Deep Learning for Natural Language Processing

Deep learning is becoming increasingly important in a technology-dominated world. However, the building of computational models that accurately represent linguistic structures is complex, as it involves an in-depth knowledge of neural networks and the understanding of advanced mathematical concepts such as calculus and statistics. This book makes these complexities accessible to those from a humanities and social sciences background by providing a clear introduction to deep learning for natural language processing. It covers both theoretical and practical aspects and assumes minimal knowledge of machine learning, explaining the theory behind natural language in an easy-to-read way. It includes pseudo code for the simpler algorithms discussed and actual Python code for the more complicated architectures, using modern deep learning libraries such as PyTorch and Hugging Face. Providing the necessary theoretical foundation and practical tools, this book will enable readers to immediately begin building real-world, practical natural language processing systems.

MIHAI SURDEANU is an associate professor in the computer science department at the University of Arizona. He works in both academia and industry on natural language processing systems that process and extract meaning from natural language.

MARCO A. VALENZUELA-ESCÁRCEGA is a research scientist in the computer science department at the University of Arizona. He has worked on natural language processing projects in both industry and academia.

Deep Learning for Natural Language Processing

A Gentle Introduction

Mihai Surdeanu
University of Arizona

Marco Antonio Valenzuela-Escárcega
University of Arizona

CAMBRIDGE
UNIVERSITY PRESS

CAMBRIDGE UNIVERSITY PRESS

Shaftesbury Road, Cambridge CB2 8EA, United Kingdom

One Liberty Plaza, 20th Floor, New York, NY 10006, USA

477 Williamstown Road, Port Melbourne, VIC 3207, Australia

314–321, 3rd Floor, Plot 3, Splendor Forum, Jasola District Centre,
New Delhi – 110025, India

103 Penang Road, #05–06/07, Visioncrest Commercial, Singapore 238467

Cambridge University Press is part of Cambridge University Press & Assessment,
a department of the University of Cambridge.

We share the University's mission to contribute to society through the pursuit of
education, learning and research at the highest international levels of excellence.

www.cambridge.org
Information on this title: www.cambridge.org/9781316515662

DOI: 10.1017/9781009026222

First published 2024

A catalogue record for this publication is available from the British Library

A Cataloging-in-Publication data record for this book is available from the Library of Congress

ISBN 978-1-316-51566-2 Hardback
ISBN 978-1-009-01265-2 Paperback

Contents

Figures

Cambridge University Press & Assessment
978-1-316-51566-2 — Deep Learning for Natural Language Processing
Mihai Surdeanu, Marco Antonio Valenzuela-Escárcega
Frontmatter
More Information

Tables

Preface

Upon encountering this publication, one might ask the obvious question, "Why do we need another deep learning and natural language processing book?" Several excellent ones have been published, covering both theoretical and practical aspects of deep learning and its application to language processing. However, from our experience teaching courses on natural language processing, we argue that, despite their excellent quality, most of these books do not target their most likely readers. The intended reader of this book is one who is skilled in a domain other than machine learning and natural language processing and whose work relies, at least partially, on the automated analysis of large amounts of data, especially textual data. Such experts may include social scientists, political scientists, biomedical scientists, and even computer scientists and computational linguists with limited exposure to machine learning.

Existing deep learning and natural language processing books generally fall into two camps. The first camp focuses on the theoretical foundations of deep learning. This is certainly useful to the aforementioned readers, as one should understand the theoretical aspects of a tool before using it. However, these books tend to assume the typical background of a machine learning researcher and, as a consequence, we have often seen students who do not have this background rapidly get lost in such material. To mitigate this issue, the second type of book that exists today focuses on the machine learning practitioner – that is, on how to use deep learning software, with minimal attention paid to the theoretical aspects. We argue that focusing on practical aspects is similarly necessary but not sufficient. Considering that deep learning frameworks and libraries have become fairly complex, the chance of misusing them due to theoretical misunderstandings is high. We have commonly seen this problem in our courses too.

This book therefore aims to bridge the theoretical and practical aspects of deep learning for natural language processing. We cover the necessary theoretical background and assume minimal machine learning background from the reader. Our aim is that anyone who took introductory linear algebra and calculus courses will be able to follow the theoretical material. To address practical aspects, this book includes pseudocode for the simpler algorithms discussed and actual Python code for the more complicated architectures. The code should be understandable to anyone who has taken a Python programming course. After reading this book, we expect that the reader will have the necessary foundation to immediately begin building real-world, practical natural language processing systems, and to expand their knowledge by reading research publications on these topics.