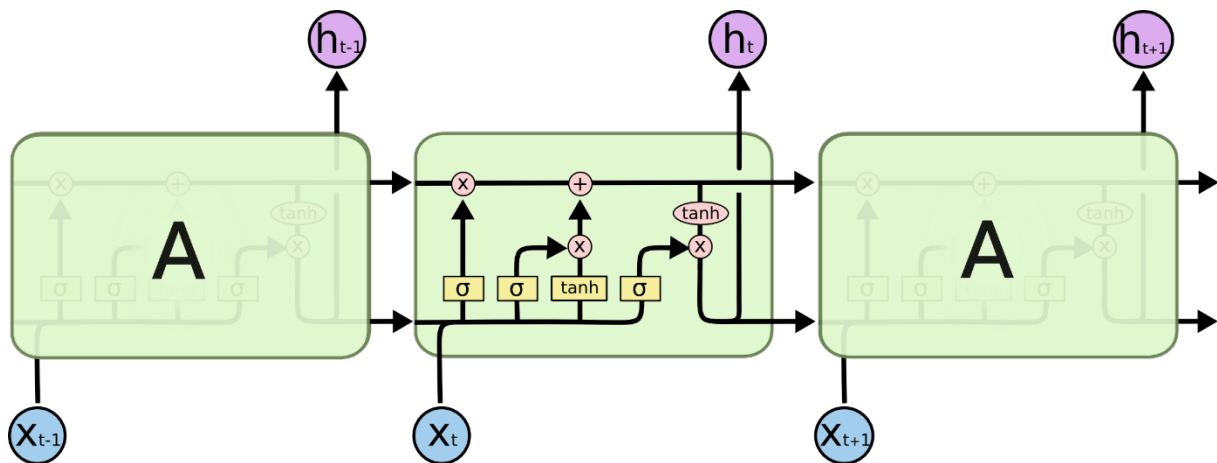


Recurrent Neural Networks (RNN) - LSTM

- Feedforward-Netz mit Rückkopplungsschleifen und zeitlich versetzter Backpropagation
- Einsatz: Spracherkennung, Übersetzer, Handschriftenerkennung
- Schwerer zu trainieren
- LSTM = Long short-term memory units, erfolgreichstes spezialisiertes RNN das das „Gradient vanishing“ Problem löst.



Hinweise zur Parametrisierung

- Vorsicht mit Overfitting (größere Netzwerke sind leistungsfähiger aber neigen leichter zum overfitt), „Testset“ aufheben, das nicht fürs Training verwendet wird, Trainingsset muß größer sein als die Zahl der Parameter; mehr Daten ist immer hilfreich
- Auf jeden Fall mehrere Epochen trainieren, nach jeder Epoche Kovergenz überprüfen
- Die Lernrate ist der wichtigste „Hyperparameter“
- Für LSTMs besser „softsign“ activation function statt „tanh“ verwenden
- Datennormalisierung nicht vergessen

Varianten von LSTM

- „peephole connections.“ [Gers & Schmidhuber (2000)]
- „coupled forget and input gates“
- Gated Recurrent Unit, oder GRU [Cho, et al. (2014)] → stellt eine Vereinfachung dar
- Depth Gated RNNs [Yao, et al. (2015)]

Literatur

<https://deeplearning4j.org/lstm.html#recurrent>

<https://colah.github.io/posts/2015-08-Understanding-LSTMs>