Acces PDF Technologies Language Human On Lectures Synthesis Processing Language Natural For Methods Network Neural

Right here, we have countless books **Technologies Language Human On Lectures Synthesis Processing Language Natural For Methods Network Neural** and collections to check out. We additionally give variant types and moreover type of the books to browse. The suitable book, fiction, history, novel, scientific research, as competently as various supplementary sorts of books are readily within reach here.

As this Technologies Language Human On Lectures Synthesis Processing Language Natural For Methods Network Neural, it ends occurring beast one of the favored books Technologies Language Human On Lectures Synthesis Processing Language Natural For Methods Network Neural collections that we have. This is why you remain in the best website to look the incredible book to have.

KEY=NATURAL - JUSTICE JORDAN

Data-intensive Text Processing with MapReduce

Morgan & Claypool Publishers **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

Learning to Rank for Information Retrieval and Natural Language Processing, Second Edition

Springer Nature **Learning to rank refers to machine learning techniques for training a model in a ranking task. Learning to rank is useful for many applications in information retrieval, natural language processing, and data mining. Intensive studies have been conducted on its problems recently, and significant progress has been made. This lecture gives an introduction to the area including the fundamental problems, major approaches, theories, applications, and future work. The author begins by showing that various ranking problems in information retrieval and natural language processing can be formalized as two basic ranking tasks, namely ranking creation (or simply ranking) and ranking aggregation. In ranking creation, given a request, one wants to generate a ranking list of offerings based on the features derived from the request and the offerings. In ranking aggregation, given a request, as well as a number of ranking lists of offerings, one wants to generate a new ranking list of the offerings. Ranking creation (or ranking) is the major problem in learning to rank. It is usually formalized as a supervised learning task. The author gives detailed explanations on learning for ranking creation and ranking aggregation, including training and testing, evaluation, feature creation, and major approaches. Many methods have been proposed for ranking creation. The methods can be categorized as the pointwise, pairwise, and listwise approaches according to the loss functions they employ. They can also be categorized according to the techniques they employ, such as the SVM based, Boosting based, and Neural Network based approaches. The author also introduces some popular learning to rank methods in details. These include: PRank, OC SVM, McRank, Ranking SVM, IR SVM, GBRank, RankNet, ListNet & ListMLE, AdaRank, SVM MAP, SoftRank, LambdaRank, LambdaMART, Borda Count, Markov Chain, and CRanking. The author explains several example**

applications of learning to rank including web search, collaborative filtering, definition search, keyphrase extraction, query dependent summarization, and re-ranking in machine translation. A formulation of learning for ranking creation is given in the statistical learning framework. Ongoing and future research directions for learning to rank are also discussed. Table of Contents: Learning to Rank / Learning for Ranking Creation / Learning for Ranking Aggregation / Methods of Learning to Rank / Applications of Learning to Rank / Theory of Learning to Rank / Ongoing and Future Work

Arabic Natural Language Processing

Morgan & Claypool Publishers **This book provides system developers and researchers in natural language processing and computational linguistics with the necessary background information for working with the Arabic language. The goal is to introduce Arabic linguistic phenomena and review the state-of-the-art in Arabic processing. The book discusses Arabic script, phonology, orthography, morphology, syntax and semantics, with a final chapter on machine translation issues. The chapter sizes correspond more or less to what is linguistically distinctive about Arabic, with morphology getting the lion's share, followed by Arabic script. No previous knowledge of Arabic is needed. This book is designed for computer scientists and linguists alike. The focus of the book is on Modern Standard Arabic; however, notes on practical issues related to Arabic dialects and languages written in the Arabic script are presented in different chapters. Table of Contents: What is "Arabic"? / Arabic Script / Arabic Phonology and Orthography / Arabic Morphology / Computational Morphology Tasks / Arabic Syntax / A Note on Arabic Semantics / A Note on Arabic and Machine Translation**

Semi-Supervised Learning and Domain Adaptation in Natural Language Processing

Morgan & Claypool Publishers **This book introduces basic supervised learning algorithms applicable to natural language processing (NLP) and shows how the performance of these algorithms can often be improved by exploiting the marginal distribution of large amounts of unlabeled data. One reason for that is data sparsity, i.e., the limited amounts of data we have available in NLP. However, in most real-world NLP applications our labeled data is also heavily biased.**

This book introduces extensions of supervised learning algorithms to cope with data sparsity and different kinds of sampling bias. This book is intended to be both readable by first-year students and interesting to the expert audience. My intention was to introduce what is necessary to appreciate the major challenges we face in contemporary NLP related to data sparsity and sampling bias, without wasting too much time on details about supervised learning algorithms or particular NLP applications. I use text classification, part-of-speech tagging, and dependency parsing as running examples, and limit myself to a small set of cardinal learning algorithms. I have worried less about theoretical guarantees ("this algorithm never does too badly") than about useful rules of thumb ("in this case this algorithm may perform really well"). In NLP, data is so noisy, biased, and non-stationary that few theoretical guarantees can be established and we are typically left with our gut feelings and a catalogue of crazy ideas. I hope this book will provide its readers with both. Throughout the book we include snippets of Python code and empirical evaluations, when relevant.

