****

**Please return to:**

*Dr. Ramesh Nath Premnath, Publishing Editor, Springer Science+Business Media Singapore Pvt Ltd, 152 Beach Road, #22-06/08 Gateway East, Singapore 189721*

*Phone +65 6396 6386, FAX +65 6298 8043/ email:* [*ramesh.premnath@springer.com*](mailto:ramesh.premnath@springer.com)

Book Proposal Form

The following questions are designed to help us get a clear picture of your book proposal and to provide us with information about the readership which will enable us to develop an effective marketing and promotion strategy. Please complete the form as fully as you can.

**Full or working title of the book:**

Neural Representations of Meaning

**Subtitle:**

Modern Developments in using machine learnt representations of natural language

|  |  |
| --- | --- |
| ***Name & Affiliation of each Author / Editor***  ***(as they should appear in the book)*** | ***Home and email Addresses (for royalty purposes)*** |
| Lyndon White School of Electrical, Electronic and Computer Engineering. The University of Western Australia | [lyndon.white@research.uwa.edu.au](mailto:lyndon.white@research.uwa.edu.au) 2A Alston Av, Como, Western Australia |
| Roberto Togneri  School of Electrical, Electronic and Computer Engineering. The University of Western Australia | roberto.togneri@uwa.edu.au  63 Ten Seldam Circle, Winthrop WA 6150, Australia |
| Wei Liu School of Computer Science and Software Engineering The University of Western Australia  Mohammed Bennamoun  School of Computer Science and Software Engineering The University of Western Australia | [wei.liu@uwa.edu.au](mailto:wei.liu@uwa.edu.au) 113A Shenton Rd, Swanbourne, WA 6010  mohammed.bennamoun@uwa.edu.au  33 Bankhurst Way, Greenwood, WA 6024, Australia. |

To apply for copyright and registration with the Library of Congress and other bibliographic services we also need the following information about each author/editor:

|  |  |  |  |  |
| --- | --- | --- | --- | --- |
| ***Last Name*** | ***First names*** | ***Date of birth*** | ***Nationality*** | ***Qualifications*** |
| White | Lyndon | 31 October 1990 | Australia | BCM, BE (hons) |
| Togneri | Roberto | 15 April 1962 | Australian | PhD |
| Liu | Wei | 14 July 1972 | Australian | PhD. |
| Bennamoun | Mohammed | 3 June 1961 | Australian | PhD. |

if possible, please attach a short curriculum vitae of each author/editor

**ABOUT YOUR BOOK**

**Please write a short text about your book (minimum of 200 words), incorporating answers to the following questions.** (What is the subject of your book? What methods, results, or topics will be of particular interest to the readers, and why? Are there any special features like illustrations, tables, a new form of presentation, or didactic approach? What main benefit will the reader derive from the book? This statement will serve as the basis for our promotional texts.)

This book will cover an introduction to modern natural languages processing via machine learning. It will focus on how neural networks create a machine interpretable representation of the meaning of natural language. Language is crucially linked to ideas: As Webster’s 1923 “English Composition and Literature” puts it: *“A sentence is a group of words expressing a complete thought”.* Thus the representation of sentences, the words that make them up is of crucial importance in current advancements in artificial intelligence and other “smart” systems being developed today.

The core focus of the book is to serve as an overview of works in the area, as it has progressed from Bengio et al’s seminal work on a “Neural Probabilistic Language Model” in 2003, to the latest techniques from this year. The book will connect these works together to allow the reader to understand how the techniques relate.

A special in the presentation of the relationship between works is to display the context as a directed graph. Each work (or technology) will be represented as a node, and it will have arrowed connecting earlier technologies. This motif will be repeated in the marginalia and/or section headings throughout the book, to allow the reader when glancing through, to quickly orientate themselves within the literature – understanding that for example, that the recursive autoencoder (RvAE) is derived from the work on autoencoders (AEs) and the work on recursive neural networks (RvNNs) which are in-turn an extension on recurrent neural networks (RNNs). This structured decomposition of the advancements in the area will allow the reader to become familiar with the state of the art.

One key distinguishing factor of the book will be the attention paid to practical implementations. For many academic works the implementation of the technique is lacking in maturity – which is understandable, for many works are a stepping stone to further advancements. The book will particularly highlight whether there are mature commercial or open source implementations of works discussed. It will also, use sidebars to alert the reader to key tips that they should take into account when implementing the reviewed techniques themselves – For example the importance of using indexing, rather than one-hot matrix multiplication to retrieve word embedding. These practical aspects, combined with the more traditional review of the literature will make this book more directly applicable to most readers.

**Please append an annotated table of contents of the work.** Additional information about the work such as relevant off-prints, a specimen chapter, etc. may be supplied separately.

