

Entwicklung eines persönlichen KI-basierten adaptiven Assistenten

Facharbeit im Informatik LK1 am Ernst-Moritz-Arndt Gymnasium

von Anaïs Siebers bei Herrn van den Boom im Schuljahr 2016/17

An dieser Stelle möchte ich mich bei meiner Familie für ihre Geduld und Unterstützung bedanken.

amazing, isn't it?  
so, what is it?  
it hurts, isn't it?  
why would you do that?  
"you can do it.  
"i can do it.  
I can't do it.  
"i can do it.  
"don't do it.  
"i can do it.  
i couldn't do it. —

~Googles Künstliche Intelligenz  
(<http://www.wired.co.uk/article/google-artificial-intelligence-poetry>,  
Letzter Zugriff: 01.03.2017, 20:14)

# Inhaltsverzeichnis

Inhaltsverzeichnis	2
1. Vorwort	4
2. Der persönliche Assistent	5
2.1 Warum ein Assistent? Wofür?	5
2.2 Was bieten bereits vorhandene Assistenten?	5
2.3 Was ist Lückenhaft?	5
2.4 Was möchte ich erreichen?	6
2.5 Ideen zur Umsetzung	7
3. Interaktion mit dem Assistenten	8
3.1 Eingabe	8
3.2 Ausgabe	8
3.3 Verwendungszweck	8
4. Analyse des Textes	10
4.1 Ziel	10
4.2 Problematik	10
4.3 Recherche	10
4.4 Anbindung	11
4.5 Ergebnis	12
5. Das Rekurrente Neuronale Netz	13
5.1 Prinzip/Ansatz	13
5.2 Aufbau eines Rekurrenten Neuronalen Netzes	14
5.2.1 Das Biologische Neuronale Netz	14
5.2.1.1 Allgemein	14
5.2.1.2 Input	15
5.2.1.3 Output	15
5.2.2 Das Künstliche Neuronale Netz	15
5.2.2.1 Layers	15
5.2.2.2 In-/Output	16
5.2.2.3 Gewichte	16
5.2.2.4 Aktivität	17
5.2.2.6 Training	18
5.2.2.5 Lernregeln	18
5.2.3 Das Rekurrente Neuronale Netz	20
5.3 Verwendung	20

5.3.1 ANNAs Neuronales Netz (siehe 12. Implementierung)	20
5.3.1.1 Layers	20
5.3.1.2 Aktivität	21
5.3.1.3 Training	21
5.3.1.4 Lernregel	21
6. Umsetzung von ANNA	23
6.1 Idee	23
6.2 Umsetzung	23
6.3 Test	24
6.4 Änderungen	25
6.5 Zweiter Test	25
6.6 Ergebnis	26
6.7 Problematik	26
7. ANNAs Oberfläche	28
7.1 Prinzip	28
7.2 Umsetzung	28
7.3 gewünschtes Ergebnis	31
7.4 Resultat	31
8. Zusammenfassung	32
8.1 Aufwand	32
8.2 Bewertung des Ergebnisses	32
9. Literaturverzeichnis	34
10. Anhang	36
11. Implementierung	53
12. Beratungsprotokolle zur Facharbeit	93
13. Selbstständigkeitserklärung	94

# 1. Vorwort

In dieser Facharbeit wird ein Ansatz für die Entwicklung eines KI-basierten Assistenten beschrieben. Der Assistent soll im Gegensatz zu anderen Assistenten die Individualität und Persönlichkeit des Nutzers würdigen. Das Prinzip des Assistenten ist an den Lern- und Wachstumsprozess des Menschen angelehnt.

Der Assistent heißt ANNA. Der Name wurde bewusst gewählt und erleichtert die Beschreibung des Assistenten im Laufe der Facharbeit. Er steht für:

**A** daptiver  
**N** atürlicher  
**N** euronaler  
**A** ssistent

Im Gegensatz zu anderen Assistenten soll er mehr auf den Benutzer und dessen Interessen eingehen. Andere Assistenten werden überwiegend von wirtschaftlichen Interessen des Anbieters bestimmt. So wird bei Google, zum Beispiel, meistens zuerst Amazon kommen, wenn man nach einer Ware sucht. Doch bei ANNA soll es möglich sein, seine eigenen Quellen auszusuchen und zu priorisieren.

Da in der Zwischenzeit unter <http://anna.is.it> eine Webseite zur Verfügung steht, auf der man den aktuellen Stand des Assistenten LIVE ausprobieren kann, habe ich in der vorliegenden Version diesbezüglich kleinere Korrekturen und Anpassungen vorgenommen.

Aufgrund der Komplexität des Themas befindet sich ANNA allerdings immer noch in einem Beta Stadium und ist Gegenstand weiterer Entwicklung und Forschung.

## 2. Der persönliche Assistent

### 2.1 WARUM EIN ASSISTENT? WOFÜR?

Es gibt heutzutage diverse sogenannte intelligente persönliche Assistenten und Wissensnavigatoren. Der führende Assistent ist laut James A. Martin, der Alexa, Cortana, Google Assistant und Siri getestet hat, Siri<sup>1</sup> (vgl. Martin 05.10.2016). Der Name Siri leitet sich von der Abkürzung des Englischen Äquivalents „Speech Interpretation and Recognition Interface“ ab. Jedoch ist Apple nicht die einzige Corporation, die einen sprachgesteuerten Assistenten anbietet. Die beiden Wirtschaftsriesen Google und Microsoft arbeiten an ihren eigenen Systemen, welche Google Now und Cortana heißen. Auch andere große bekannte Firmen, wie die IBM oder Amazon investieren in sprachgesteuerte KI-basierte Recherche Assistenten.

In Zukunft werden viele Menschen diese Assistenten nutzen, da sie intuitiv zu bedienen sind und dadurch die Schnittstelle zwischen Mensch und Informationsverarbeitung einfacher wird.

### 2.2 WAS BIETEN BEREITS VORHANDENE ASSISTENTEN?

Die auf dem Markt vertretenen Assistenten bieten bereits heute ein umfangreiches Angebot. Nimmt man Siri als Beispiel, so hat sie als moderner hochwertiger Assistent grundsätzlich verschiedene Funktionen:

- Sie bietet Unterstützung bei auf den Geräten vorhandenen Apps (Anwendungen) beziehungsweise den dahinter ablaufenden Prozessen. Ein klassisches Beispiel hierfür ist „Siri schreibe eine Mail an ...“. Dann nutzt Siri den Prozess des Mail-Schreibens der Mail App.
- Eine weitere Funktion ist die Funktion der Recherche. Das beste Beispiel hierfür ist: „Was ist ... Siri?“. Fragt man Siri dies, so wird sie als Antwort den Wikipedia-Text beziehungsweise einen bei Google gefundenen Link ausgeben.