Natural Language Processing for Historical Texts

Morgan & Claypool Publishers **More and more historical texts are becoming available in digital form. Digitization of paper documents is motivated by the aim of preserving cultural heritage and making it more accessible, both to laypeople and scholars. As digital images cannot be searched for text, digitization projects increasingly strive to create digital text, which can be searched and otherwise automatically processed, in addition to facsimiles. Indeed, the emerging field of digital humanities heavily relies on the availability of digital text for its studies. Together with the increasing availability of historical texts in digital form, there is a growing interest in applying natural language processing (NLP) methods and tools to historical texts. However, the specific linguistic properties of historical texts -- the lack of standardized orthography, in particular -- pose special challenges for NLP. This book aims to give an introduction to NLP for historical texts and an overview of the state of the art in this field. The book starts with an overview of methods for the acquisition of historical texts (scanning and OCR), discusses text encoding and annotation schemes, and presents examples of corpora of historical texts in a variety of languages. The book then discusses specific methods, such as creating part-of-speech taggers for historical languages or handling spelling variation. A final chapter analyzes the relationship between NLP and the digital humanities. Certain recently emerging textual genres, such as SMS, social media, and chat messages, or newsgroup and forum postings share a number of properties with historical texts, for example, nonstandard orthography and grammar, and profuse use of abbreviations. The methods**

and techniques required for the effective processing of historical texts are thus also of interest for research in other domains. Table of Contents: Introduction / NLP and Digital Humanities / Spelling in Historical Texts / Acquiring Historical Texts / Text Encoding and Annotation Schemes / Handling Spelling Variation / NLP Tools for Historical Languages / Historical Corpora / Conclusion / Bibliography

Explainable Natural Language Processing

Springer Nature **This book presents a taxonomy framework and survey of methods relevant to explaining the decisions and analyzing the inner workings of Natural Language Processing (NLP) models. The book is intended to provide a snapshot of Explainable NLP, though the field continues to rapidly grow. The book is intended to be both readable by first-year M.Sc. students and interesting to an expert audience. The book opens by motivating a focus on providing a consistent taxonomy, pointing out inconsistencies and redundancies in previous taxonomies. It goes on to present (i) a taxonomy or framework for thinking about how approaches to explainable NLP relate to one another; (ii) brief surveys of each of the classes in the taxonomy, with a focus on methods that are relevant for NLP; and (iii) a discussion of the inherent limitations of some classes of methods, as well as how to best evaluate them. Finally, the book closes by providing a list of resources for further research on explainability.**

Web Corpus Construction

Springer Nature **The World Wide Web constitutes the largest existing source of texts written in a great variety of languages. A feasible and sound way of exploiting this data for linguistic research is to compile a static corpus for a given language. There are several adavantages of this approach: (i) Working with such corpora obviates the problems encountered when using Internet search engines in quantitative linguistic research (such as non-transparent ranking algorithms). (ii) Creating a corpus from web data is virtually free. (iii) The size of corpora compiled from the WWW may exceed by several orders of magnitudes the size of language resources offered elsewhere. (iv) The data is locally available to the user, and it can be linguistically post-processed and queried with the tools preferred by her/him. This book addresses the main practical tasks in the creation of web corpora up to giga-token size. Among these tasks are the sampling process (i.e., web crawling) and the usual cleanups including boilerplate removal and removal of duplicated content. Linguistic processing and problems with linguistic processing coming from the different kinds of**

noise in web corpora are also covered. Finally, the authors show how web corpora can be evaluated and compared to other corpora (such as traditionally compiled corpora). For additional material please visit the companion website: sites.morganclaypool.com/wcc Table of Contents: Preface / Acknowledgments / Web Corpora / Data Collection / Post-Processing / Linguistic Processing / Corpus Evaluation and Comparison / Bibliography / Authors' Biographies

Ontology-Based Interpretation of Natural Language

Morgan & Claypool Publishers **For humans, understanding a natural language sentence or discourse is so effortless that we hardly ever think about it. For machines, however, the task of interpreting natural language, especially grasping meaning beyond the literal content, has proven extremely difficult and requires a large amount of background knowledge. This book focuses on the interpretation of natural language with respect to specific domain knowledge captured in ontologies. The main contribution is an approach that puts ontologies at the center of the interpretation process. This means that ontologies not only provide a formalization of domain knowledge necessary for interpretation but also support and guide the construction of meaning representations. We start with an introduction to ontologies and demonstrate how linguistic information can be attached to them by means of the ontology lexicon model lemon. These lexica then serve as basis for the automatic generation of grammars, which we use to compositionally construct meaning representations that conform with the vocabulary of an underlying ontology. As a result, the level of representational granularity is not driven by language but by the semantic distinctions made in the underlying ontology and thus by distinctions that are relevant in the context of a particular domain. We highlight some of the challenges involved in the construction of ontology-based meaning representations, and show how ontologies can be exploited for ambiguity resolution and the interpretation of temporal expressions. Finally, we present a question answering system that combines all tools and techniques introduced throughout the book in a real-world application, and sketch how the presented approach can scale to larger, multi-domain scenarios in the context of the Semantic Web.**

Automated Essay Scoring

Springer Nature **This book discusses the state of the art of automated essay scoring, its challenges and its potential. One of the earliest applications of artificial intelligence to language data (along with machine translation and speech**

recognition), automated essay scoring has evolved to become both a revenue-generating industry and a vast field of research, with many subfields and connections to other NLP tasks. In this book, we review the developments in this field against the backdrop of Elias Page's seminal 1966 paper titled "The Imminence of Grading Essays by Computer." Part 1 establishes what automated essay scoring is about, why it exists, where the technology stands, and what are some of the main issues. In Part 2, the book presents guided exercises to illustrate how one would go about building and evaluating a simple automated scoring system, while Part 3 offers readers a survey of the literature on different types of scoring models, the aspects of essay quality studied in prior research, and the implementation and evaluation of a scoring engine. Part 4 offers a broader view of the field inclusive of some neighboring areas, and Part \ref{part5} closes with summary and discussion. This book grew out of a week-long course on automated evaluation of language production at the North American Summer School for Logic, Language, and Information (NASSLLI), attended by advanced undergraduates and early-stage graduate students from a variety of disciplines. Teachers of natural language processing, in particular, will find that the book offers a useful foundation for a supplemental module on automated scoring. Professionals and students in linguistics, applied linguistics, educational technology, and other related disciplines will also find the material here useful.

