

Rekurrente / rückgekoppelte neuronale Netzwerke

Forschungsseminar Deep Learning 2018

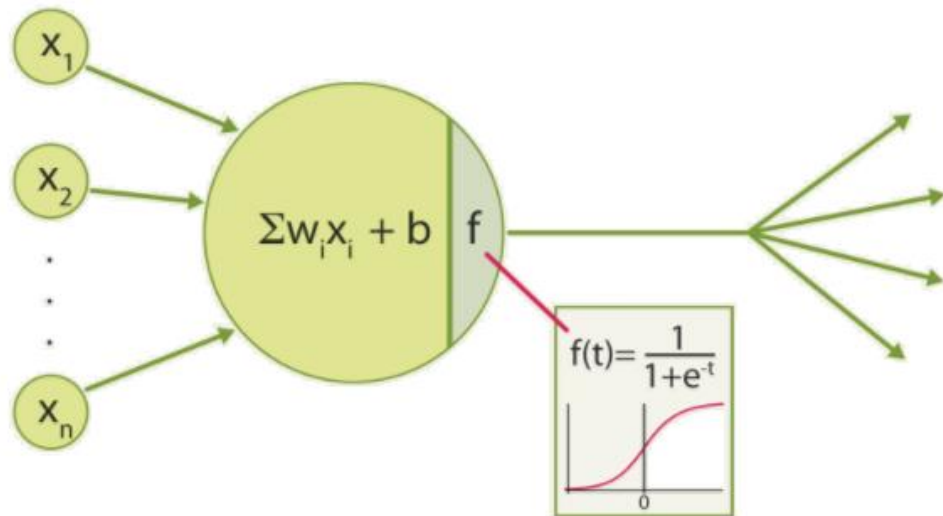
Universität Leipzig

12.01.2018

Vortragender: Andreas Haselhuhn

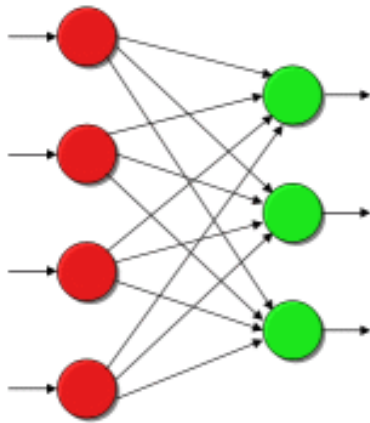
Neuronale Netzwerke

- Neuron besteht aus:
 - Eingängen
 - Summenfunktion
 - Aktivierungsfunktion
 - Ausgängen

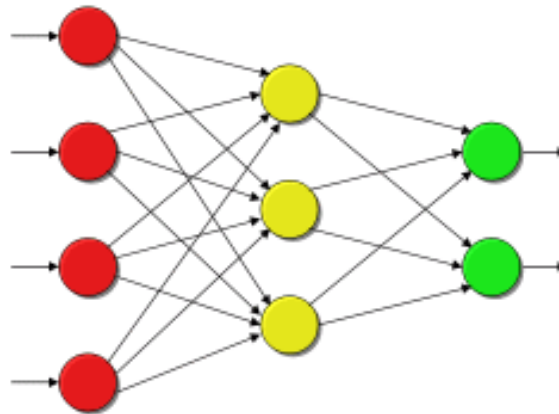


Einschichtige / Mehrschichtige feedforward-Netze

- Keine Kreise (baumähnliche Struktur)
- Kann keine oder nur schwer Vorhersagen treffen
- Wird meist zur Mustererkennung eingesetzt



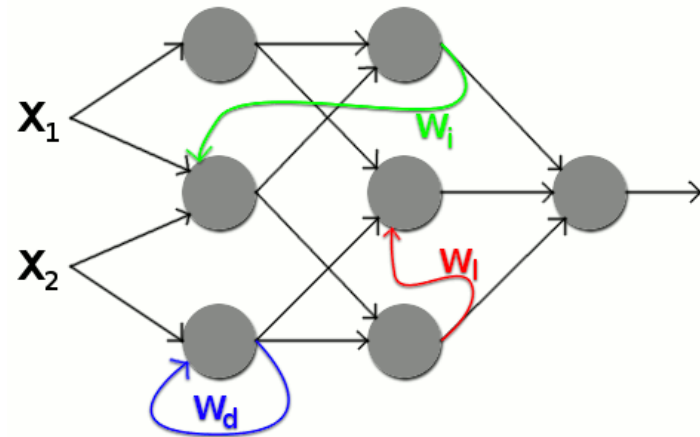
<http://www.neuronalesnetz.de/pattern.html>



