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 9, 2017

_____

Unterschrift

Abstract

...

Zusammenfassung

…

Contents

List of Figures

List of Tables

Abbreviations

NLP   *Natural Language Processing* (linguistische Datenverarbeitung)

IR   *Information Retrieval* (Informationsrückgewinnung)

SGD   *Stochastic Gradient Descent*

NNLM   *(Feedforward) Neural Net Language Model*

RNNLM   *Recurrent Neural Net Language Model*

CBOW   *Continous Bag-of-Words*

1 Introduction

This is just a test

2 Evaluation

…

3 Discussion and Conclusion

…

Bibliography

[1] Kheira Leila Arras, Jan Heyd, The-Anh Ly, and Andreas Müller. Exploring semantic word similarities in German News Articles. 2014.

[2] R.E. Bellman. *Adaptive control processes: A guided tour*. Princeton University Press (Princeton, NJ), 1961.

[3] R. Bender, A. Ziegler, and St. Lange. Logistische Regression. (14):12–14, 2002.

[4] Yoshua Bengio, Réjean Ducharme, Pascal Vincent, and Christian Jauvin. A Neural Probabilistic Language Model. *Journal of Machine Learning Research*, 3:1137–1155, 2003.

[5] Yoshua Bengio, Pascal Lamblin, Dan Popovici, and Hugo Larochelle. Greedy Layer-Wise Training of Deep Networks. *Advances in neural information processing systems*, 19(1):153, 2007.

[6] Peter F. Brown, Peter V. DeSouza, Robert L. Mercer, Vincent J. Della Pietra, and Jenifer C. Lai. Class-Based n-gram Models of Natural Language. *Computational Linguistics*, (1950), 1992.

[7] Michael Curtiss, Iain Becker, and Tudor Bosman. Unicorn: a system for searching the social graph. Technical report, 2013.

[8] Joshua Goodman. A Bit of Progress in Language Modeling. page 73, 2001.

[9] Zellig Harris. *Distributional Structure*, volume 1. 1954.

[10] J J Hopfield. Neural networks and physical systems with emergent collective computational abilities. *Proceedings of the National Academy of Sciences of the United States of America*, 79(8):2554–2558, 1982.

[11] H Kučera and W N Francis. *Computational Analysis of Present-Day American English*. Brown University Press, 1967.

[12] T Mikolov, M Karafiat, L Burget, J Cernocky, and S Khudanpur. Recurrent Neural Network based Language Model. *Interspeech*, (September):1045–1048, 2010.

[13] Tomas Mikolov, Kai Chen, Greg Corrado, and Jeffrey Dean. Distributed Representations of Words and Phrases and their Compositionality. pages 1–9, 2013.

[14] Tomas Mikolov, Greg Corrado, Kai Chen, and Jeffrey Dean. Efficient Estimation of Word Representations in Vector Space. *Proceedings of the International Conference on Learning Representations (ICLR 2013)*, pages 1–12, 2013.

[15] Tomas Mikolov, Wen-tau Yih, and Geoffrey Zweig. Linguistic Regularities in Continuous Space Word Representations. 2012.

Bibliography

[16] Zach Miners. Yahoo buys SkyPhrase to better understand natural language. *PCWorld*, 2013.

[17] Marvin Minsky and Seymour Papert. *Perceptron - An Essay in Computational Geometry*. MIT Press, 1969.

[18] Frederic Morin and Y Bengio. Hierarchical probabilistic neural network language model. *Proceedings of the Tenth International Workshop on Artificial Intelligence and Statistics*, pages 246–252, 2005.

[19] Karl Pearson. On lines and planes of closest fit to systems of points in space. *The London, Edinburgh, and Dublin Philosophical Magazine and Journal of Science*, 2:559–572, 1901.

[20] F Peng. Augmentating Naive Bayes Classifiers with Statistical Language Models. *Computer Science Department Faculty Publication Series*, Paper 91, 2003.

[21] Xin Rong. word2vec Parameter Learning Explained. pages 1–19.

[22] F Rosenblatt. The perceptron: a probabilistic model for information storage and organization in the brain. *Psychological review*, 65(6):386–408, 1958.

[23] David E. Rumelhart, Geoffrey E. Hinton, and Ronald J. Williams. Learning internal representations by error propagation, 1986.

[24] David Shapiro, Doug Platts, and Magico Martinez. Google hummingbird explained. *icrossing*, 2013.

[25] Richard Socher, Alex Perelygin, Jean Y Wu, Jason Chuang, Christopher D Manning, Andrew Y Ng, and Christopher Potts. Recursive Deep Models for Semantic Compositionality Over a Sentiment Treebank. 2013.

[26] Bernard Widrow. An Adaptive 'Adaline' Neuron Using Chemical 'Memistors', 1960.

[27] Daojian Zeng, Kang Liu, Siwei Lai, Guangyou Zhou, and Jun Zhao. Relation Classification via Convolutional Deep Neural Network. pages 1–10, 2014.

A First appendix

A.1 test

...