



Technische Universität Berlin

Quality and Usability Lab

Part-of-Speech Tagging  
with Neural Networks  
for a Conversational Agent

**Master Thesis**

Master of Science (M.Sc.)

**Author** Andreas Müller

**Major** Computer Engineering

**Matriculation No.** 333471

**Date** 18th May 2018

**1st supervisor** Prof. Dr.-Ing. Sebastian Möller

**2nd supervisor** Prof. Dr. ???



# Eidesstattliche Erklärung

Hiermit versichere ich, dass ich die vorliegende Arbeit selbstständig verfasst und keine anderen als die angegebenen Quellen und Hilfsmittel benutzt habe. Alle Ausführungen, die anderen veröffentlichten oder nicht veröffentlichten Schriften wörtlich oder sinngemäß entnommen wurden, habe ich kenntlich gemacht.

Die Arbeit hat in gleicher oder ähnlicher Fassung noch keiner anderen Prüfungsbehörde vorgelegen.

Berlin, den December 23, 2017

---

Unterschrift



# Abstract

...



# **Zusammenfassung**

...





# Contents

List of Figures	X
List of Tables	XI
Abbreviations	XII
1 Introduction	1
2 ALEX : Artificial Conversational Agent	2
2.1 System Overview . . . . .	2
2.2 Language Model . . . . .	2
2.3 Tagging Interface . . . . .	2
3 Evaluation and Comparison	3
4 Discussion and Conclusion	4
A First appendix	5
A.1 test . . . . .	5

# List of Figures

# List of Tables

# Abbreviations

<b>NLP</b>	<i>Natural Language Processing</i>
<b>Alex</b>	<i>Artificial Conversational Agent</i>
<b>HMM</b>	<i>Hidden Markov Model</i>
<b>FNN</b>	<i>(Feed-forward) Neural Network</i>
<b>RNN</b>	<i>Recurrent Neural Network</i>

# 1 Introduction

This is just a test

## 2 ALEX : Artificial Conversational Agent

...

### 2.1 System Overview

...

### 2.2 Language Model

...

### 2.3 Tagging Interface

The modular structure of ALEX allows for easier separation of various functions and therefore easier replaceability of certain functionalities. One of these modules is the tagger, which is used to train a language model on the one hand and to assign tags to the words of a given input sentence on the other hand.

The implementation of this tagger utilizes a Hidden Markov Model (HMM), which is a statistical model that is particularly used for pattern recognition, speech recognition and part-of-speech tagging. ALEX uses an already existing implementation of the HMM Tagger from the Natural Language Toolkit (NLTK)<sup>1</sup>, called `HiddenMarkovModelTagger`.

---

<sup>1</sup> The Natural Language Toolkit is a collection of *Python* programming libraries for natural language processing, see <http://nltk.org>

## 3 Evaluation and Comparison

...

## 4 Discussion and Conclusion

...



# **A First appendix**

## **A.1 test**

...