<http://www.neuronalesnetz.de/units.html>

Rekurrente Neuronale Netze

- Auch Verbindungen zwischen der aktuellen und den vorangegangenen Schichten möglich
 - Es bilden sich Kreisstrukturen
- Kann Vorhersagen machen (ähnlich menschlicher Kognition)
- Hat ein Gedächtnis
- Wird eingesetzt bei:
 - Übersetzungen
 - Satzvervollständigung
 - Kurzen für Aktien berechnen
 - Wetter vorhersagen
 - Attraktorennetze
 - Tumorzellen Erkennung (im Stadium: aus dieser Zelle könnte eine Tumorzelle werden)



https://de.wikipedia.org/wiki/Rekurrentes_neuronales_Netz#/media/File:Neuronal-Networks-Feedback.png

Typen rekurrenter neuronaler Netzwerke

- Direkte Rückkopplung (direct feedback) Abb.: 1
 - wird der eigene Ausgang eines Neurons als weiterer Eingang genutzt
- Indirekte Rückkopplung (indirect feedback) Abb.: 2
 - verbindet den Ausgang eines Neurons mit einem Eingang eines Neuron der voran gehenden Schichten

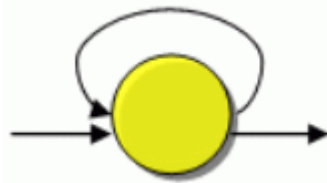


Abb.: 1,
<http://www.neuronalesnetz.de/rekurrente.html>

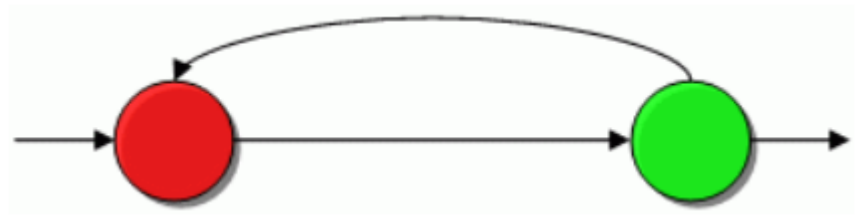
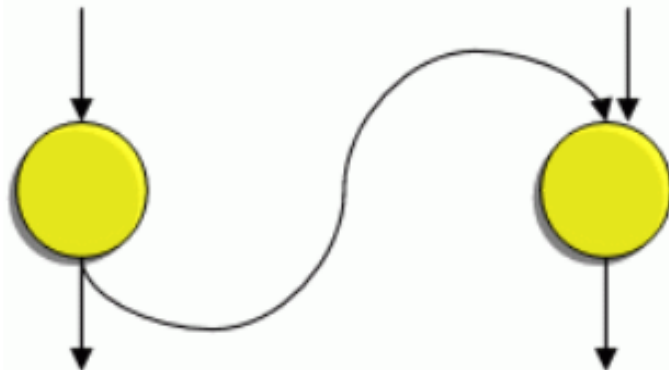


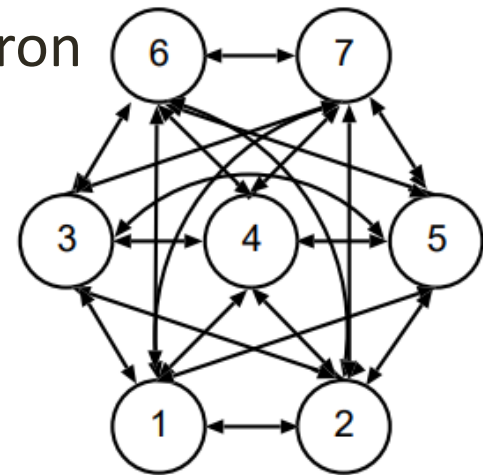
Abb.: 2, <http://www.neuronalesnetz.de/rekurrente.html>