**Part A: Introductory**

* Chapter 1: Introduction: 2-3 pages
  + Introduce the book, and the utility of using machine learning for natural language processing
* Chapter 2: Introduction to machine learning for representations (10 pages)
  + This chapter can be skipped by readers already familiar with machine learning
  + This will not be a full introduction to machine learning, which of-course could be an entire book on its own.
  + It will cover the crucial basic techniques used in the works discussed in part B. The main focus will be on introducing neural networks, going into reasonable detail on hyper-parameters such as activation functions and layer sizes, as well as covering training techniques.
  + It will not cover techniques that are special to natural language processing – those will be discussed in chapters 4,5, and 6.
* Chapter 3: Current Challenges in Natural Language Processing (10 pages)
  + Whereas chapter 2 introduces the Machine Learning domain, this chapter introduces the Natural Language Processing domain.
  + This chapter will cover the current tasks that natural language processing is being deployed for
  + It provide forward references to the works (in Part B) that use machine learnt representation to accomplish these tasks
  + Particular tasks to be discussed include
    - Language Modelling
    - Sentiment Analysis
    - Image Captioning
    - Image generation
    - Machine Translation
    - Paraphrase and plagiarism detection
  + Limited discussion will be presented here on prior, non-machine learning techniques.

**Part B: Representations**

* Chapter 4: Word Representations (10 pages)
  + The core work that began to draw a lot of attention to this area.
  + This will cover crucial works such as Skip-Grams, CBOW, GLoVe, and of course the original neural probabilistic language model.
  + As well as more recent techniques based on matrix factorisations.
* Chapter 5: Word Sense Representations (5-10 pages)
  + In this chapter, technologies for representing the multiple meanings of a single word can have will be discussed.
  + This is a growing area, and is particularly important in languages (including English, but other languages even more so), where polysemous and homonymous words are common.
  + It leads naturally to the next section on phrase representation. Rather than a single word having many meanings, the next chapter will discuss how a single meaning may take multiple words to express.
* Chapter 7: Phrase Representations (5-10 pages)
  + This will cover phrases,
  + Phrases range from:
    - multi-token words: for example: “et al”, “word sense”;
    - to collocations: “young adult”, “5 year old”
    - to longer sentence clauses: “the fast train”, “once upon a time”
* Chapter 8: Sentence representations and beyond (5-10 pages)
  + This chapter takes the previous discussion of phrases to the next level: sentences.
  + This will include discussions of works on recursive structure
  + As well work leveraging recurrent neural networks.
  + Methods that do not strongly consider order (including Sum of Word Embeddings; paragraph vectors) will also be discussed here.
  + Many of these techniques extent to arbitrary length sequences of words.
* Chapter 9: Character-Based Representations (5 pages)
  + This short chapter will discuss some of the recent works which directly modelling only the characters (letter), but using this to accomplish tasks from much larger structures.
  + Starting from Zhang and LeCun’s “Text Understanding from Scratch” (2015), and concluding with the very most recent works.
  + It will draw the book to a close by retouching on many of the tasks more commonly associated with prior sections and will discuss how they are attempted from a fully uninformed system that is learning only from letters.
  + This is a challenging area with fewer works to be discussed.

**Part C: Conclusion**

* Chapter 8: Conclusion (10 pages)
  + This will conclude the book.
  + It will summarise the prior charters
  + Discuss the progression of the field.
  + It will discuss the role machine learnt representations have within larger systems.
  + It will conclude with an outlook on the future.

**List at least three unique selling points (USP).** A USP is a compelling reason for why a customer should buy your book instead of a similar book. Ideally, each point will show how a key feature (e.g., 100 color photos) leads to a customer benefit (e.g., fast and reliable diagnosis). Sample USPs for a professional computer science book:

- Numerous step-by-step tutorials help the reader to learn quickly

- A special chapter on next generation Flash prepares readers for the future

- Includes ten tips on how to protect flash sites from hackers

* Unique Graphical Marginalia of Related works allows the reader to immediately locate themselves in the literature.
* Introductory Chapter on Machine Learning allows nonexperts to quickly get up to speed to understand how it is revolutionising natural language processing
* Includes clear breakdown of which works have available commercial or open-source implementations available today; to allowed readers to understand which works they can leverage today.
* Practical tips on implementation of the techniques provided in sidebars, saving the reader time and effort

**Keywords for Catalogue index and Search Engine Optimization**: Please give the 5 most important keywords in the language of publication.

**Natural Language Processing, Machine Learning, vector representations, word embeddings, Learned Representations**

*Many thanks for taking the time to complete this questionnaire. Your detailed responses will be very helpful in our assessment of the potential market for your proposal and will enable us to reach a decision regarding publication that much sooner.*