### 2.3 WAS IST LÜCKENHAFT?

Das größte Problem, dass die Assistenten heutzutage haben, ist, dass ihnen Grundmechanismen menschlicher Kommunikation fehlen und nur sehr schwache Interaktion

bereitstellen. Ein gutes Beispiel hierfür sind Namen. Wenn man einen Namen ausspricht, der nicht aus dem deutschen Sprachraum kommt, so ist die Wahrscheinlichkeit groß, dass Siri den Namen falsch schreibt und falsch wiedergibt. Dies ist darauf zurückzuführen, dass der Wortschatz und die Grammatik, die Siri vorliegen auf einem, zum Beispiel dem deutschen Sprachraum, ausgelegt sind. Man kann Siri nun aber nicht verbessern und etwas wie: „mit t in der Mitte“, sagen, da Siri den Bezug nicht herstellen kann.

Außerdem können die Systeme das Suchverhalten und die Interessen einer Person noch nicht lernen. Zwar bietet Cortana zum Beispiel Personalisierungsfeatures an, welche jedoch nur den Wohnort für das Wetter und ein Notizbuch von Cortana, in dem man seine Interessen und Präferenzen speichern kann, enthalten. Auch Apple bietet die Möglichkeit der Personalisierung an. Eine Funktion ist zum Beispiel dass man dem System seinen Namen beibringen kann. Diese Einstellungen sind allerdings alle über ein weiteres Menü vorzunehmen und ermöglichen keine leichtere nach der Person ausgerichtete Suche und Recherche.

Neben den technischen Aspekten ist es auch sehr wichtig, dass der Nutzer die Auswahl der herangezogenen Informationsquellen beeinflussen kann und hierbei nicht von den Interessen der Betreiber geleitet wird. Momentan fehlt vor allem die Würdigung der Benutzerpersönlichkeit, die Möglichkeit für Informationen eine eigene Wertehierarchie vorzugeben sowie eigenes oder - persönlich validiertes - fremdes Wissen zu speichern und in die Suchergebnisse einzugliedern.

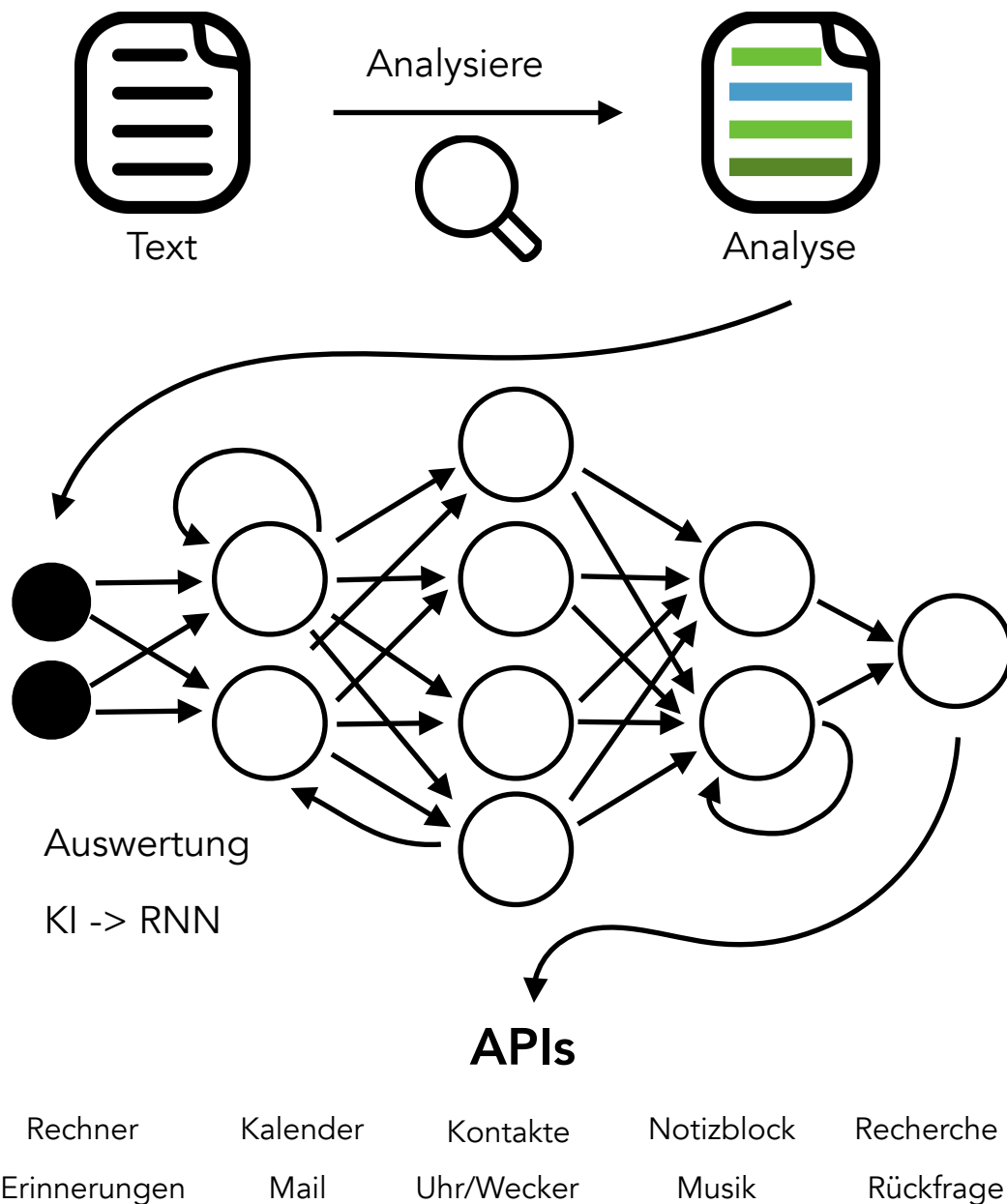
## 2.4 WAS MÖCHTE ICH ERREICHEN?

Mit diesem Projekt möchte ich Grundlagen für einen adaptiven Assistenten implementieren, der ein Neuronales Netz nutzt, um das Verhalten und die Persönlichkeit des Nutzers zu erlernen. Nachdem der Assistent das Verhalten des Nutzers gelernt hat, soll er auch besser Antworten auf Fragen liefern und somit das Kriterium eines effektiven Assistenten erfüllen.

Dies impliziert ein gegenseitiges Kennenlernen bei dem sich der Benutzer das Vokabular und Verständnis von ANNA aneignet und ANNA im Gegenzug vom Benutzer lernt sowie seine Vorlieben und Abneigungen zunehmend einbezieht.

## 2.5 IDEEN ZUR UMSETZUNG

Für die Analyse des Textes und für die Handhabung und das Interface der Datenbank werde ich zunächst nach bereits vorhandenen Bibliotheken suchen. Mein eigentliches Projekt im Rahmen der Facharbeit wird die Umsetzung eines Rekurrenten Neuronales Netzes sein. Dieses wird für die Künstliche Intelligenz von ANNA zuständig sein.





## 3. Interaktion mit dem Assistenten

### 3.1 EINGABE

Zunächst muss der Assistent eine Eingabe übergeben bekommen, auf welche er dann reagieren soll. Diese Eingabe kann zu einem späteren Zeitpunkt auch über Spracheingabe erfolgen.

Der Nutzer müsste sich zunächst authentifizieren und könnte danach mit ANNA interagieren. Die Identifikation ist notwendig, damit Datenschutz gewährleistet werden kann und damit es möglich ist ANNA für sich zu individualisieren, also die Möglichkeit zur privaten Datenspeicherung verfügbar ist.