Typen rekurrenter neuronaler Netzwerke

- Seitliche Rückkopplung (lateral feedback)
 - verbindet den Ausgang eines Neurons mit einem Eingang eines anderen Neurons der selben Schicht
- Vollständige Verbindung
 - Hier hat jeder Ausgang eine Verbindung zu jedem anderen Neuron



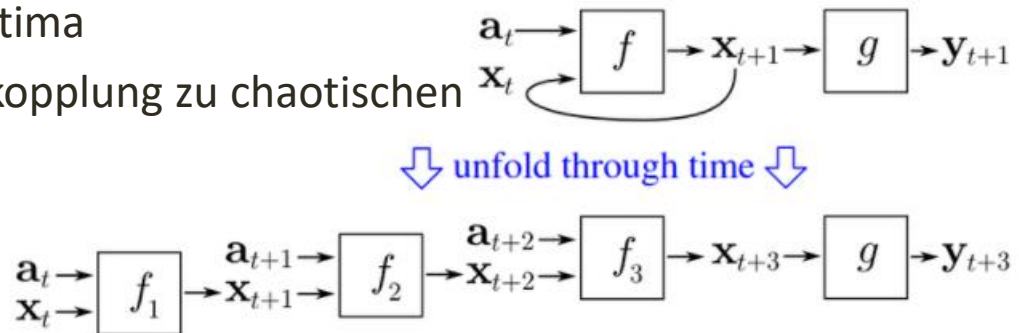
<http://www.neuronalesnetz.de/rekurrente.html>



https://lionel.kr.hs-niederrhein.de/~ueberholz/Lehre/NumInf2_WS1415/numinf_11_nn.pdf : Seite 14

Training von RNN's

- Backpropagation
 - Eingabe wird durch ein Netz geschickt und mit der erwarteten Ausgabe verglichen
 - Abweichungen von der erwarteten Ausgabe werden von der Ausgabeschicht zur Eingabeschicht zurück gesendet
 - Gewichtungen der Neuronen werden anhand ihres Einflusses auf den Fehler geändert
- Backpropagation through time (BPTT)
 - Ähnlich dem Backpropagation
 - Netzwerk wird durch die Zeit „aufgefaltet“
 - Anfällig für lokale Optima
 - kann durch die Rückkopplung zu chaotischen Zuständen führen

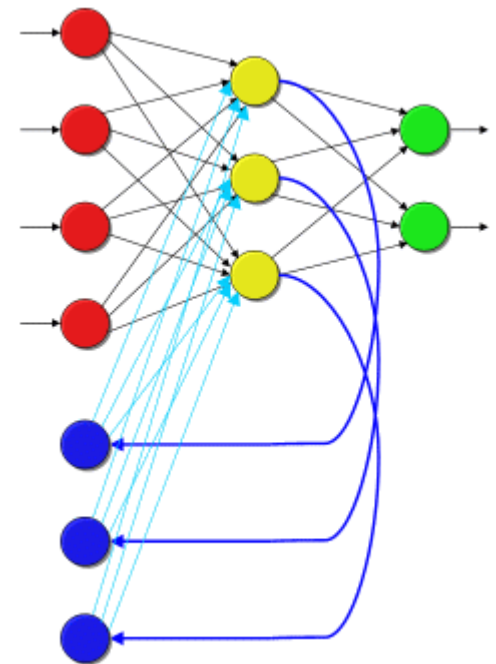


Beispiele für die Umsetzung mit RNN's

- 1. Simple Recurrent Networks
- 2. RNN zur Vorhersage eines Wertes (z.B. Aktienkurse / Skalarwertes)
- 3. RNN zur Sequenzen- / Wort- vorhersage

