

Neural Networks: Architectures and Applications for NLP

Session 00: Organisatorisches

Julia Kreutzer & Julian Hitschler

25. Oktober 2016

Institut für Computerlinguistik, Heidelberg

1. Vorstellung
2. Kursaufbau
3. Scheinerwerb

Vorstellung

- Wer sind wir?
 - {kreutzer, hirschler} at cl.uni-heidelberg.de
 - Vorlesung, Referate: Julia
 - Übung, Projekte: Julian
 - Sprechstunde: Julia Di 16-17, Julian Do 11-12, R 107
- Wer seid ihr?
 - BA, MA, ...
 - CL, ...

Einführung in Architekturen und Anwendungen von Neuronalen Netzen für NLP

- Was sind (tiefe) Neuronale Netze?
- Warum können sie erfolgreich für NLP eingesetzt werden?
- Welche Architekturen eignen sich für welche Aufgaben?
- Welche Herausforderungen gibt es?

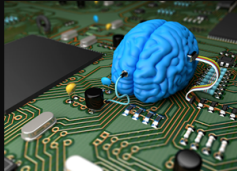
Nach diesem Kurs solltet ihr

- grundlegende Architekturen von Künstlichen Neuronalen Netzen kennen
- die Techniken zum Training dieser Netze verstanden haben
- (einfache) Modelle für NLP-Anwendungen implementieren und trainieren können

Deep Learning



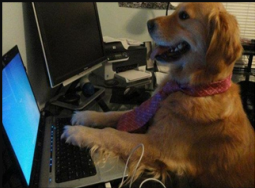
What society thinks I do



What my friends think I do



What other computer scientists think I do



What mathematicians think I do



What I think I do

```
from theano import *
```

What I actually do

Abbildung 1: Deep Learning, Quelle

Aufbau des Kurses

Zwei Teile:

1. Vorlesung & Übung
25.10.16 - 13.12.16
2. Seminar & Projekte
20.12.16 - 02.02.17

Termine:

Di, 14:15 - 15:45, INF 325, SR 24

Do, 09:15 - 10:45, INF 346, SR 10

- Kursthemen: [Kursseite](#), [Wikiseite](#)
- Literatur:
 - [\[Goldberg, 2015\]](#)
 - [\[Cho, 2015\]](#)
- Online-Kurs: [\[Socher, 2016\]](#)

Scheinerwerb

- regelmäßige Teilnahme an VL **und** Übung (verpflichtend!)
- alle Übungsblätter bearbeiten
- Referat **oder** Implementierungsprojekt
- wöchentliche Fragen im Seminarteil

Commitment! im LSF und im Kurs

Übungsaufgaben

- 4 theoretische Übungen
 - mathematische Voraussetzungen für das Verständnis von NNs
- 2 praktische Übungen
 - Implementierung in Python, Theano, Tensorflow
 - installiert auf `last`
- alle Übungsaufgaben müssen selbst substantiell bearbeitet werden
- Abgabefristen sind einzuhalten
- gegenseitige Hilfe ist erwünscht und erlaubt
- Übungsaufgaben werden korrigiert, aber nicht benotet

Implementierungsprojekte

- Alternative zu Referaten
- keine Gruppenarbeit (Ausnahmen bitte frühzeitig absprechen)
- Umfang:
 - Implementierung
 - einfaches trainierbares NN für Klassifikationsproblem (nur eigener Code, ohne Verwendung eines Frameworks) oder
 - mäßig komplexes NN für Sequenzklassifikation oder probabilistische Sequenzmodellierung (unter Verwendung eines Frameworks)
 - Dokumentation
- Ab dem 22.12. sind fünf Übungstermine ganz der gemeinsamen Arbeit an den Projekten gewidmet
- Abgabetermin: 31.3.2017, 23:59
- Weitere Details in der ersten Übungssitzung (Donnerstag)

- per E-Mail bis **Fr, 28. 10.** 2 präferierte Paper auswählen
- Proseminar kürzere (+), Hauptseminar komplexere Paper (++)
- evtl. zu zweit, je nach Anzahl der Interessenten
- Vorbesprechung mind. **1 Woche vor Präsentation** mit Folien
- bei Fragen während der Vorbereitung: Sprechstunde / E-mail
- max. 45min Zeit für Referat, Fragen & Diskussion

