

Masterarbeit

Entwicklung eines Deep Learning Frameworks mit Theano zur datengetriebenen Modellbildung dynamischer Mehrgrößensysteme

Philipp Quentin, Matrikelnr.: 29234307

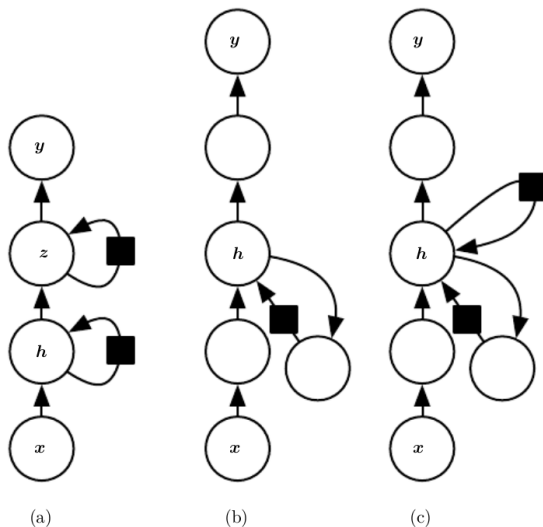


Abbildung 1: Varianten tiefer Rekurrenter Neuronaler Netze [2]

Um einen effizienten und sicheren Betrieb industrieller chemischer Prozesse gewährleisten zu können, sind Techniken zur frühzeitigen Fehlerdetektion und -diagnose erforderlich. Zu diesem Zweck können regelbasierte Expertensysteme, (physikalische) Prozessmodelle oder datengetriebene Verfahren eingesetzt werden. Aufgrund der Komplexität und Größe chemischer Prozessanlagen sind die erstgenannten Ansätze sehr zeit- und arbeitsintensiv bzw. mitunter nicht realisierbar.

Unter den datengetriebenen Verfahren versprechen tiefe (rekurrente) Neuronale Netze Vorteile bei der Modellbildung. Hierzu gehören eine bessere Generalisierungsfähigkeit als flache Neuronale Netze, was in einer geringeren Anzahl an erforderlichen Parametern und somit einer geringeren Neigung zu Overfitting resultiert. Zudem versprechen sie eine hohe Robustheit gegenüber topologischen Designentscheidungen, also der Anzahl an Schichten und Neuronen,

und somit ein relativ einfaches Engineering. Rekurrente Neuronale Netze zeichnen sich darüber hinaus durch eine implizite Termselektion aus, was sie für die Modellierung dynamischer Prozesse besonders attraktiv macht.

Im Rahmen dieser Arbeit sind Untersuchungen an (tiefen) KNN, RNN und LSTM bezüglich deren Leistungsfähigkeit zur Bildung von Simulationsmodellen durchzuführen.

Die Arbeit umfasst im Einzelnen folgende Teilaufgaben:

- Einarbeitung in die Modellklasse Rekurrenter Neuronaler Netze und in die Python Bibliothek Theano.
- Entwicklung und Implementierung eines Deep Learning Frameworks mit Theano zur Identifikation nichtlinearer Gleichungsfehler- und Ausgangsfehlermodelle.
- Auswahl geeigneter Benchmarkprobleme, wie bspw. dem Tennessee Eastman Prozess.
- Durchführung von Studien zur Untersuchung folgender Aspekte:
 - Toleranz der Ergebnisse gegenüber Designentscheidungen, wie der Wahl der Netztopologie (Anzahl an Schichten und Neuronen).
 - Vergleich der Ergebnisgüte tiefer rekurrenter mit flachen bzw. reinen vorwärtsgerichteten Architekturen.
 - Begründete Auswahl und ggf. Vergleich der Trainingsverfahren.
- Dokumentation der Ergebnisse und Kolloquiumsvortrag.

Literatur

- [1] P. Bashivan u. a. “Learning Representations from EG with Deep Recurrent-Convolutional Neural Networks”. In: *ICLR 2016*. 2016.
- [2] I. Goodfellow, Y. Bengio und A. Courville. *Deep Learning*. 2016. URL: <http://www.deeplearningbook.org>.
- [3] M. Hermans und B. Schrauwen. “Training and Analysing Deep Recurrent Neural Networks”. In: *Advances in Neural Information Processing Systems 26*. Hrsg. von C. J. C. Burges u. a. Curran Associates, Inc., 2013, S. 190–198. URL: <http://papers.nips.cc/paper/5166-training-and-analysing-deep-recurrent-neural-networks.pdf>.
- [4] S. El Hihi und Y. Bengio. “Hierarchical Recurrent Neural Networks for Long-Term Dependencies”. In: *Advances in Neural Information Processing Systems 8 (NIPS'95)*. Hrsg. von D.S. Touretzky, M. Mozer und M.E. Hasselmo. 1996.
- [5] S. Hochreiter und J. Schmidhuber. “Long Short-Term Memory”. In: *Neural Computation* 9 (1997), S. 1735–1780.
- [6] M. Långkvist, L. K. und A. Loutfi. “A review of unsupervised feature learning and deep learning for time-series modeling”. In: *Pattern Recognition Letters* 42 (2014), S. 11–24.
- [7] H.-T. Su, T. J. McAvoy und P. Werbos. “Long-term predictions of chemical processes using recurrent neural networks. A parallel training approach”. In: *Industrial & Engineering Chemistry Research* 31 (1992), S. 1338–1352.
- [8] I. Sutskever. “Training Recurrent Neural Networks”. Diss. University of Toronto, 2013.