

1. Record Nr.	UNINA9910872197803321
Autore	Nugues Pierre M
Titolo	Python for Natural Language Processing : Programming with NumPy, scikit-learn, Keras, and PyTorch / / by Pierre M. Nugues
Pubbl/distr/stampa	Cham : , : Springer Nature Switzerland : , : Imprint : Springer, , 2024
ISBN	9783031575495
Edizione	[3rd ed. 2024.]
Descrizione fisica	1 online resource (536 pages)
Collana	Cognitive Technologies, , 2197-6635
Disciplina	006.35 000
Soggetti	Natural language processing (Computer science) Computational linguistics Python (Computer program language) Artificial intelligence User interfaces (Computer systems) Human-computer interaction Natural Language Processing (NLP) Computational Linguistics Python Artificial Intelligence User Interfaces and Human Computer Interaction
Lingua di pubblicazione	Inglese
Formato	Materiale a stampa
Livello bibliografico	Monografia
Nota di contenuto	Preface to the third edition -- Preface to the second edition -- Preface to the first edition -- 1. An Overview of Language Processing -- 2. A Tour of Python -- 3. Corpus Processing Tools -- 4. Encoding and Annotation Scheme -- 5. Python for Numerical Computations -- 6. Topics in Information Theory and Machine Learning -- 7. Linear and Logistic Regression -- 8. Neural Networks -- 9. Counting and Indexing Words -- 10. Dense Vector Representations -- 11. Word Sequences -- 12. Words, Parts of Speech, and Morphology -- 13. Subword Segmentation -- 14. Part-of-Speech and Sequence Annotation -- 15. Self-Attention and Transformers -- 16. Pretraining an Encoder: The BERT Language Model -- 17. Sequence-to-Sequence Architectures:

Sommario/riassunto

Since the last edition of this book (2014), progress has been astonishing in all areas of Natural Language Processing, with recent achievements in Text Generation that spurred a media interest going beyond the traditional academic circles. Text Processing has meanwhile become a mainstream industrial tool that is used, to various extents, by countless companies. As such, a revision of this book was deemed necessary to catch up with the recent breakthroughs, and the author discusses models and architectures that have been instrumental in the recent progress of Natural Language Processing. As in the first two editions, the intention is to expose the reader to the theories used in Natural Language Processing, and to programming examples that are essential for a deep understanding of the concepts. Although present in the previous two editions, Machine Learning is now even more pregnant, having replaced many of the earlier techniques to process text. Many new techniques build on the availability of text. Using Python notebooks, the reader will be able to load small corpora, format text, apply the models through executing pieces of code, gradually discover the theoretical parts by possibly modifying the code or the parameters, and traverse theories and concrete problems through a constant interaction between the user and the machine. The data sizes and hardware requirements are kept to a reasonable minimum so that a user can see instantly, or at least quickly, the results of most experiments on most machines. The book does not assume a deep knowledge of Python, and an introduction to this language aimed at Text Processing is given in Ch. 2, which will enable the reader to touch all the programming concepts, including NumPy arrays and PyTorch tensors as fundamental structures to represent and process numerical data in Python, or Keras for training Neural Networks to classify texts. Covering topics like Word Segmentation and Part-of-Speech and Sequence Annotation, the textbook also gives an in-depth overview of Transformers (for instance, BERT), Self-Attention and Sequence-to-Sequence Architectures. .