Die Authentifizierung befindet sich momentan noch in der Implementierung.

### 3.2 AUSGABE

ANNA gibt die Ergebnisse der Suche nach Gewichten sortiert aus. Die Gewichte stellen sich aus den vorherigen Suchen oder der Einschränkungen durch Filter, die das Suchen erleichtern sollen, zusammen.

ANNA speichert die unterschiedlichen Präferenzen der Suchquellen zu den verschiedenen Nutzern ab und kann somit bei der nächsten Suche automatisch „von alleine“ direkt in den bevorzugten Quellen des jeweiligen Nutzers suchen (siehe Anhang I). So sucht der eine Nutzer bevorzugt bei [chefkoch.de](http://chefkoch.de) und der andere Nutzer präferiert [wikipedia.de](http://wikipedia.de).

### 3.3 VERWENDUNGSZWECK

Der persönliche Assistent ANNA kann sowohl privat als auch im beruflichen Umfeld im deutschsprachigen Raum eingesetzt werden. Im Gegensatz zu anderen Assistenten orientiert sich ANNA nicht an kommerziellen/monetären Interessen sondern an den persönlichen Axiomatik des Nutzers mit dem Ergebnis, dass er das gewünschte Ergebnis seiner Suche auf der ersten Seite an oberster Stelle findet.

Dies erhöht auch die Chancen für kleinere Anbieter, die bei den größeren Suchmaschinen aufgrund von Werbeschaltungen größerer Unternehmen meistens erst auf den Seiten nach der

Seite 5 auftauchen. Wenn darüber hinaus der Nutzer auch entscheiden kann, welches Gewicht Regionalität der Ergebnisse für ihn hat, ist dies ein wichtiger Faktor für Pluralismus und Diversität von Informationen und Angebot (siehe 5.1 Prinzip/Ansatz).

## 4. Analyse des Textes

### 4.1 ZIEL

Die Analyse des eingegebenen Textes ist die Basis des persönlichen Assistenten. Sie ist dafür zuständig, dass ANNA die Eingabe des Users verwerten kann. Ziel der Analyse ist, dass die Bestandteile des Satzes analysiert und in einer für das Rekurrente Neuronale Netz verwendbaren Art aufbereitet und zur Verfügung gestellt werden.

### 4.2 PROBLEMATIK

Die deutsche Sprache ist eine sehr komplexe Sprache und einzelne Wörter können mehrere Bedeutungen haben, sogenannte Homonyme. Außerdem kann fast jedes Wort verschiedene Genera oder Kasus enthalten. Das Problem ist das Erkennen des richtigen Genus und Kasus.

Nimmt man den Satz „Ich heiße Anaïs.“ als Beispiel, dann muss zunächst das „Ich“ analysiert werden. Dieses „Ich“ kann sowohl Subjekt, Akkusativ, Neutrum, Singular als auch Subjekt, Dativ, Singular, Neutrum, ... Abgesehen von der Vielzahl an Möglichkeiten was die Beugungen von „Ich“ in diesem Satz angeht, stellt auch das nächste Wort „heiße“ ein Problem, da es sowohl Verb als auch Adjektiv sein kann.

### 4.3 RECHERCHE

Bei meiner Recherche habe ich mich zunächst damit beschäftigt, welche bereits vorhandenen Bibliotheken in Frage kommen könnten. Dabei bin ich auf die Bibliothek LanguageTool (siehe Anhang II) gestoßen, die unter anderem als Plug-In bei OpenOffice Verwendung findet. LanguageTool ist ein kostenloses Programm, welches eingegebenen Texte auf Grammatik- und Rechtschreibfehler überprüft und sogar die Sprachwahl korrigiert.

Das in Java entwickelte LanguageTool ist quelloffen und unter einer freien Lizenz stehend. Es hat für den Einstieg eine `Example.class` (siehe Anhang III). Anhand dieser Klasse ist es möglich, die Analysenmöglichkeiten, die LanguageTool bietet, auszuprobieren und zu vergleichen.

Es gibt mehrere Möglichkeiten Sätze in LanguageTool zu analysieren. Zunächst muss die Sprache festgelegt werden, da LanguageTool international verwendet werden kann. Da ANNA nur für den deutschen Sprachraum gedacht ist, habe ich für meine Implementierung die deutsche Sprache fest eingestellt.

Mithilfe der Werkzeuge von LanguageTool kann man Methoden wie `check(testsatz)` benutzen. Diese gibt dann die Grammatikregeln, die für einen Satz angewendet wurden aus. Die Methode `getLemmas()`, auf der anderen Seite, gibt zurück, was die Grundform eines Wortes ist, beziehungsweise sein kann. Das Wort „heiße“ kann zum Beispiel das Verb „heißen“ sein (Ich heiße...), es kann aber auch das feminine Adjektiv von „heiß“ (Die heiße Suppe...) (siehe Anhang IV).

Für die Anbindung von LanguageTool galt es auch, insbesondere die Verbindung zum Rekurrenten Neuronalen Netz zu beachten. LanguageTool soll für das Rekurrente Neuronale Netz eine Erleichterung des Verständnis sein und daher war das Ergebnis der Analyse von LanguageTool darauf ausgelegt, dem Rekurrenten Neuronalen Netz, die gefundenen Informationen übersichtlich bereitzustellen.

Da ANNA später mithilfe der Auswertung der Sätze arbeiten soll, ist eine Auswertung und adequate Form der Speicherung für die Wörter und ihre Informationen sinnvoll. Diese Speicherung wird durch eine angelegte Hilfsklassenstruktur vorgenommen. Die Sätze werden zunächst in Wörter geteilt und dann anhand der Wortklassen gespeichert. Es wird dabei das Lemma und der Kasus-Numerus-Genus gespeichert.

Die Wortklassen sind die Arten der Wörter. Für die Einteilung wurden die Wortklassensysteme von Morphy (Vollständiges Klassensystem, großes und kleines Tag Set)<sup>42</sup> (Lezius 01.08.1998: 5) benutzt, da sie ein erforschtes System darstellen, welches von LanguageTool verwendet wurde und welches mit SGML-Abkürzungen beschrieben ist (siehe Anhang V).

Im Laufe der weiteren Recherche, auch im Bezug auf das Rekurrente Neuronale Netz (siehe Rekurrentes Neuronales Netz), stellte sich dann heraus, dass die Lemmata am Ende für die Verarbeitung am Wichtigsten sind. Sie stellen einen einheitlichen Nenner der Wörter her. Daher bekommt das Rekurrente Neuronale Netz als Eingabe die Grundform der Wörter.