Introduction to Chinese Natural Language Processing

Morgan & Claypool Publishers **This book introduces Chinese language-processing issues and techniques to readers who already have a basic background in natural language processing (NLP). Since the major difference between Chinese and Western languages is at the word level, the book primarily focuses on Chinese morphological analysis and introduces the concept, structure, and interword semantics of Chinese words. The following topics are covered: a general introduction to Chinese NLP; Chinese characters, morphemes, and words and the characteristics of Chinese words that have to be considered in NLP applications; Chinese word segmentation; unknown word detection; word meaning and Chinese linguistic resources; interword semantics based on word collocation and NLP techniques for collocation extraction. Table of Contents: Introduction / Words in Chinese / Challenges in Chinese Morphological Processing / Chinese Word Segmentation / Unknown Word Identification / Word Meaning / Chinese Collocations / Automatic Chinese Collocation Extraction / Appendix / References / Author Biographies**

Natural Language Processing for Social Media Second Edition

Morgan & Claypool Publishers **In recent years, online social networking has revolutionized interpersonal communication. The newer research on language analysis in social media has been increasingly focusing on the latter's impact on our daily lives, both on a personal and a professional level. Natural language processing (NLP) is one of the most promising avenues for social media data processing. It is a scientific challenge to develop powerful methods and algorithms which extract relevant information from a large volume of data coming from multiple sources and languages in various formats or in free form. We discuss the challenges in analyzing social media texts in contrast with traditional documents. Research methods in information extraction, automatic categorization and clustering, automatic summarization and indexing, and statistical machine translation need to be adapted to a new kind of data. This book reviews the current research on NLP tools and methods for processing the non-traditional information from social media data that is available in large amounts (big data), and shows how innovative NLP approaches can integrate appropriate linguistic information in various fields such as social media monitoring, healthcare, business intelligence, industry, marketing, and security and defence. We review the existing evaluation metrics for NLP and social media applications, and the new efforts in evaluation campaigns or shared tasks on new datasets collected from social media. Such tasks are organized by the Association for Computational Linguistics (such as SemEval tasks) or by the National Institute of Standards and Technology via the Text REtrieval Conference (TREC) and the Text Analysis Conference (TAC). In the concluding chapter, we discuss the importance of this dynamic discipline and its great potential for NLP in the coming decade, in the context of changes in mobile technology, cloud computing, virtual reality, and social networking. In this second edition, we have added information about recent progress in the tasks and applications presented in the first edition. We discuss new methods and their results. The number of research projects and publications that use social media data is constantly increasing due to continuously growing amounts of social media data and the need to automatically process them. We have added 85 new references to the more than 300 references from the first edition. Besides updating each section, we have added a new application (digital marketing) to the section on media monitoring and we have augmented the section on healthcare applications with an extended**

discussion of recent research on detecting signs of mental illness from social media.

Discourse Processing

Morgan & Claypool Publishers **Discourse Processing here is framed as marking up a text with structural descriptions on several levels, which can serve to support many language-processing or text-mining tasks. We first explore some ways of assigning structure on the document level: the logical document structure as determined by the layout of the text, its genre-specific content structure, and its breakdown into topical segments. Then the focus moves to phenomena of local coherence. We introduce the problem of coreference and look at methods for building chains of coreferring entities in the text. Next, the notion of coherence relation is introduced as the second important factor of local coherence. We study the role of connectives and other means of signaling such relations in text, and then return to the level of larger textual units, where tree or graph structures can be ascribed by recursively assigning coherence relations. Taken together, these descriptions can inform text summarization, information extraction, discourse-aware sentiment analysis, question answering, and the like. Table of Contents: Introduction / Large Discourse Units and Topics / Coreference Resolution / Small Discourse Units and Coherence Relations / Summary: Text Structure on Multiple Interacting Levels**

Linguistic Structure Prediction

Morgan & Claypool Publishers **A major part of natural language processing now depends on the use of text data to build linguistic analyzers. We consider statistical, computational approaches to modeling linguistic structure. We seek to unify across many approaches and many kinds of linguistic structures. Assuming a basic understanding of natural language processing and/or machine learning, we seek to bridge the gap between the two fields. Approaches to decoding (i.e., carrying out linguistic structure prediction) and supervised and unsupervised learning of models that predict discrete structures as outputs are the focus. We also survey natural language processing problems to which these methods are being applied, and we address related topics in probabilistic inference, optimization, and experimental methodology. Table of Contents: Representations and Linguistic Data / Decoding: Making Predictions / Learning Structure from Annotated Data / Learning Structure from Incomplete Data / Beyond Decoding: Inference**

Spoken Dialogue Systems

Morgan & Claypool Publishers **Considerable progress has been made in recent years in the development of dialogue systems that support robust and efficient human-machine interaction using spoken language. Spoken dialogue technology allows various interactive applications to be built and used for practical purposes, and research focuses on issues that aim to increase the system's communicative competence by including aspects of error correction, cooperation, multimodality, and adaptation in context. This book gives a comprehensive view of state-of-the-art techniques that are used to build spoken dialogue systems. It provides an overview of the basic issues such as system architectures, various dialogue management methods, system evaluation, and also surveys advanced topics concerning extensions of the basic model to more conversational setups. The goal of the book is to provide an introduction to the methods, problems, and solutions that are used in dialogue system development and evaluation. It presents dialogue modelling and system development issues relevant in both academic and industrial environments and also discusses requirements and challenges for advanced interaction management and future research. Table of Contents: Preface / Introduction to Spoken Dialogue Systems / Dialogue Management / Error Handling / Case Studies: Advanced Approaches to Dialogue Management / Advanced Issues / Methodologies and Practices of Evaluation / Future Directions / References / Author Biographies**