1. Simple Recurrent Networks [1]

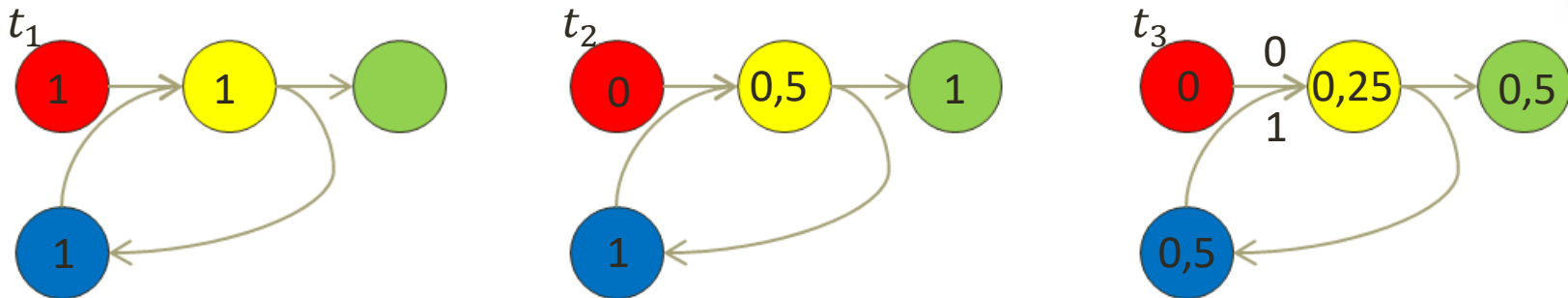
- Für jede Hiddenschicht existieren Kontext-Einheiten (KU) (blau)
 - Für jedes Neuron eine KU
 - KU verbunden zu jedem Neuron der Hidden-Schicht
 - KU befinden sich in der Schicht $i-1$
 - Erhalten Ausgabe des Neurons als Eingang
 - Verarbeiten diese und geben das Ergebnis in Zeitschritt $t+1$ zurück
 - Gewichtungen werden beim lernen (mit Backpropagation) angepasst und sind nicht fix



<http://www.neuronalesnetz.de/rekurrenente1.html>

1. Simple Recurrent Networks [1]

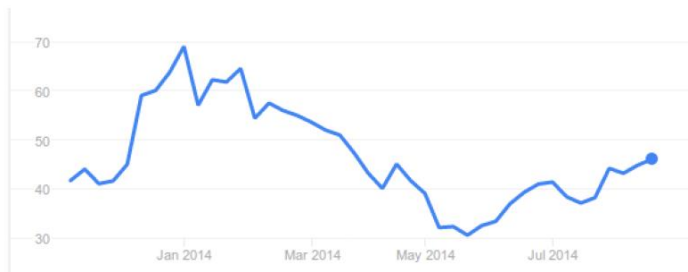
- Durch die KU besitzt das Simple RNN Teilinformationen aus sämtlichen vorangegangenen Zeitpunkten.
 - Da KU im Schritt „t“ eine Kopie der Hidden-Units aus t-1 haben
 - Im Schritt t-1 hatten sie eine Kopie von t-2



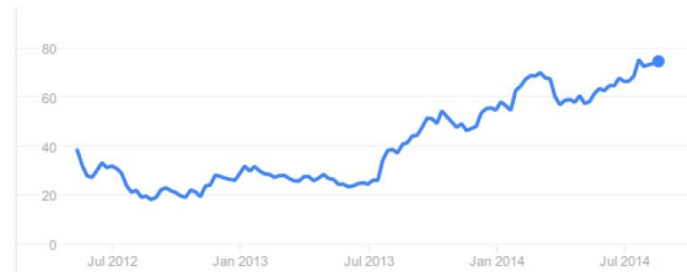
- KU mit modifizierbaren Gewicht kann als dynamischen Gedächtnis des RNN betrachtet werden.
- Simple RNN kann sehr flexible willkürliche Sequenzen lernen
- Einsatz für sich wiederholende Tätigkeiten (Tee/Kaffee Bestellung)

2. RNN zur Vorhersage eines Wertes (Skalar) [2]

- Aktienkurse



Company A

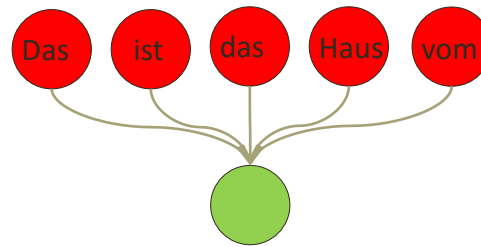
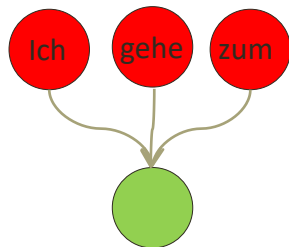