Das Ergebnis einer Analyse mit LanguageTool durch die modifizierte `Example.class` ist eine Liste mit allen zu den Wörtern gefundenen Daten (siehe Anhang VI). Diese setzt sich aus dem eingegebenen Wort, dem/den dazugehörigen Lemma/ta und der Wortform, die durch die Wortklassensysteme definiert ist, zusammen

Hier ist beispielsweise das Wort Ich analysiert:

```
Ich[Ich/SUB:AKK:PLU:NEU,Ich/SUB:AKK:SIN:NEU,Ich/SUB:DAT:PLU:NEU,Ich/SUB:DAT:SIN:NEU,Ich/SUB:GEN:PLU:NEU,Ich/SUB:GEN:SIN:NEU,Ich/SUB:NOM:PLU:NEU,Ich/SUB:NOM:SIN:NEU,ich/PRO:PER:NOM:SIN:ALG,B-NP]
```

```
Ich[Ich/SUB:AKK:PLU:NEU,Ich/SUB:AKK:SIN:NEU,Ich/SUB:DAT:PLU:NEU,Ich/SUB:DAT:SIN:NEU,Ich/SUB:GEN:PLU:NEU,Ich/SUB:GEN:SIN:NEU,Ich/SUB:NOM:PLU:NEU,Ich/SUB:NOM:SIN:NEU,ich/PRO:PER:NOM:SIN:ALG,B-NP]
```

```
Ich[Ich/SUB:AKK:PLU:NEU,Ich/SUB:AKK:SIN:NEU,Ich/SUB:DAT:PLU:NEU,Ich/SUB:DAT:SIN:NEU,Ich/SUB:GEN:PLU:NEU,Ich/SUB:GEN:SIN:NEU,Ich/SUB:NOM:PLU:NEU,Ich/SUB:NOM:SIN:NEU,ich/PRO:PER:NOM:SIN:ALG,B-NP]
```

Substantiv	Nominativ	Genitiv	Dativ	Akkusativ	Neutrum
Singular	Plural	Pronomen	Personal	Allgemein	
B-NP -> chunk(B = Anfang eines chunks; NP = Nomen)					

Der obere Abschnitt zeigt die direkte Analyse von LanguageTool und der untere das Ergebnis, welches zunächst geteilt und mithilfe der Hilfsklassen geordnet gespeichert und im Anschluss in einen String zurückverwandelt wurde.

Die bunten SGML-Abkürzungen sind anhand der „Wortklassensysteme von Morphy“<sup>2</sup> (ebd.) übersetzt (siehe Anhang VII). Die Besonderheit am Ende des Satzes zeigt einen Chunk, der für die Grammatiküberprüfung von LanguageTool benutzt wird.

Ein chunk<sup>3</sup> (vgl. [wikidot.com](http://wikidot.com) 07.09.2014) markiert in LanguageTool einen Teil, der auf die KNG-Kongruenz (Kasus-Numerus-Genus Übereinstimmung) überprüft werden muss. Dabei signalisiert B den Beginn eines chunks. E wäre, auf der anderen Seite, das Endzeichen eines chunks. Das NP zeigt, dass es sich um ein Nomen handelt. Es könnte auch VP (Verb) sein. Ein Beispiel für die sinnvolle Verwendung von chunks wäre „die hübsche groß Frau“. Hier ist das groß nicht angepasst. Daher würde mittels des chunks der Abschnitt von „die“ bis „Frau“ markiert und anhand des NP „Frau“ angepasst werden.

# 5. Das Rekurrente Neuronale Netz

## 5.1 PRINZIP/ANSATZ

Das Rekurrente Neuronale Netz ist der wesentliche Bestandteil der Künstlichen Intelligenz von ANNA. Es ist an das menschliche Neuronale Netz angelehnt und somit an ein biologisches Konzept. Daher soll ANNA auch im Sinne der Bionik dem menschlichen Verhalten angelehnt sein. Dies soll unter anderem das Lernverfahren und den Denkprozess beinhalten.

Doch damit ANNA den menschlichen Denkprozess nachahmen kann, muss zunächst betrachtet werden, in welchen Punkten sich Computer und Menschen im Wesentlichen unterscheiden. Hier sieht man man einen Vergleich zwischen dem menschlichen Gehirn und dem Computer<sup>4</sup> (vgl. Kriesel 2007: 4):

	<b>Gehirn</b>	<b>Computer</b>
Anzahl Recheneinheiten	$\approx 10^{11}$	$\approx 10^9$
Art Recheneinheiten	Neurone	Transistoren
Art der Berechnung	massiv parallel	i.d.R. seriell
Datenspeicherung	assoziativ	adressbasiert

Deutlich sind die Unterschiede zwischen dem menschlichen Gehirn und dem Computer zu sehen. Es sind alles technische beziehungsweise biologische Unterschiede. Abgesehen von der Rechenleistung, die noch beim Menschen höher ist als beim Computer, aber jedoch in naher Zukunft über die des Menschen steigen wird<sup>5</sup> (vgl. Schmidhuber 2016: 2:54 min), unterscheiden sich die anderen Kategorien grundlegend.

Obwohl diese Unterschiede wie die Rechenleistung und Datenspeicherung zunächst nicht zu ändern sind, kann jedoch in der Art der Verarbeitung etwas verändert werden. Momentan wird bei vielen Assistenten eine sehr große Datenbank zur Speicherung von Informationen genutzt. Aber dies ist für das Projekt nicht möglich. Eine Datenbank für die Wörter und ihre Priorisierung ist notwendig, aber es ist nicht möglich riesige Datenbanken mit Verbindungen zwischen Wörtern anzulegen. Daher soll ANNA dem Menschen ähnlich handeln und anhand von Assoziationen zu einem Ergebnis kommen. Diese Assoziationen sollen im Rekurrenten Neuronalen Netz repräsentiert werden.

Die aktuellen Assistenten werden von großen Forscherteams entwickelt und geprüft. Da diese Ressourcen nicht zur Verfügung stehen, gilt es einen anderen Lösungsansatz zu wählen. Seit kurzer Zeit wird intensiv an dem Verfahren DeepLearning geforscht. Dies ist ein Verfahren, das selbstlernend ist und sich von alleine korrigiert. Das „Schweizer Forschungsinstitut für Künstliche Intelligenz IDSIA (Univ. Lugano & SUPSI)“ hat es geschafft, dass mithilfe von DeepLearning ein Computer eine Sprache gelernt hat, die keiner von den Forschern beherrschte<sup>5</sup> (ebd.).

Das Ziel von ANNA ist das Speichern des Suchverhaltens des Nutzers und seine Wertehierarchie im Bezug auf Quellen (siehe 8. ANNAs Ausgabe).

Zunächst müssen die wesentlichen Merkmale von ANNA festgelegt werden. Das Wichtigste eines neuronalen Netzes ist die Klassifikation. „Klassifikation ist ein allgemeiner Vorgang, um verschiedene Objekte oder Ereignisse einer oder mehreren Klassen zuzuordnen und darauf aufbauend Entscheidungen zu treffen.“<sup>6</sup> (Zöllner 2014: 7). Da ANNA auf den User ausgelegt sein soll, gilt es, sich nach dem User zu richten, und daher wird ANNAs Ergebnis des Rekurrenten Neuronalen Netzes nach der Wahl des User ausgerichtet sein. Dadurch ergibt sich eine automatische Klassifikation.

