

---

# *Modern Approaches in Natural Language Processing*



---

# *Contents*

---

Preface	v
Foreword	1
1 Introduction	3
2 Chapter 1	5
3 Introduction: Deep Learning for NLP	7
4 Foundations/Applications of Modern NLP	9
5 Recurrent neural networks and their applications in NLP	11
6 Convolutional neural networks and their applications in NLP	13
7 Introduction: Transfer Learning for NLP	15
8 Transfer Learning for NLP I	17
9 Attention and Self-Attention for NLP	19
10 Transfer Learning for NLP II	21
11 Introduction: Resources for NLP	23
12 Resources and Benchmarks for NLP	25
13 Software for NLP: The huggingface transformers module	27
14 Acknowledgements	29



---

## *Preface*

---

In the last few years, there have been several breakthroughs concerning the methodologies used in Natural Language Processing (NLP). These breakthroughs originate from both new modeling frameworks as well as from improvements in the availability of computational and lexical resources.

In this seminar, we are planning to review these frameworks starting with a methodology that can be seen as the beginning of modern NLP: Word Embeddings.

We will further discuss the integration of embeddings into end-to-end trainable approaches, namely convolutional and recurrent neural networks. As Attention-based models and transfer learning approaches are the foundation of most of the recent state-of-the-art models, we will cover these two topics extensively in the second part of our seminar.

We will furthermore talk about software implementations of these methods and benchmark tasks/data sets for evaluating state-of-the-art models.

This book is the outcome of the seminar “Modern Approaches in Natural Language Processing” which took place in summer 2020 at the Department of Statistics, LMU Munich.



**FIGURE 1:** Creative Commons License

This book is licensed under the Creative Commons Attribution-NonCommercial-ShareAlike 4.0 International License<sup>1</sup>.

---

<sup>1</sup><http://creativecommons.org/licenses/by-nc-sa/4.0/>



---

## ***Foreword***

---

*Author: Christoph Molnar*

This book is the result of an experiment in university teaching. Each semester, students of the Statistics Master can choose from a selection of seminar topics. Usually, every student in the seminar chooses a scientific paper, gives a talk about the paper and summarizes it in the form of a seminar paper. The supervisors help the students, they listen to the talks, read the seminar papers, grade the work and then ... hide the seminar papers away in (digital) drawers. This seemed wasteful to us, given the huge amount of effort the students usually invest in seminars. An idea was born: Why not create a book with a website as the outcome of the seminar? Something that will last at least a few years after the end of the semester. In the summer term 2020, some Statistics Master students signed up for our seminar entitled “Limitations of Interpretable Machine Learning”. When they came to the kick-off meeting, they had no idea that they would write a book by the end of the semester.

We were bound by the examination rules for conducting the seminar, but otherwise we could deviate from the traditional format. We deviated in several ways:

1. Each student project is part of a book, and not an isolated seminar paper.
2. We gave challenges to the students, instead of papers. The challenge was to investigate a specific limitation of interpretable machine learning methods.
3. We designed the work to live beyond the seminar.
4. We emphasized collaboration. Students wrote some chapters in teams and reviewed each others texts.

---

## **Technical Setup**

The book chapters are written in the Markdown language. The simulations, data examples and visualizations were created with R ([R Core Team, 2018](#)). To combine R-code and Markdown, we used rmarkdown. The book was compiled

with the bookdown package. We collaborated using git and github. For details, head over to the book's repository<sup>2</sup>.

---

<sup>2</sup>[link/to/repo](#)



# 1

---

## *Introduction*

---

*Author:*

*Supervisor:*

---

### 1.1 Intro About the Seminar Topic

---

### 1.2 Outline of the Booklet



# 2

## Chapter 1

*Authors: Author 1, Author 2*

*Supervisor: Supervisor*

### 2.1 Lorem Ipsum

Lorem ipsum dolor sit amet, consetetur sadipscing elitr, sed diam nonumy eirmod tempor invidunt ut labore et dolore magna aliquyam erat, sed diam voluptua. At vero eos et accusam et justo duo dolores et ea rebum. Stet clita kasd gubergren, no sea takimata sanctus est Lorem ipsum dolor sit amet. Lorem ipsum dolor sit amet, consetetur sadipscing elitr, sed diam nonumy eirmod tempor invidunt ut labore et dolore magna aliquyam erat, sed diam voluptua. At vero eos et accusam et justo duo dolores et ea rebum. Stet clita kasd gubergren, no sea takimata sanctus est Lorem ipsum dolor sit amet.