Company B

[2]
Seite 9

- Wortvorhersage

- „Das ist das Haus vom ...“, „Ich gehe zum ...“, „Das Flugzeug fliegt mit einer Höhe von 100 Meter nach ...“

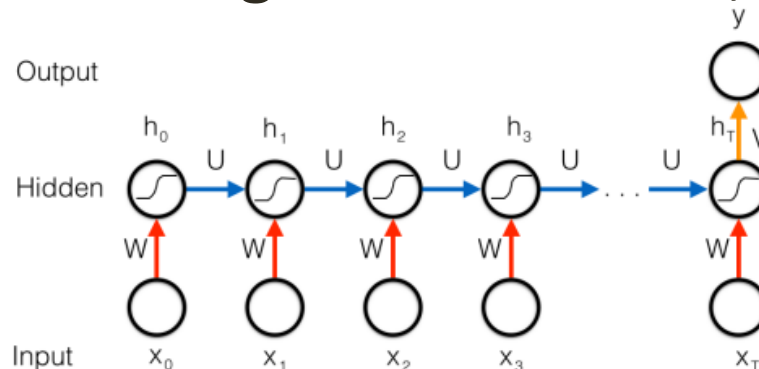


2. RNN zur Vorhersage eines Wertes (Skalar) [2]

- Problem: Verschiedene Anzahl an Eingaben
 - Somit müssten immer neue Eingabe- und Hidden-Knoten erstellt werden
 - Man müsste jedes Netz neu anlernen
- Lösung RNN mit Zeitabhängigkeiten
- Lernen nur mit Backpropagation through time (BPTT)

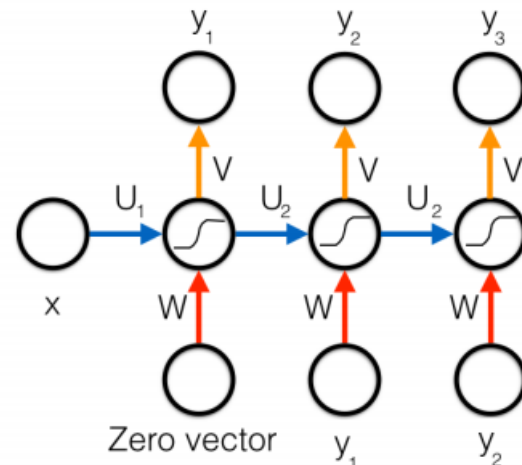
2. RNN zur Vorhersage eines Wertes (Skalar) [2]

- Variable Anzahl an Knoten
 - X ist Eingabe
 - H Hidden-Schicht und U Verbindungen zwischen diesen
 - V ist die Ausgabe
 - Wenn T wächst so gibt es immer W, U, V



3. RNN für Sequenzen z.B. Übersetzungen [2], [3]

- „Ich habe ich gesagt.“ -> „I have I said.“ ?
- Verschied lange Sätze bzw. Texte
- Lösung: ähnlicher Ansatz wie in 2.
 - Nur wird jetzt jedes mal ein Wert zurückgegeben



3. RNN für Sequenzen z.B. Übersetzungen [2], [3]

- Vorteil: es können zwischen Ergebnisse ausgegeben werden
- Nachteil:
 - es muss ein Wert „Zero“ eingeführt werden, da auch eine Ausgabe in $T=0$ erwartet wird
 - Es muss einen Wert(Knoten) X_0 geben der den Startwert enthält
 - BPTT ist noch anfälliger für Optima

Verweise / Quellen

- [1] <http://www.neuralesnetz.de/rekurrente.html>
- [2] A Tutorial on Deep Learning Part 2: Autoencoders, Convolutional Neural Networks and Recurrent Neural Networks
- [3] <http://www.wildml.com/2015/09/recurrent-neural-networks-tutorial-part-1-introduction-to-rnns/>
- [4] https://lionel.kr.hs-niederrhein.de/~ueberholz/Lehre/NumInf2_WS1415/numinf_11_nn.pdf
- [5] <https://www.heise.de/ct/ausgabe/2016-6-Die-Mathematik-neuronaler-Netze-einfache-Mechanismen-komplexe-Konstruktion-3120565.html>
- [6] https://de.wikipedia.org/wiki/Rekurrentes_neurales_Netz#/media/File:Neuronal-Networks-Feedback.png