Diese Art von neuronalen Netzen haben eine Lernart, die sich „überwachtes Lernen (engl. supervised learning)“<sup>6</sup> (Zöllner 2014: 8) nennt. Der Nutzen der Aktion wird vom Benutzer bewertet, indem er angibt, ob das gewünschte Ergebnis ausgegeben wurde oder nicht. Es gibt auch „unüberwachtes Lernen (engl. unsupervised learning)“ und „bestärkendes Lernen (engl. reinforcement learning)“<sup>6</sup> (ebd.).

## 5.2 AUFBAU EINES REKURRENTEN NEURONALEN NETZES

### 5.2.1 Das Biologische Neuronale Netz

#### 5.2.1.1 Allgemein

---

Ein biologisches Neuronales Netz ist ein Netz aus vielen Neuronen, die miteinander verbunden sind (siehe Anhang VIII). „Ein Neuron ist nichts weiter als ein Schalter mit einem Informationseingang und -ausgang.“<sup>4</sup> (Kriesel 2007: 19, 2.2). Dem Neuron wird über die Dendriten an die Synapse eine Information übergeben. Dendriten sind zuständig für das Empfangen von elektrischen Impulsen. Ein Neuron hat viele Dendriten.<sup>4</sup> (vgl. ebd. 20, 2.2.1.2)

### 5.2.1.2 Input

---

Es gibt zwei Arten von Synapsen: chemische und elektrische Synapsen. Die Synapsen unterscheiden sich in ihrer Funktionsweise. Elektrische Synapsen lösen eine fest vorgeschriebene Reaktionen aus wie zum Beispiel Fluchtreflexe, weil sie eine direkte Verbindung zum Zellkern haben. Chemische Synapsen, auf der anderen Seite, bekommen ein elektrisches Signal, welches eine chemische Reaktion auslöst und dadurch einen Übermittlungsstoff (sog. Neurotransmitter) freisetzt. Dieser wird, wenn er an der Zellwand andockt, wieder in ein elektrisches Signal verwandelt (siehe Anhang IX)<sup>4</sup> (vgl. ebd. 20, 2.2.1.1).

Durch die Neurotransmitter ist eine größere Varietät an Intensität und Überleitung von Informationen möglich. „Die Regulierungsvielfalt ist enorm, und dass die Synapsen hier auch noch variabel sind [...] ist einer der zentralen Punkte bei Betrachtung der Lernfähigkeit des Gehirns.“<sup>4</sup> (ebd.).

### 5.2.1.3 Output

---

Die im Zellkern (Soma) ankommenden Impulse werden aufaddiert bis sie einen bestimmten Wert erreichen. Ist dieser Schwellwert erreicht, dann gibt das Neuron ein Signal weiter, da die Schwelle des Axons nun überschritten werden kann. Das Axon leitet nun den Impuls zum Dendriten des nächsten Neurons weiter. Es ist isoliert und kann somit die Impulse besser weiterleiten (siehe Anhang X)<sup>4</sup> (vgl. ebd. 21, 2.2.1.3 und 2.2.1.4).

Zusammenfassend lässt sich sagen, dass ein Neuron über die Dendriten „Impulse von durchschnittlich  $10^3$  bis  $10^4$  anderen Neuronen“<sup>4</sup> (ebd. 33, 2.5) erhält. Diese werden dann von Synapsen an das Soma weitergeleitet und dort kumuliert bis der Schwellwert des Axons überschritten ist. Dann wird der Impuls weitergegeben. Der nun übergebene Impuls ist eine Zusammenfassung der anderen kumulierten Impulse <sup>4</sup> (vgl. ebd. 33, 2.5).

## 5.2.2 Das Künstliche Neuronale Netz

### 5.2.2.1 Layers

---

Das Künstliche Neuronale Netz ist eine vereinfachte Übertragung der komplexen Prozesse des Gehirns auf eine technische Ebene. Es gibt drei sog. Units. Die Input-Unit oder auch das Input-Layer, die Hidden-Unit/das Hidden-Layer und die Output-Units/das Output-Layer



(siehe Anhang XI)<sup>7</sup> (neuronaalesnetz.de: 5). Unit ist der englische Name für die Neuronen (die auch Perzeptronen genannt werden)<sup>6</sup> (vgl. Zöllner 2014: 10, 3.3).

### 5.2.2.2 In-/Output

Der In- und Output der Neuronen lässt sich von den Faktoren des biologischen Neuronalen Netzes ableiten. Dort kommt man zu diverse Abrissen<sup>4</sup> (vgl. Kriesel 2007: 33f., 2.5):

<b>Input eines Neurons</b>	<b>Output eines Neurons</b>	<b>Generell</b>
<u>Vektorieller Input</u> Der Input hat einen Vektor. Er kommt aus vielen Richtungen von vielen Neuronen und nach der Kumulation verlässt er als neuer Vektor das Neuron.	<u>Aufkumulieren der Inputs</u> Die diversen Impulse werden kumuliert. Technisch: Es werden übergebene Summen kumuliert. Bei einer bestimmten Summe wird etwas weitergegeben.	<u>Nichtlineare Kennlinie</u> Aufgrund des skalaren Outputs und der vielen vektoriellen Inputs sind In- und Output nicht proportional.
<u>Synapsen verändern Input</u> Die Synapsen nehmen die Gewichtung vor. Sie verändern den Wert der ankommenden Information beziehungsweise den ursprünglichen Impuls.	<u>Skalarer Output</u> Es gibt nur einen Output, der beim nächsten Neuron wieder einen Vektor darstellt. In einem Neuron werden mehrere Vektor-Inputs zu einem skalaren Output.	<u>Gewichte einstellbar</u> An den Synapsen variieren die Gewichte, damit das Netz variabler ist.

### 5.2.2.3 Gewichte

Die schon häufig angesprochenen Gewichte haben drei mögliche Auslegungen:

- positives Gewicht: das Neuron x wirkt bestärkend auf ein anderes Neuron
- negatives Gewicht: das Neuron x wirkt hemmend auf ein anderes Neuron
- Gewicht = null: kein Einfluss des Neurons x auf ein weiteren Neuron

Die Gewichtsverteilung macht daher auch das Wissen eines Netzes aus. Daher wird beim Lernen auch die Gewichtsverteilung verändert<sup>7</sup> (neuronaalesnetz.de: 6).

So sieht ein technisches/künstliches Neuron aus: siehe Anhang XII.

#### 5.2.2.4 Aktivität

Die Eingabe Funktion (engl. Input Function) hat folgende Formel:

$$in_i = \sum_{j=0}^n w_{j,i} a_j \quad \text{Abb. 1}$$

Der Wert der Eingabefunktion stellt die Summe aus den Werten der Neuronen, die ein Signal/einen Wert geschickt haben, beziehungsweise die Summe der Werte des Input-Layers des Neuronalen Netzes <sup>6 (ebd.)</sup> dar.  $w_j$  stellt hierbei die Gewichtung der Synapse dar und  $a$  ist der Ausgangswert.