[R Core Team \(2018\)](#)

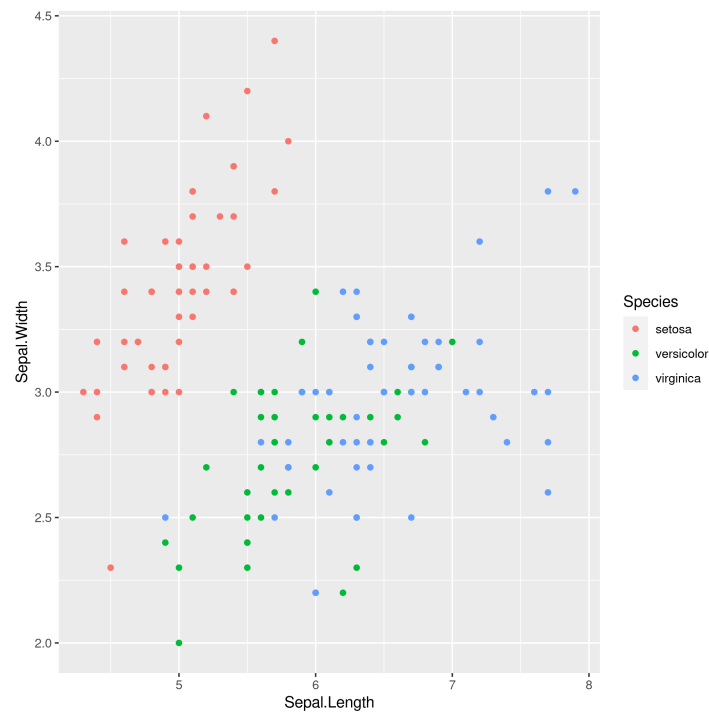
### 2.2 Using Figures

Referencing can be done by using the chunk label e.g. `\@ref(fig:ch01-figure01)` for [2.1](#).

**NOTE!!!** Do not use underscores in chunk labels! This will crash the compilation ...

### 2.3 Using Tex

HTML rendering uses MathJax while pdf rendering uses LaTeX:



**FIGURE 2.1:** This is the caption of the figure!

$$f(x) = x^2$$

---

## 2.4 Using Stored Results

	Estimate	Std. Error	t value	Pr(> t )
(Intercept)	2.1713	0.2798	7.760	0.0000
Sepal.Width	0.4959	0.0861	5.761	0.0000
Petal.Length	0.8292	0.0685	12.101	0.0000
Petal.Width	-0.3152	0.1512	-2.084	0.0389
Speciesversicolor	-0.7236	0.2402	-3.013	0.0031
Speciesvirginica	-1.0235	0.3337	-3.067	0.0026

# 3

---

## *Introduction: Deep Learning for NLP*

---

*Authors: Author 1, Author 2*

*Supervisor: Supervisor*



# 4

## *Foundations/Applications of Modern NLP*

*Authors: Author 1, Author 2*

*Supervisor: Supervisor*





# 5

---

## *Recurrent neural networks and their applications in NLP*

---

*Authors: Author 1, Author 2*

*Supervisor: Supervisor*



# 6

---

## *Convolutional neural networks and their applications in NLP*

---

*Authors: Author 1, Author 2*

*Supervisor: Supervisor*



# 7

---

## *Introduction: Transfer Learning for NLP*

---

*Authors: Author 1, Author 2*

*Supervisor: Supervisor*



# 8

## *Transfer Learning for NLP I*

*Authors: Author 1, Author 2*

*Supervisor: Supervisor*





# 9

## *Attention and Self-Attention for NLP*

*Authors: Author 1, Author 2*

*Supervisor: Supervisor*



# 10

---

## *Transfer Learning for NLP II*

---

*Authors: Author 1, Author 2*

*Supervisor: Supervisor*



# 11

---

## *Introduction: Resources for NLP*

---

*Authors: Author 1, Author 2*

*Supervisor: Supervisor*



# 12

---

## *Resources and Benchmarks for NLP*

---

*Authors: Author 1, Author 2*

*Supervisor: Supervisor*





# 13

---

## *Software for NLP: The huggingface transformers module*

---

*Authors: Author 1, Author 2*

*Supervisor: Supervisor*



# 14

---

## *Acknowledgements*

---

The most important contributions are from the students themselves. The success of such projects highly depends on the students. And this book is a success, so thanks a lot to all the authors! The other important role is the supervisor. Thanks to all the supervisors who participated! Special thanks to Christian Heumann<sup>1</sup> who enabled us to conduct the seminar in such an experimental way, supported us and gave valuable feedback for the seminar structure. Thanks a lot as well to the entire Department of Statistics<sup>2</sup> and the LMU Munich<sup>3</sup> for the infrastructure.

The authors of this work take full responsibilities for its content.

---

<sup>1</sup><https://www.misoda.statistik.uni-muenchen.de/personen/professoren/heumann/index.html>

<sup>2</sup><https://www.statistik.uni-muenchen.de/>

<sup>3</sup><http://www.en.uni-muenchen.de/index.html>



---

## ***Bibliography***

---

R Core Team (2018). *R: A Language and Environment for Statistical Computing*. R Foundation for Statistical Computing, Vienna, Austria.