- verständliche Einführung in die Problematik der Arbeit geben
- selbst gewählten Schwerpunkt setzen
- Zeit dementsprechend einteilen
- Annahmen, Methodik, Schlussfolgerungen der Arbeit hinterfragen
- Folien nach dem Vortrag im Wiki hochladen

Vorläufiger Plan:

20.12. NN mit Struktur

Tree-LSTM (Lutz Büch) [Tai et al., 2015]

Recursive Autoencoders (++) [Socher et al., 2011]

10.01. Erweiterungen für RNNs

Adaptive Computation Time (++) [Graves, 2016]

Larger Context (++) [Wang and Cho, 2016]

17.01. Natural Language Understanding

Text Classification (+) [Zhang et al., 2015]

Neural Conversational Model (+) [Vinyals and Le, 2015]

24.01. Convolutional and Recurrent Neural Networks

Relation Classification (+) [Vu et al., 2016]

Neural Machine Translation with character-level decoder (++)
[Chung et al., 2016]

31.01. Erweiterungen für Seq2Seq

Hybrid Neural Machine Translation (++)
[Luong and Manning, 2016]

Copy Mechanism (++) [Gu et al., 2016]

- mindestens eines der Paper für jede Woche lesen
- 3 Fragen bis montags vor dem Referat per E-Mail an Julia
- können auch Anmerkungen, Kritikpunkte, Ideen sein



Cho, K. (2015).

Natural language understanding with distributed representation.

arXiv preprint arXiv:1511.07916.



Chung, J., Cho, K., and Bengio, Y. (2016).

A character-level decoder without explicit segmentation for neural machine translation.

In Proceedings of the 54th Annual Meeting of the Association for Computational Linguistics (Volume 1: Long Papers), pages 1693–1703, Berlin, Germany. Association for Computational Linguistics.

References ii



Goldberg, Y. (2015).

A primer on neural network models for natural language processing.

arXiv preprint arXiv:1510.00726.



Graves, A. (2016).

Adaptive computation time for recurrent neural networks.

arXiv preprint arXiv:1603.08983.



Gu, J., Lu, Z., Li, H., and Li, V. O. (2016).

Incorporating copying mechanism in sequence-to-sequence learning.

In Proceedings of the 54th Annual Meeting of the Association for Computational Linguistics (Volume 1: Long Papers), pages 1631–1640, Berlin, Germany. Association for Computational Linguistics.



Luong, M.-T. and Manning, C. D. (2016).

Achieving open vocabulary neural machine translation with hybrid word-character models.

In Proceedings of the 54th Annual Meeting of the Association for Computational Linguistics (Volume 1: Long Papers), pages 1054–1063, Berlin, Germany. Association for Computational Linguistics.



Socher, R. (2016).

Deep learning for natural language processing.

Stanford University Lecture.



Socher, R., Huang, E. H., Pennin, J., Manning, C. D., and Ng, A. Y. (2011).

Dynamic pooling and unfolding recursive autoencoders for paraphrase detection.

In Advances in Neural Information Processing Systems, pages 801–809.



Tai, K. S., Socher, R., and Manning, C. D. (2015).

Improved semantic representations from tree-structured long short-term memory networks.

pages 1556–1566.



Vinyals, O. and Le, Q. (2015).

A neural conversational model.

arXiv preprint arXiv:1506.05869.



Vu, N. T., Adel, H., Gupta, P., and Schütze, H. (2016).

Combining recurrent and convolutional neural networks for relation classification.

In Proceedings of the 2016 Conference of the North American Chapter of the Association for Computational Linguistics: Human Language Technologies, pages 534–539, San Diego, California. Association for Computational Linguistics.



Wang, T. and Cho, K. (2016).

Larger-context language modelling with recurrent neural network.

In Proceedings of the 54th Annual Meeting of the Association for Computational Linguistics (Volume 1: Long Papers), pages 1319–1329, Berlin, Germany. Association for Computational Linguistics.



Zhang, X., Zhao, J., and LeCun, Y. (2015).

Character-level convolutional networks for text classification.

In *Advances in Neural Information Processing Systems*, pages 649–657.