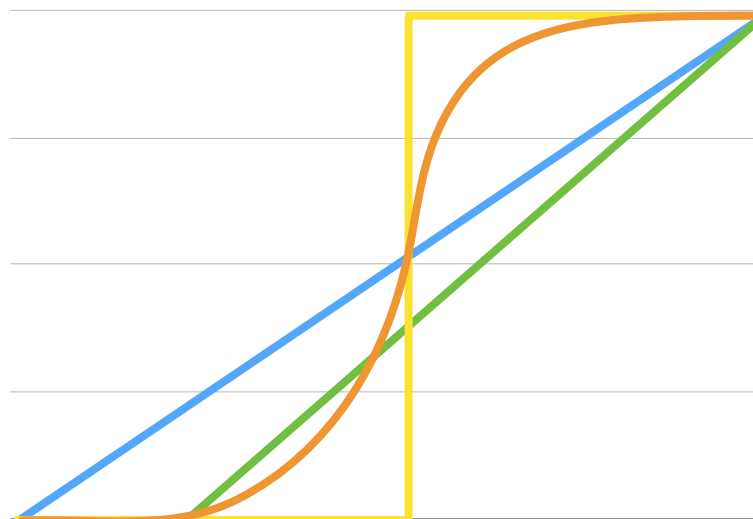
Hier ist die Aktivierungsfunktion (engl. Activation Function):

$$a_i = g(in_i) = g\left(\sum_{j=0}^n w_{j,i} a_j\right) \quad \text{Abb. 2}$$

Diese Funktion ist dafür zuständig, den Wert zu berechnen, der dann gegen den Schwellwert geprüft wird. Abhängig vom Ergebnis gibt das Neuron dann ein positives Gewicht an das nächste Neuron weiter (Bias Weight). <sup>6 (ebd.)</sup>

Dieser Wert, der gegen den Schwellwert geprüft ist, wird mit der Aktivierungsfunktion ausgerechnet und beschreibt die Aktivität und die Inputs des Neurons. Man kann sie in einem 2-dimensionalen Graphen darstellen. Die x-Achse stellt hierbei die Aktivität dar während die y-Achse die Anzahl der Input zeigt. Man unterscheidet zwischen insgesamt vier verschiedene Aktivitätsfunktionen, die sich in der Aktivierungsfunktion als  $g(x)$  widerspiegelt:

- Lineare Aktivitätsfunktion
- Lineare Aktivitätsfunktion mit Schwelle
- Schwellwertfunktion (Signumsfunktion)
- Sigmoide Aktivitätsfunktion



Die für kognitive Prozesse verwendete Aktivierungsfunktion der Aktivitätsfunktion (siehe Abbildung 2) ist die Sigmoidale Aktivierungsfunktion<sup>7</sup> ([neuronalesnetz.de](http://neuronalesnetz.de): 9). Die Formel für die Sigmoidale Funktion ist wie folgt<sup>6</sup> (Zöllner 2014: 11):

$$g(x) = \frac{1}{(1 + e^{-x})} \quad \text{Abb. 4}$$

Im Gegensatz zu den anderen Aktivierungsfunktionen ist die Sigmoidale Aktivierungsfunktion flexibler (siehe Anhang XIII). Jedoch ist ein Nachteil der Sigmoiden Funktion, dass sie, je weiter sie sich vom ursprünglichen Wert (dem Schwellwert) ist, nicht mehr viel hinzulernt. Daher sieht man auch am Graphen, dass sich die Funktion asymptotisch für  $x = +\infty$  an  $y = 1$  nähert (siehe Anhang XIII).

#### 5.2.2.6 Training

---

Abhängig von der Lernart<sup>(siehe 6.1: 13)</sup> werden die Neuronen und das Neuronale Netz unterschiedlich gespeichert und es werden ihre Ergebnisse geprüft. Das Training ist dafür da, dass das Ergebnis des Netzes überprüft und dahingehend verändert wird, damit es stimmt; entweder anhand einer Rückmeldung des Nutzers, einer vergebenen Lösung, einer automatisch vorgegebenen Lösung oder teilweise auch selbstständig. Dies hängt auch von der Lernregel und Trainingsphase ab.

Während ein Neuronales Netz trainiert wird, werden Fehler behoben. Dieses stellt sich in der Umverteilung der Gewichte dar.

#### 5.2.2.5 Lernregeln

---

Es gibt diverse Lernregeln (auch Lernverfahren<sup>4</sup> (vgl. Kriesel 2007: 41)) für ein neuronales Netz. Die Lernregeln sind wichtig für die Justierung der Gewichte. Sie sind Teil des Trainings. Die durch die Lernregeln erzielten Ergebnisse werden in der Trainingsphase dann immer wieder überprüft und werden zur Veränderung der Gewichte genutzt. Die bekanntesten Lernregeln sind: die Herb-Regel, die Delta-Regel, Perceptron, Backpropagation, Competitive Learning, Adeline und Kohonen.<sup>7</sup> (vgl. [neuronalesnetz.de](http://neuronalesnetz.de): 26):

	Biologische Grundlage	Funktionsweise	Vorteile	Nachteile	Lernphase
<b>Herb-Regel</b>	teilweise	Aktivität zweier verbundenen Neuronen führt zu einer Veränderung der Gewichte.	biologisch erklärbar, einfach	nicht sehr belastbar	alles
<b>Delta-Regel</b>	eher nicht	Das Ergebnis der Neuronen eines Layers wird mit dem erwarteten Ergebnis verglichen und im Anschluss wird das Gewicht proportional zu dieser Differenz verändert.	einfache Implementierung	geringe Mächtigkeit	überwachtes Lernen
<b>Perceptron</b>	mathematisches Modell eines Neurons <sup>9</sup> (vgl. <a href="http://cs.stanford.edu">cs.stanford.edu</a> )	Perceptronen lernen indem sie den Fehler zwischen dem Ergebnis und der Erwartung/dem Ziel und der Lernrate betrachten. <sup>8</sup> (vgl. Souza 2016: 105)	eines der ältesten und einfachsten Verfahren	es wurde unter anderem von Adaline überholt, da es eine geringe Mächtigkeit hat <sup>10</sup> (vgl. Raschka 2013-2017)	überwachtes Lernen
<b>Backpropagation</b>	eher nicht	Erweiterung von Perceptron: Der Fehler, der an der Ausgabeschicht berechnet wird, wird zurückgeleitet und anhand von ihm wird das Netz dann verändert.	große Mächtigkeit, mehrere Layer möglich	Unübersichtlichkeit durch Komplexität und viele Layers	überwachtes Lernen
<b>Competitive Learning</b>	teilweise	Der beste Inputwert wird zum Outputwert und die betroffenen Gewichte werden so geändert, dass sie dem Input näher kommen.	einziges Verfahren für unüberwachtes Lernen	keine Kategorien mehr, „Vorherrschaft“ eines Neurons	unüberwachtes Lernen
<b>Adaline</b>	teilweise	Erweiterung von Perceptron: Mit dem Input werden ständig Werte vorhergesagt, die dann im Anschluss ausgewertet werden. <sup>10</sup> (vgl. Raschka 2013-2017)	die Größe des Fehlers ist bekannt und es gibt ständig eine Rückmeldung <sup>10</sup> (vgl. Raschka 2013-2017)	altes Single-Layer Künstliches Neuronales Netz -> nur wenige Informationen können verarbeitet werden	überwachtes Lernen
<b>Kohonen</b>	teilweise	Self Organizing Maps oder SOMs: Neuronen werden in einem Koordinatensystem angeordnet und ihre Position wird dann verschoben. Es wird im Anschluss nach den Knoten/Knubbeln geschaut.	Bild-/Tonerkennung, Kinematik, finden des kürzesten Weges zwischen zwei Punkten	nicht für kognitive Prozesse geeignet	unüberwachtes Lernen

Für Neuronale Netze, die mehrere HiddenLayers haben, bietet sich Backpropagation an, da es die größte Mächtigkeit hat und für komplexe kognitive Prozesse, die „konkrete Anwendungsprobleme lösen sollen“<sup>7</sup> (ebd. 18).