Deep Learning in Natural Language Processing

Springer **In recent years, deep learning has fundamentally changed the landscapes of a number of areas in artificial intelligence, including speech, vision, natural language, robotics, and game playing. In particular, the striking success of deep learning in a wide variety of natural language processing (NLP) applications has served as a benchmark for the advances in one of the most important tasks in artificial intelligence. This book reviews the state of the art of deep learning research and its successful applications to major NLP tasks, including speech recognition and understanding, dialogue systems, lexical analysis, parsing, knowledge graphs, machine translation, question answering, sentiment analysis, social computing, and natural language generation from images. Outlining and analyzing various research frontiers of NLP in the deep learning era, it features self-contained, comprehensive chapters written by leading researchers in the field. A glossary of technical terms and commonly used acronyms in the intersection of deep**

learning and NLP is also provided. The book appeals to advanced undergraduate and graduate students, post-doctoral researchers, lecturers and industrial researchers, as well as anyone interested in deep learning and natural language processing.

Dependency Parsing

Morgan & Claypool Publishers **Dependency-based methods for syntactic parsing have become increasingly popular in natural language processing in recent years. This book gives a thorough introduction to the methods that are most widely used today. After an introduction to dependency grammar and dependency parsing, followed by a formal characterization of the dependency parsing problem, the book surveys the three major classes of parsing models that are in current use: transition-based, graph-based, and grammar-based models. It continues with a chapter on evaluation and one on the comparison of different methods, and it closes with a few words on current trends and future prospects of dependency parsing. The book presupposes a knowledge of basic concepts in linguistics and computer science, as well as some knowledge of parsing methods for constituency-based representations. Table of Contents: Introduction / Dependency Parsing / Transition-Based Parsing / Graph-Based Parsing / Grammar-Based Parsing / Evaluation / Comparison / Final Thoughts**

Introduction to Natural Language Processing

MIT Press **A survey of computational methods for understanding, generating, and manipulating human language, which offers a synthesis of classical representations and algorithms with contemporary machine learning techniques. This textbook provides a technical perspective on natural language processing—methods for building computer software that understands, generates, and manipulates human language. It emphasizes contemporary data-driven approaches, focusing on techniques from supervised and unsupervised machine learning. The first section establishes a foundation in machine learning by building a set of tools that will be used throughout the book and applying them to word-based textual analysis. The second section introduces structured representations of language, including sequences, trees, and graphs. The third section explores different approaches to the representation and analysis of linguistic meaning, ranging from formal logic to neural word embeddings. The final section offers chapter-length treatments of three transformative applications of natural language processing: information extraction, machine translation, and text**

generation. End-of-chapter exercises include both paper-and-pencil analysis and software implementation. The text synthesizes and distills a broad and diverse research literature, linking contemporary machine learning techniques with the field's linguistic and computational foundations. It is suitable for use in advanced undergraduate and graduate-level courses and as a reference for software engineers and data scientists. Readers should have a background in computer programming and college-level mathematics. After mastering the material presented, students will have the technical skill to build and analyze novel natural language processing systems and to understand the latest research in the field.

Advances in Artificial Intelligence

29th Canadian Conference on Artificial Intelligence, Canadian AI 2016, Victoria, BC, Canada, May 31 - June 3, 2016. Proceedings

Springer **This book constitutes the refereed proceedings of the 29th Canadian Conference on Artificial Intelligence, Canadian AI 2016, held in Victoria, BC, Canada, in May/June 2016. The 12 full papers and 27 short papers presented were carefully reviewed and selected from 97 submissions. The focus of the conference was on the following subjects: actions and behaviours, audio and visual recognition, natural language processing, reasoning and learning, streams and distributed computing.**

Natural Language Processing for Historical Texts

Springer Nature **More and more historical texts are becoming available in digital form. Digitization of paper documents is motivated by the aim of preserving cultural heritage and making it more accessible, both to laypeople and scholars. As digital images cannot be searched for text, digitization projects increasingly strive to create digital text, which can be searched and otherwise automatically processed, in addition to facsimiles. Indeed, the emerging field of digital**

humanities heavily relies on the availability of digital text for its studies. Together with the increasing availability of historical texts in digital form, there is a growing interest in applying natural language processing (NLP) methods and tools to historical texts. However, the specific linguistic properties of historical texts -- the lack of standardized orthography, in particular -- pose special challenges for NLP. This book aims to give an introduction to NLP for historical texts and an overview of the state of the art in this field. The book starts with an overview of methods for the acquisition of historical texts (scanning and OCR), discusses text encoding and annotation schemes, and presents examples of corpora of historical texts in a variety of languages. The book then discusses specific methods, such as creating part-of-speech taggers for historical languages or handling spelling variation. A final chapter analyzes the relationship between NLP and the digital humanities. Certain recently emerging textual genres, such as SMS, social media, and chat messages, or newsgroup and forum postings share a number of properties with historical texts, for example, nonstandard orthography and grammar, and profuse use of abbreviations. The methods and techniques required for the effective processing of historical texts are thus also of interest for research in other domains. Table of Contents: Introduction / NLP and Digital Humanities / Spelling in Historical Texts / Acquiring Historical Texts / Text Encoding and Annotation Schemes / Handling Spelling Variation / NLP Tools for Historical Languages / Historical Corpora / Conclusion / Bibliography

