

Neural Networks: Architectures and Applications for NLP, Übungsblatt 4

Institut für Computerlinguistik - Universität Heidelberg

Wintersemester 2016/2017

Übungsaufgaben

Abgabe bis 15.12.2016, 9.15 Uhr

Name:

Matrikelnummer:

Diese Übung können Sie in Theano oder TensorFlow implementieren.

A.1

Implementieren Sie das Skip-Gram-Modell von Mikolov et al. (2013). Berechnen Sie Word-Embeddings auf drei englisch- oder deutschsprachigen Büchern¹ von Project Gutenberg (www.gutenberg.org). Wählen Sie zehn Wörter aus, die in allen drei Büchern vorkommen, und visualisieren Sie ihre Embeddings für alle drei Bücher in zwei Dimensionen mittels einer geeigneten Dimensionsreduktionstechnik.

Literatur

Mikolov, T., Sutskever, I., Chen, K., Corrado, G. S., and Dean, J. (2013). Distributed representations of words and phrases and their compositionality. In Burges, C. J. C., Bottou, L., Welling, M., Ghahramani, Z., and Weinberger, K. Q., editors, *Advances in Neural Information Processing Systems 26*, pages 3111–3119. Curran Associates, Inc.

¹Alle drei Bücher sollten in der gleichen Sprache verfasst sein.