Book Is For If you are an R programmer, analyst, or data scientist who wants to gain experience in performing text data mining and analytics with R, then this book is for you. Exposure to working with statistical methods and language processing would You Will Learn Get acquainted with some of the highly efficient R packages such as OpenNLP and RWeka to perform various steps in the text mining process Access and manipulate data from different sources such as JSON and HTTP Process text using regular expressions Get to know the different approac such as POS tagging, to get started with text analysis Explore different dimensionality reduction techniques, such as Principal Component Analysis (PCA), and understand its implementation in R Discover the underlying themes or topics that are present in an unstructured collection of documents, using comm as Latent Dirichlet Allocation (LDA) Build a baseline sentence completing application Perform entity extraction and named entity recognition using R In Detail Text Mining (or text data mining or text analytics) is the process of extracting useful and high-quality information from text by devising patterns and trend an extensive ecosystem to mine text through its many frameworks and packages. Starting with basic information about the statistics concepts used in text mining, this book will teach you how to access, cleanse, and process text using the R language and will equip you with the tools and the associated knowle tagging, chunking, and entailment approaches and their usage in natural language processing. Moving on, this book will teach you different dimensionality reduction techniques and their implementation in R. Next, we will cover pattern recognition in text data utilizing classification mechanisms, perform entity re develop an ontology learning framework. By the end of the book, you will develop a practical application from the concepts learned, and will understand how text mining can be leveraged to analyze the massively available data on social media. Style and approach This book takes a hands-on, example-driven a mining process with lucid implementation in R.

Deep learning methods are achieving state-of-the-art results on challenging machine learning problems such as describing photos and translating text from one language to another. In this new laser-focused Ebook, finally cut through the math, research papers and patchwork descriptions about natural langua Using clear explanations, standard Python libraries and step-by-step tutorial lessons you will discover what natural language processing is, the promise of deep learning in the field, how to clean and prepare text data for modeling, and how to develop deep learning models for your own natural language proces

Recipes to help you overcome your data science hurdles using Java About This Book This book provides modern recipes in small steps to help an apprentice cook become a master chef in data science Use these recipes to obtain, clean, analyze, and learn from your data Learn how to get your data science a production and enterprise environments effortlessly Who This Book Is For This book is for Java developers who are familiar with the fundamentals of data science and want to improve their skills to become a pro. What You Will Learn Find out how to clean and make datasets ready so you can acquire actual in noise and outliers Develop the skills to use modern machine learning techniques to retrieve information and transform data to knowledge. retrieve information from large amount of data in text format. Familiarize yourself with cutting-edge techniques to store and search large volumes of data and retrieve inform amounts of data in text format Develop basic skills to apply big data and deep learning technologies on large volumes of data Evolve your data visualization skills and gain valuable insights from your data Get to know a step-by-step formula to develop an industry-standard, large-scale, real-life data product Ga visualize data and interact with users through data insights In Detail If you are looking to build data science models that are good for production, Java has come to the rescue. With the aid of strong libraries such as MLlib, Weka, DL4j, and more, you can efficiently perform all the data science tasks you need to. provides modern recipes to solve your common and not-so-common data science-related problems. We start with recipes to help you obtain, clean, index, and search data. Then you will learn a variety of techniques to analyze, learn from, and retrieve information from data. You will also understand how to han deeply from data, and visualize data. Finally, you will work through unique recipes that solve your problems while taking data science to production, writing distributed data science applications, and much more—things that will come in handy at work. Style and approach This book contains short yet very effecti most common problems. Some recipes cater to very specific, rare pain points. The recipes cover different data sets and work very closely to real production environments

If you are a Java programmer who wants to learn about the fundamental tasks underlying natural language processing, this book is for you. You will be able to identify and use NLP tasks for many common problems, and integrate them in your applications to solve more difficult problems. Readers should be fa with Java software development.

A Guide to Corpus-Building for Applications
A Programming Handbook for Visual Designers and Artists
Text Processing in Java
Advanced Approaches in Analyzing Unstructured Data
Computational Linguistics and Intelligent Text Processing
The Art of Image Processing with Java
Machine Learning in Java
Analyzing Text with the Natural Language Toolkit
Think Java
Transformers for Natural Language Processing
How to Find, Organize, and Manipulate It
Taming Text

From lambda expressions and JavaFX 8 to new support for network programming and mobile development, Java 8 brings a wealth of changes. This cookbook helps you get up to speed right away with hundreds of hands-on recipes across a broad range of Java topics. You'll learn useful techniques for everything from debugging and data structures to GUI development and functional programming. Each recipe includes self-contained code solutions that you can freely use, along with a discussion of how and why they work. If you are familiar with Java basics, this cookbook will bolster your knowledge of the language in general and Java 8's main APIs in particular. Recipes include: Methods for compiling, running, and debugging Manipulating, comparing, and rearranging text Regular expressions for string- and pattern-matching Handling numbers, dates, and times Structuring data with collections, arrays, and other types Object-oriented and functional programming techniques Directory and filesystem operations Working with graphics, audio, and video GUI development, including JavaFX and handlers Network programming on both client and server Database access, using JPA, Hibernate, and JDBC Processing JSON and XML for data storage Multithreading and concurrency

bull; Demonstrates how Python is the perfect language for text-processing functions. bull; Provides practical pointers and tips that emphasize efficient, flexible, and maintainable approaches to text-processing challenges. bull; Helps programmers develop solutions for dealing with the increasing amounts of data with which we are all inundated.