Sentiment Analysis and Opinion Mining

Springer Nature **Sentiment analysis and opinion mining is the field of study that analyzes people's opinions, sentiments, evaluations, attitudes, and emotions from written language. It is one of the most active research areas in natural language processing and is also widely studied in data mining, Web mining, and text mining. In fact, this research has spread outside of computer science to the management sciences and social sciences due to its importance to business and society as a whole. The growing importance of sentiment analysis coincides with the growth of social media such as reviews, forum discussions, blogs, micro-blogs, Twitter, and social networks. For the first time in human history, we now have a huge volume of opinionated data recorded in digital form for analysis. Sentiment analysis systems are being applied in almost every business and social domain because opinions are central to almost all human activities and are key influencers of our behaviors. Our beliefs and perceptions of reality, and the choices we make, are largely conditioned on how others see and evaluate the world. For this reason, when we need to make a decision we often seek out the opinions of others. This is true not only for individuals but also for organizations. This book is a**

comprehensive introductory and survey text. It covers all important topics and the latest developments in the field with over 400 references. It is suitable for students, researchers and practitioners who are interested in social media analysis in general and sentiment analysis in particular. Lecturers can readily use it in class for courses on natural language processing, social media analysis, text mining, and data mining. Lecture slides are also available online. Table of Contents: Preface / Sentiment Analysis: A Fascinating Problem / The Problem of Sentiment Analysis / Document Sentiment Classification / Sentence Subjectivity and Sentiment Classification / Aspect-Based Sentiment Analysis / Sentiment Lexicon Generation / Opinion Summarization / Analysis of Comparative Opinions / Opinion Search and Retrieval / Opinion Spam Detection / Quality of Reviews / Concluding Remarks / Bibliography / Author Biography

Natural Language Processing

A Machine Learning Perspective

Cambridge University Press **This undergraduate textbook introduces essential machine learning concepts in NLP in a unified and gentle mathematical framework.**

Cross-Language Information Retrieval

Springer Nature **Search for information is no longer exclusively limited within the native language of the user, but is more and more extended to other languages. This gives rise to the problem of cross-language information retrieval (CLIR), whose goal is to find relevant information written in a different language to a query. In addition to the problems of monolingual information retrieval (IR), translation is the key problem in CLIR: one should translate either the query or the documents from a language to another. However, this translation problem is not identical to full-text machine translation (MT): the goal is not to produce a human-readable translation, but a translation suitable for finding relevant documents. Specific translation methods are thus required. The goal of this book is to provide a comprehensive description of the specific problems arising in CLIR, the solutions proposed in this area, as well as the remaining problems. The book starts with a general description of the monolingual IR and CLIR problems. Different classes of approaches to translation are then presented: approaches using an MT system, dictionary-based translation**

and approaches based on parallel and comparable corpora. In addition, the typical retrieval effectiveness using different approaches is compared. It will be shown that translation approaches specifically designed for CLIR can rival and outperform high-quality MT systems. Finally, the book offers a look into the future that draws a strong parallel between query expansion in monolingual IR and query translation in CLIR, suggesting that many approaches developed in monolingual IR can be adapted to CLIR. The book can be used as an introduction to CLIR. Advanced readers can also find more technical details and discussions about the remaining research challenges in the future. It is suitable to new researchers who intend to carry out research on CLIR. Table of Contents: Preface / Introduction / Using Manually Constructed Translation Systems and Resources for CLIR / Translation Based on Parallel and Comparable Corpora / Other Methods to Improve CLIR / A Look into the Future: Toward a Unified View of Monolingual IR and CLIR? / References / Author Biography

Discourse Processing

Springer Nature **Discourse Processing here is framed as marking up a text with structural descriptions on several levels, which can serve to support many language-processing or text-mining tasks. We first explore some ways of assigning structure on the document level: the logical document structure as determined by the layout of the text, its genre-specific content structure, and its breakdown into topical segments. Then the focus moves to phenomena of local coherence. We introduce the problem of coreference and look at methods for building chains of coreferring entities in the text. Next, the notion of coherence relation is introduced as the second important factor of local coherence. We study the role of connectives and other means of signaling such relations in text, and then return to the level of larger textual units, where tree or graph structures can be ascribed by recursively assigning coherence relations. Taken together, these descriptions can inform text summarization, information extraction, discourse-aware sentiment analysis, question answering, and the like. Table of Contents: Introduction / Large Discourse Units and Topics / Coreference Resolution / Small Discourse Units and Coherence Relations / Summary: Text Structure on Multiple Interacting Levels**

Natural Language Processing for Global and Local Business

IGI Global **The concept of natural language processing has become one of the preferred methods to better understand consumers, especially in recent years when digital technologies and research methods have developed exponentially. It has become apparent that when responding to international consumers through multiple platforms and speaking in the same language in which the consumers express themselves, companies are improving their standings within the public sphere. Natural Language Processing for Global and Local Business provides research exploring the theoretical and practical phenomenon of natural language processing through different languages and platforms in terms of today's conditions. Featuring coverage on a broad range of topics such as computational linguistics, information engineering, and translation technology, this book is ideally designed for IT specialists, academics, researchers, students, and business professionals seeking current research on improving and understanding the consumer experience.**