### 5.2.3 Das Rekurrente Neuronale Netz

Ein Rekurrentes Neuronales Netz ist ein Neuronales Netz bei dem die einzelnen Neuronen auch auf sich selber oder andere vorherige Neuronen zeigen (siehe Anhang XIV).

Es wird häufig für künstliche Intelligenz genutzt. Im Zusammenhang mit ANNA ist das Neuronale Netz aufgrund von Backpropagation rekurrent.

Zunächst wird der Output mithilfe des sog. Forward-Pass ausgerechnet. Dann wird der Output mit dem erwarteten Wert verglichen und es wird der Fehler bestimmt, wie bei der Perceptron Lernregel. Wenn die Fehlerquote klein genug ist, wird das Training beendet. Ist die Fehlerquote groß, wird das Netz rekurrent wieder aufgerufen. Allerdings werden zunächst vom OutputLayer über die HiddenLayers zum InputLayer die Werte verändert, bis die Fehlerquote klein genug ist. Dies nennt sich Backward-Pass.

## 5.3 VERWENDUNG

### 5.3.1 ANNAs Neuronales Netz (siehe 12. Implementierung)

#### 5.3.1.1 Layers

---

Das InputLayer von ANNA ist ein Layer, das  $n$  Neuronen hat. Die Anzahl der Neuronen hängt von der Anzahl der Informationen, die ANNA übergeben bekommt, ab. Jede übergebene Information hat ein Gewicht und jedes Neuron des Netzes hat alle Gewichte aller Informationen. Das heißt, dass man  $x$  Informationen hat und im Anschluss dann  $n = x$  Neuronen hat, die je eine Liste an  $g = x$  Gewichten haben. Das bedeutet, dass das InputLayer aus insgesamt  $x^x$  Gewichten besteht.

Die Anzahl der HiddenLayers wird mithilfe der Formel  $(\text{int}) \text{Math.pow}((\text{anzahlDerNeuronen} + 1), 0.5)$  berechnet. Diese Formel wird benutzt, da es das Gleiche ist wie  $\sqrt{(\text{anzahlDerNeuronen} + 1)}$ . Die Wurzel aus  $\text{anzahlDerNeuronen}$  plus eins bietet sich an, da sie sich abhängig von der  $\text{anzahlDerNeuronen}$  verändert und immer in einem klaren Verhältnis zu denselben ist. Außerdem ist das Netz aufgrund der HiddenLayer sehr mächtig.

Die Anzahl der Output Neuronen ist für ANNA eins, weil es darum geht, die erwartete Lösung aus allen Input Gewichten zu errechnen. Daher gilt, aus n Input-Gewichten wird eine Lösung errechnet:

$$f: \mathbb{R}^n \rightarrow \mathbb{B}^1$$

Abb. 5<sup>4</sup> (Kriesel 2007: 6)

#### 5.3.1.2 Aktivität

---

Für ANNAs Rekurrentes Neuronales Netz gilt die Sigmoidale Aktivierungsfunktion. Sie ist aufgrund der großen Differenzierbarkeit beim Trainingsalgorithmus am geeignetsten. Außerdem ist sie die am häufigsten gebrauchte Funktion für künstliche Neuronale Netze und ist daher sehr „erprobt“

#### 5.3.1.3 Training

---

Von den drei Lernarten ist das überwachte Lernen die Lernart, die ANNAs Rekurrentes Neuronales Netz nutzen wird. Gegen die anderen beiden Varianten spricht, dass beim unüberwachten Lernen zum Beispiel Informationen eines Satzes ausgelassen werden und dass ANNA keine Erwartung des Nutzers bekommt und sich somit nicht an ihn anpassen kann. Das bestärkende Lernen wiederum könnte genutzt werden, da der Nutzen der Aktion bewertet wird.

#### 5.3.1.4 Lernregel

---

Die von ANNA genutzte Lernregel ist Backpropagation. Im Bezug auf ANNA sieht Backpropagation als Pseudocode wie folgt aus<sup>11</sup> (vgl. Papagelis & Kim):

```
solange (die Epochenzahl < die maximale Epochenzahl)
    und (die Fehlerquote > gewünschte Fehlerquote)
    // forward (siehe 7.2 Umsetzung) - erstes Durchlaufen des Netzes (Perceptron)
    für jedes Layer des Netzwerk
        für jedes Neuron der Layer
            1. Summe der Input Gewichte berechnen
```

## 2. Berechne die Aktivierungsfunktion des Neuron

*// Backpropagation*

**für** jedes Neuron im OutputLayer

Berechne den Fehler des Neurons

**für** jedes Neuron im HiddenLayer

1. Berechne den Fehler des Neurons

2. Speichere ein neues Gewicht für jedes Neuron des Netzes

*// Allgemeiner Fehler*

Berechne den allgemeinen Fehler des Netzes

## 6. Umsetzung von ANNA

### 6.1 IDEE

Die Idee hinter dem Rekurrenten Neuronalen Netz von ANNA ist, dass es sich die Wahl des Nutzers und das Prinzip merkt. Dadurch lernt es und kann sich, natürlich erst nach längerem Nutzen, auf den Nutzer einstellen und ihm die Ergebnisse anders präsentieren. Dies wäre auch möglich, wenn für jedes Mal, wenn der Nutzer eine bestimmte Option anklickt, diese Option um eins an Gewicht zunimmt. Allerdings kann ANNA sich durch das Rekurrente Neuronale Netz auch ein allgemeines Bild machen und die Auswirkung der Wahl berücksichtigen. So ergibt sich zum Beispiel, dass der User Vegetarier ist, wenn er viele vegetarische Gerichte sucht, oder aus einer Auswahl an Gerichten immer die vegetarischen wählt und häufig Tierschutzseiten besucht.

Daher bekommt ANNA von dem Interface, welches dem Nutzer mehrere mögliche Aktionen anbietet, diese Aktionen und zusätzlich die tatsächliche Auswahl des Users. Die Aktionen bekommen bei der ersten Benutzung zufällige Gewichte zugewiesen. Ansonsten werden die zuletzt gespeicherten Gewichte der Aktionen genommen. Die Gewichte werden in das Rekurrente Neuronale Netz eingegeben und als Ergebnis wird die Auswahl des Nutzers vorgegeben, damit sich das Netz danach ausrichtet. Danach werden die neu errechneten Gewichte gespeichert und in der Datenbank abgelegt, damit bei jeder weiteren Suche die bereits vorgenommene Justierung wieder vorhanden ist.