Currently used at many colleges, universities, and high schools, this hands-on introduction to computer science is ideal for people with little or no programming experience. The goal of this concise book is not just to teach you Java, but to help you think like a computer scientist. You'll learn how to

program—a useful skill by itself—but you'll also discover how to use programming as a means to an end. Authors Allen Downey and Chris Mayfield start with the most basic concepts and gradually move into topics that are more complex, such as recursion and object-oriented programming. Each brief chapter covers the material for one week of a college course and includes exercises to help you practice what you've learned. Learn one concept at a time: tackle complex topics in a series of small steps with examples Understand how to formulate problems, think creatively about solutions, and write programs clearly and accurately Determine which development techniques work best for you, and practice the important skill of debugging Learn relationships among input and output, decisions and loops, classes and methods, strings and arrays Work on exercises involving word games, graphics, puzzles, and playing cards

"In this course we will cover the essence of natural language processing (NLP) using Java. This video course will commence by walking you through basic NLP tasks including data acquisition, data cleaning, finding parts of text, and determining the end of sentences. These serve as the basis for other NLP tasks such as classifying text and determining the relationship between text elements. This will be followed by the use of tokenization techniques. Tokenization is used for almost all NLP tasks. You will learn how text can be split to reveal information such as names, dates, and even the grammatical structure of a sentence. These types of activity can lead to insights into the relationships between text elements and embedded meaning in a document. "--Resource description page.

Being the first book in the market to dive deep into the Transformers, it is a step-by-step guide for data and AI practitioners to help enhance the performance of language understanding and gain expertise with hands-on implementation of transformers using PyTorch, TensorFlow, Hugging Face, Trax, and AllenNLP.

While most other image processing texts approach this subject from an engineering perspective, The Art of Image Processing with Java places image processing within the realm of both engineering and computer science students by emphasizing software design. Ideal for students studying computer science or software engineering, it clearly teaches

The free book "Fundamentals of Computer Programming with C#" is a comprehensive computer programming tutorial that teaches programming, logical thinking, data structures and algorithms, problem solving and high quality code with lots of examples in C#. It starts with the first steps in programming and software development like variables, data types, conditional statements, loops and arrays and continues with other basic topics like methods, numeral systems, strings and string processing, exceptions, classes and objects. After the basics this fundamental programming book enters into more advanced programming topics like recursion, data structures (lists, trees, hash-tables and graphs), high-quality code, unit testing and refactoring, object-oriented principles (inheritance, abstraction, encapsulation and polymorphism) and their implementation the C# language. It also covers fundamental topics that each good developer should know like algorithm design, complexity of algorithms and problem solving. The book uses C# language and Visual Studio to illustrate the programming concepts and explains some C# / .NET specific technologies like lambda expressions, extension methods and LINQ. The book is written by a team of developers lead by Svetlin Nakov who has 20+ years practical software development experience. It teaches the major programming concepts and way of thinking needed to become a good software engineer and the C# language in the meantime. It is a great start for anyone who wants to become a skillful software engineer. The books does not teach technologies like databases, mobile and web development, but shows the true way to master the basics of programming regardless of the languages, technologies and tools. It is good for beginners and intermediate developers who want to put a solid base for a successful career in the software engineering industry. The book is accompanied by free video lessons, presentation slides and mind maps, as well as hundreds of exercises and live examples. Download the free C# programming book, videos, presentations and other resources from http://introprogramming.info. Title: Fundamentals of Computer Programming with C# (The Bulgarian C# Programming Book) ISBN: 9789544007737 ISBN-13: 978-954-400-773-7 (9789544007737) ISBN-10: 954-400-773-3 (9544007733) Author: Svetlin Nakov & Co. Pages: 1132 Language: English Published: Sofia, 2013 Publisher: Faber Publishing, Bulgaria Web site: http://www.introprogramming.info License: CC-Attribution-Share-Alike Tags: free, programming, book, computer programming, programming fundamentals, ebook, book programming, C#, CSharp, C# book, tutorial, C# tutorial; programming concepts, programming fundamentals, compiler, Visual Studio, .NET, .NET Framework, data types, variables, expressions, statements, console, conditional statements, control-flow logic, loops, arrays, numeral systems, methods, strings, text processing, StringBuilder, exceptions, exception handling, stack trace, streams, files, text files, linear data structures, list, linked list, stack, queue, tree, balanced tree, graph, depth-first search, DFS, breadth-first search, BFS, dictionaries, hash tables, associative arrays, sets, algorithms, sorting algorithm, searching algorithms, recursion, combinatorial algorithms, algorithm complexity, OOP, object-oriented programming, classes, objects, constructors, fields, properties, static members, abstraction, interfaces, encapsulation, inheritance, virtual methods, polymorphism, cohesion, coupling, enumerations, generics, namespaces, UML, design patterns, extension methods, anonymous types, lambda expressions, LINQ, code quality, high-quality code, high-quality classes, high-quality methods, code formatting, self-documenting code, code refactoring, problem solving, problem solving methodology, 9789544007737, 9544007733

The Text Mining Handbook
Getting Started with Natural Language Processing in Java
Python 3 Text Processing with NLTK 3 Cookbook
Java for Data Science
Solutions and Examples for Java Developers
Techniques for building machine learning and neural network models for NLP, 2nd Edition
UNIX Text Processing
Text Processing with Ruby
Programming for Corpus Linguistics
Fundamentals of Computer Programming with C#
Mining the Social Web
Natural Language Annotation for Machine Learning