Semantic Role Labeling

Morgan & Claypool Publishers **This book is aimed at providing an overview of several aspects of semantic role labeling. Chapter 1 begins with linguistic background on the definition of semantic roles and the controversies surrounding them. Chapter 2 describes how the theories have led to structured lexicons such as FrameNet, VerbNet and the PropBank Frame Files that in turn provide the basis for large scale semantic annotation of corpora. This data has facilitated the development of automatic semantic role labeling systems based on supervised machine learning techniques. Chapter 3 presents the general principles of applying both supervised and unsupervised machine learning to this task, with a description of the standard stages and feature choices, as well as giving details of several specific systems. Recent advances include the use of joint inference to take advantage of context sensitivities, and attempts to improve performance by closer integration of the syntactic parsing task with semantic role labeling. Chapter 3 also discusses the impact the granularity of the semantic roles has on system performance. Having outlined the basic approach with respect to English, Chapter 4 goes on to discuss applying the same techniques to other languages, using**

Chinese as the primary example. Although substantial training data is available for Chinese, this is not the case for many other languages, and techniques for projecting English role labels onto parallel corpora are also presented. Table of Contents: Preface / Semantic Roles / Available Lexical Resources / Machine Learning for Semantic Role Labeling / A Cross-Lingual Perspective / Summary

Linguistic Fundamentals for Natural Language Processing II

100 Essentials from Semantics and Pragmatics

Springer Nature **Meaning is a fundamental concept in Natural Language Processing (NLP), in the tasks of both Natural Language Understanding (NLU) and Natural Language Generation (NLG). This is because the aims of these fields are to build systems that understand what people mean when they speak or write, and that can produce linguistic strings that successfully express to people the intended content. In order for NLP to scale beyond partial, task-specific solutions, researchers in these fields must be informed by what is known about how humans use language to express and understand communicative intents. The purpose of this book is to present a selection of useful information about semantics and pragmatics, as understood in linguistics, in a way that's accessible to and useful for NLP practitioners with minimal (or even no) prior training in linguistics.**

Finite-State Text Processing

Morgan & Claypool Publishers **Weighted finite-state transducers (WFSTs) are commonly used by engineers and computational linguists for processing and generating speech and text. This book first provides a detailed introduction to this formalism. It then introduces Pynini, a Python library for compiling finite-state grammars and for combining, optimizing, applying, and searching finite-state transducers. This book illustrates this library's conventions and use with a series of case studies. These include the compilation and application of context-dependent rewrite rules, the construction of morphological analyzers and generators, and text generation and processing applications.**

Embeddings in Natural Language Processing

Theory and Advances in Vector Representations of Meaning

Morgan & Claypool Publishers **Embeddings have undoubtedly been one of the most influential research areas in Natural Language Processing (NLP). Encoding information into a low-dimensional vector representation, which is easily integrable in modern machine learning models, has played a central role in the development of NLP. Embedding techniques initially focused on words, but the attention soon started to shift to other forms: from graph structures, such as knowledge bases, to other types of textual content, such as sentences and documents. This book provides a high-level synthesis of the main embedding techniques in NLP, in the broad sense. The book starts by explaining conventional word vector space models and word embeddings (e.g., Word2Vec and GloVe) and then moves to other types of embeddings, such as word sense, sentence and document, and graph embeddings. The book also provides an overview of recent developments in contextualized representations (e.g., ELMo and BERT) and explains their potential in NLP. Throughout the book, the reader can find both essential information for understanding a certain topic from scratch and a broad overview of the most successful techniques developed in the literature.**

Artificial Intelligence in Daily Life

Springer Nature **Given the exponential growth of Artificial Intelligence (AI) over the past few decades, AI and its related applications have become part of daily life in ways that we could never have dreamt of only a century ago. Our routines have been changed beyond measure by robotics and AI, which are now used in a vast array of services. Though AI is still in its infancy, we have already benefited immensely. This book introduces readers to basic Artificial Intelligence concepts, and helps them understand the relationship between AI and daily life. In the interest of clarity, the content is divided into four major parts. Part I (AI Concepts) presents fundamental concepts of and information on AI; while Part II (AI Technology) introduces readers to the five core AI Technologies that provide the building blocks for**

various AI applications, namely: Machine Learning (ML), Data Mining (DM), Computer Vision (CV), Natural Languages Processing (NLP), and Ontology-based Search Engine (OSE). In turn, Part III (AI Applications) reviews major contemporary applications that are impacting our ways of life, working styles and environment, ranging from intelligent agents and robotics to smart campus and smart city projects. Lastly, Part IV (Beyond AI) addresses related topics that are vital to the future development of AI. It also discusses a number of critical issues, such as AI ethics and privacy, the development of a conscious mind, and autonomous robotics in our daily lives.

Spoken Dialogue Systems

Springer Nature **Considerable progress has been made in recent years in the development of dialogue systems that support robust and efficient human-machine interaction using spoken language. Spoken dialogue technology allows various interactive applications to be built and used for practical purposes, and research focuses on issues that aim to increase the system's communicative competence by including aspects of error correction, cooperation, multimodality, and adaptation in context. This book gives a comprehensive view of state-of-the-art techniques that are used to build spoken dialogue systems. It provides an overview of the basic issues such as system architectures, various dialogue management methods, system evaluation, and also surveys advanced topics concerning extensions of the basic model to more conversational setups. The goal of the book is to provide an introduction to the methods, problems, and solutions that are used in dialogue system development and evaluation. It presents dialogue modelling and system development issues relevant in both academic and industrial environments and also discusses requirements and challenges for advanced interaction management and future research. Table of Contents: Preface / Introduction to Spoken Dialogue Systems / Dialogue Management / Error Handling / Case Studies: Advanced Approaches to Dialogue Management / Advanced Issues / Methodologies and Practices of Evaluation / Future Directions / References / Author Biographies**

Computational Modeling of Narrative

Morgan & Claypool Publishers **The field of narrative (or story) understanding and generation is one of the oldest in natural language processing (NLP) and artificial intelligence (AI), which is hardly surprising, since storytelling is such a fundamental and familiar intellectual and social activity. In recent years, the demands of interactive entertainment and**

interest in the creation of engaging narratives with life-like characters have provided a fresh impetus to this field. This book provides an overview of the principal problems, approaches, and challenges faced today in modeling the narrative structure of stories. The book introduces classical narratological concepts from literary theory and their mapping to computational approaches. It demonstrates how research in AI and NLP has modeled character goals, causality, and time using formalisms from planning, case-based reasoning, and temporal reasoning, and discusses fundamental limitations in such approaches. It proposes new representations for embedded narratives and fictional entities, for assessing the pace of a narrative, and offers an empirical theory of audience response. These notions are incorporated into an annotation scheme called NarrativeML. The book identifies key issues that need to be addressed, including annotation methods for long literary narratives, the representation of modality and habituality, and characterizing the goals of narrators. It also suggests a future characterized by advanced text mining of narrative structure from large-scale corpora and the development of a variety of useful authoring aids. This is the first book to provide a systematic foundation that integrates together narratology, AI, and computational linguistics. It can serve as a narratology primer for computer scientists and an elucidation of computational narratology for literary theorists. It is written in a highly accessible manner and is intended for use by a broad scientific audience that includes linguists (computational and formal semanticists), AI researchers, cognitive scientists, computer scientists, game developers, and narrative theorists.

Recognizing Textual Entailment

Models and Applications

Morgan & Claypool Publishers **In the last few years, a number of NLP researchers have developed and participated in the task of Recognizing Textual Entailment (RTE). This task encapsulates Natural Language Understanding capabilities within a very simple interface: recognizing when the meaning of a text snippet is contained in the meaning of a second piece of text. This simple abstraction of an exceedingly complex problem has broad appeal partly because it can be conceived also as a component in other NLP applications, from Machine Translation to Semantic Search to Information Extraction. It also avoids commitment to any specific meaning representation and reasoning framework, broadening its appeal within the research community. This level of abstraction also facilitates evaluation, a crucial component of any**

technological advancement program. This book explains the RTE task formulation adopted by the NLP research community, and gives a clear overview of research in this area. It draws out commonalities in this research, detailing the intuitions behind dominant approaches and their theoretical underpinnings. This book has been written with a wide audience in mind, but is intended to inform all readers about the state of the art in this fascinating field, to give a clear understanding of the principles underlying RTE research to date, and to highlight the short- and long-term research goals that will advance this technology.

Natural Language Processing for Social Media, Third Edition

Springer Nature **In recent years, online social networking has revolutionized interpersonal communication. The newer research on language analysis in social media has been increasingly focusing on the latter's impact on our daily lives, both on a personal and a professional level. Natural language processing (NLP) is one of the most promising avenues for social media data processing. It is a scientific challenge to develop powerful methods and algorithms that extract relevant information from a large volume of data coming from multiple sources and languages in various formats or in free form. This book will discuss the challenges in analyzing social media texts in contrast with traditional documents. Research methods in information extraction, automatic categorization and clustering, automatic summarization and indexing, and statistical machine translation need to be adapted to a new kind of data. This book reviews the current research on NLP tools and methods for processing the non-traditional information from social media data that is available in large amounts, and it shows how innovative NLP approaches can integrate appropriate linguistic information in various fields such as social media monitoring, health care, and business intelligence. The book further covers the existing evaluation metrics for NLP and social media applications and the new efforts in evaluation campaigns or shared tasks on new datasets collected from social media. Such tasks are organized by the Association for Computational Linguistics (such as SemEval tasks), the National Institute of Standards and Technology via the Text REtrieval Conference (TREC) and the Text Analysis Conference (TAC), or the Conference and Labs of the Evaluation Forum (CLEF). In this third edition of the book, the authors added information about recent progress in NLP for social media applications, including more about the modern techniques provided by deep neural networks (DNNs) for modeling language and analyzing social media data.**

Bitext Alignment

Morgan & Claypool Publishers **This book provides an overview of various techniques for the alignment of bitexts. It describes general concepts and strategies that can be applied to map corresponding parts in parallel documents on various levels of granularity. Bitexts are valuable linguistic resources for many different research fields and practical applications. The most predominant application is machine translation, in particular, statistical machine translation. However, there are various other threads that can be followed which may be supported by the rich linguistic knowledge implicitly stored in parallel resources. Bitexts have been explored in lexicography, word sense disambiguation, terminology extraction, computer-aided language learning and translation studies to name just a few. The book covers the essential tasks that have to be carried out when building parallel corpora starting from the collection of translated documents up to sub-sentential alignments. In particular, it describes various approaches to document alignment, sentence alignment, word alignment and tree structure alignment. It also includes a list of resources and a comprehensive review of the literature on alignment techniques. Table of Contents: Introduction / Basic Concepts and Terminology / Building Parallel Corpora / Sentence Alignment / Word Alignment / Phrase and Tree Alignment / Concluding Remarks**

Argumentation Mining

Morgan & Claypool Publishers **Argumentation mining is an application of natural language processing (NLP) that emerged a few years ago and has recently enjoyed considerable popularity, as demonstrated by a series of international workshops and by a rising number of publications at the major conferences and journals of the field. Its goals are to identify argumentation in text or dialogue; to construct representations of the constellation of claims, supporting and attacking moves (in different levels of detail); and to characterize the patterns of reasoning that appear to license the argumentation. Furthermore, recent work also addresses the difficult tasks of evaluating the persuasiveness and quality of arguments. Some of the linguistic genres that are being studied include legal text, student essays, political discourse and debate, newspaper editorials, scientific writing, and others. The book starts with a discussion of the linguistic perspective, characteristics of argumentative language, and their relationship to certain other notions such as subjectivity. Besides the connection to linguistics, argumentation has for a long time been a topic in Artificial**

Intelligence, where the focus is on devising adequate representations and reasoning formalisms that capture the properties of argumentative exchange. It is generally very difficult to connect the two realms of reasoning and text analysis, but we are convinced that it should be attempted in the long term, and therefore we also touch upon some fundamentals of reasoning approaches. Then the book turns to its focus, the computational side of mining argumentation in text. We first introduce a number of annotated corpora that have been used in the research. From the NLP perspective, argumentation mining shares subtasks with research fields such as subjectivity and sentiment analysis, semantic relation extraction, and discourse parsing. Therefore, many technical approaches are being borrowed from those (and other) fields. We break argumentation mining into a series of subtasks, starting with the preparatory steps of classifying text as argumentative (or not) and segmenting it into elementary units. Then, central steps are the automatic identification of claims, and finding statements that support or oppose the claim. For certain applications, it is also of interest to compute a full structure of an argumentative constellation of statements. Next, we discuss a few steps that try to 'dig deeper': to infer the underlying reasoning pattern for a textual argument, to reconstruct unstated premises (so-called 'enthymemes'), and to evaluate the quality of the argumentation. We also take a brief look at 'the other side' of mining, i.e., the generation or synthesis of argumentative text. The book finishes with a summary of the argumentation mining tasks, a sketch of potential applications, and a--necessarily subjective--outlook for the field.

Recent Trends in Image Processing and Pattern Recognition

Third International Conference, RTIP2R 2020, Aurangabad, India, January 3–4, 2020, Revised Selected

Papers, Part I

Springer Nature **This two-volume set constitutes the refereed proceedings of the Third International Conference on Recent Trends in Image Processing and Pattern Recognition (RTIP2R) 2020, held in Aurangabad, India, in January 2020. The 78 revised full papers presented were carefully reviewed and selected from 329 submissions. The papers are organized in topical sections in the two volumes. Part I: Computer vision and applications; Data science and machine learning; Document understanding and Recognition. Part II: Healthcare informatics and medical imaging; Image analysis and recognition; Signal processing and pattern recognition; Image and signal processing in Agriculture.**

Research Anthology on Implementing Sentiment Analysis Across Multiple Disciplines

IGI Global **The rise of internet and social media usage in the past couple of decades has presented a very useful tool for many different industries and fields to utilize. With much of the world's population writing their opinions on various products and services in public online forums, industries can collect this data through various computational tools and methods. These tools and methods, however, are still being perfected in both collection and implementation. Sentiment analysis can be used for many different industries and for many different purposes, which could better business performance and even society. The Research Anthology on Implementing Sentiment Analysis Across Multiple Disciplines discusses the tools, methodologies, applications, and implementation of sentiment analysis across various disciplines and industries such as the pharmaceutical industry, government, and the tourism industry. It further presents emerging technologies and developments within the field of sentiment analysis and opinion mining. Covering topics such as electronic word of mouth (eWOM), public security, and user similarity, this major reference work is a comprehensive resource for computer scientists, IT professionals, AI scientists, business leaders and managers, marketers, advertising agencies, public administrators, government officials, university administrators, libraries, students and faculty of higher education, researchers, and academicians.**

Neural Network Methods in Natural Language Processing

Morgan & Claypool Publishers **Neural networks are a family of powerful machine learning models and this book focuses on their application to natural language data. The first half of the book (Parts I and II) covers the basics of supervised machine learning and feed-forward neural networks, the basics of working with machine learning over language data, and the use of vector-based rather than symbolic representations for words. It also covers the computation-graph abstraction, which allows to easily define and train arbitrary neural networks, and is the basis behind the design of contemporary neural network software libraries. The second part of the book (Parts III and IV) introduces more specialized neural network architectures, including 1D convolutional neural networks, recurrent neural networks, conditioned-generation models, and attention-based models. These architectures and techniques are the driving force behind state-of-the-art algorithms for machine translation, syntactic parsing, and many other applications. Finally, we also discuss tree-shaped networks, structured prediction, and the prospects of multi-task learning.**

Advanced Machine Learning Technologies and Applications

First International Conference, AMLTA 2012, Cairo, Egypt, December 8-10, 2012, Proceedings

Springer **This book constitutes the refereed proceedings of the First International Conference on Advanced Machine Learning Technologies and Applications, AMLTA 2012, held in Cairo, Egypt, in December 2012. The 58 full papers presented were carefully reviewed and selected from 99 intial submissions. The papers are organized in topical sections on rough sets and applications, machine learning in pattern recognition and image processing, machine learning in multimedia computing, bioinformatics and cheminformatics, data classification and clustering, cloud computing and recommender systems.**

Language Technologies for the Challenges of the Digital Age

27th International Conference, GSCL 2017, Berlin, Germany, September 13-14, 2017, Proceedings

Springer **This open access volume constitutes the refereed proceedings of the 27th biennial conference of the German Society for Computational Linguistics and Language Technology, GSCL 2017, held in Berlin, Germany, in September 2017, which focused on language technologies for the digital age. The 16 full papers and 10 short papers included in the proceedings were carefully selected from 36 submissions. Topics covered include text processing of the German language, online media and online content, semantics and reasoning, sentiment analysis, and semantic web description languages.**