### 6.2 UMSETZUNG

Für die Übertragung ins Rekurrente Neuronale Netz werden die x Möglichkeiten des User genommen und in der Variablen `wort` mit ihrer ID und ihrem Gewicht in einer Hilfsklasse `Wort` gespeichert. Jedes `Wort` wird dann in der Klasse `Woerter` gespeichert. `Woerter` ist eine `ArrayList` von `Wort`.

Die Wörter werden dann so in einer Matrix, einem zweidimensionalen `doubleArray`, gespeichert und normalisiert, dass sie als Input für das Rekurrente Neuronale Netz fungieren können. Die Matrix besteht aus 1 Zeile und x Spalten. jede Spalte enthält eine ID (siehe 12. Implementierung: `Woerter - rohdatenZuMatrix()`).



Zuvor wurde noch das Neuronale Netz initialisiert. Es besteht aus einem **InputLayer** mit x Neuronen, **Math.pow(x+1, 0.5)** **HiddenLayer** mit x Neuronen und einem **OutputLayer** mit einem **Neuron** (siehe oben).

Dann wird die Matrix als Input Werte eingespeichert, das heißt, dass jedes **Neuron** eine ID „repräsentiert“. Außerdem, falls die Werte nicht zum ersten Mal benutzt werden, wird das Gewicht der Wörter als **inputGewichtIn** festgelegt. Ansonsten werden alle Neuronen zu Beginn mit dem gleichen Gewicht ausgestattet. In diesem Fall ist es 0.1 (siehe 12. Implementierung: **Neuron - initialisiereInputNeuron(Wort wort)**).

Die Neuronen in den anderen Layern werden auch initialisiert. Allerdings ist das **inputGewichtIn** immer das Output Gewicht des Neurons der vorherigen Layer. Das eigene Output Gewicht ist jedoch ein zufälliger Wert.

Im Anschluss wird der Output festgelegt. Die Wahl des Users wird genommen und aus **woerter** wird das gleiche Wort herausgesucht. Dann wird von diesem die ID als „OutputZiel“ gespeichert.

Nachdem die Voreinstellungen vorgenommen wurden, wird das Rekurrente Neuronale Netz trainiert. Dabei wird zunächst mit Perceptronen (siehe oben) das Netz **forward**, also in eine Richtung (Richtung Output Layer), gerechnet und anschließend rekursiv mit Backpropagation überarbeitet und justiert.

Zwischendurch wird, auch zur Kontrolle, immer der Zwischenstand in der Konsole ausgegeben. Danach wird geprüft, ob das Ergebnis des Netzes der gewünschte Output ist. Dies sollte in der Regel so sein. Dann werden die Gewichte der letzten Output Layer den Wörtern aus **woerter** zugewiesen und dann in der Datenbank gespeichert, damit sie bei der nächsten Verwendung nutzbar sind.

## 6.3 TEST

Um den zuvor ausgeführten Ansatz zu überprüfen, nahm ich ein Beispielergebnis, welches mir in einer Datenbank zur Verfügung stand. Die den Daten zugrunde liegende Suchanfrage war „Ich suche ein Haus in Bonn.“. Das Ergebnis auf diese Anfrage war die verwendete Tabelle (siehe Anhang XVII). Da die Daten lediglich zur Probe gedacht waren, wandelte ich

die Tabelle aus der Datenbank zunächst in eine csv-Datei um. Dabei speicherte ich nur die Schlagwörter in der Datei (siehe Anhang XVII).

Dann lies ich mein Programm zum Test laufen. Zunächst wurden die Schlagwörter einzeln zu einem `wort` in `woerter` (siehe Implementierung: `Woerter.java`). Sie bekamen eine ID und ein Gewicht zu gewiesen. Im Anschluss initialisierte sich das Rekurrente Neuronale Netz und trainierte mit den Daten.

Zur Probe suchte ich in den Ergebnissen zunächst nach „Bonn“, welches unter den Schlagworten vertreten ist. Das Rekurrente Neuronale Netz hatte auch immer ein Output Neuron mit einem Ergebnis. Allerdings war das Ergebnis des Neurons nicht dasselbe, wie die ID des gewünschten Ergebnis, also Bonn, sondern, wie sich nach mehrfachem Testen herausstellte, ein rein zufälliger Wert.

## 6.4 ÄNDERUNGEN

Daraufhin beschäftigte ich mich erneut intensiv mit dem Ergebnis des Rekurrenten Neuronalen Netzes. Es stellte sich heraus, dass das Ergebnis des Netzes eine Zahl zwischen eins und null sein muss. Zuvor versuchte ich, das Netz mit Zahlen zwischen null und hundert zu erstellen. Doch nun veränderte ich dies. Also werden die Gewichte mit zufälligen Zahlen zwischen null und eins initialisiert und die ID wird während der Übertragung in die Matrix „normalisiert“. Das bedeutet, dass der eigentliche Wert durch hundert geteilt wird. Dies ist zwar nicht stabil für große Datenmenge, reichte jedoch vorerst für das Testbeispiel.

Außerdem wird das Ergebnis des Netzes nicht mehr gegen den gewünschten Output - die Id - geprüft sondern das Ziel ist, dass das Ergebnis der eins möglichst nahe kommt. Trotzdem blieb der gewünschte Output die ID dessen, was der Nutzer ausgewählt hat. Das Ergebnis des Rekurrenten Neuronalen Netzes ist nämlich im Idealfall eine Bestätigung, dass das Netz den vorhergesagten Wert ermittelt hat, also zum gleichen Ergebnis gekommen ist, nachdem es trainiert hat, oder nicht. Wenn das Ergebnis eins ist, hat das Netz den vorhergesagten Wert gefunden wohingegen null das Gegenteil besagt.

## 6.5 ZWEITER TEST

Ich wiederholte den Test vom ersten Mal und erhielt bei der Auswahl „Bonn“ dieses Mal den Wert 0.9619680730994274 (siehe Anhang XVII). Dieser ist der eins sehr nahe. Um sicher

zu gehen, dass es nicht erneut ein zufälliger Wert ist, wiederholte ich den Test mit Bonn und erhielt erneut 0.9619680730994274. Daher versuchte ich es mit anderen Wörtern und erhielt folgende Ergebnisse:

Auswahl	Ergebnis
Bonn	0.9619680730994274
Friedrich-Wilhelms-Universität	0.9670255664315609
Wissen	0.8947317284670677
Bildung	0.9202232163700961
Kultur	0.9509324176063766

Die Ergebnisse waren sehr zufriedenstellend. Die Nähe zur eins zeigt, dass sich das Rekurrente Neuronale Netz nach dem gewünschten Ziel ausgerichtet hat. Die Abweichung von der eins lässt sich mit der maximalen Epochenzahl erklären. Diese ist jedoch notwendig, damit der das Netz nicht überlastet.

## 6.6 ERGEBNIS

Das vollständige UML-Diagramm sieht wie folgt aus: siehe Anhang XV. Der gesamte Quelltext findet sich unter 12. Implementierung. Die Grundstruktur des Neuronales Netzes ist mithilfe von „Neural Network Programming in Java“<sup>8</sup> (vgl. Souza 2016) implementiert worden. Der Quelltext ist zum Nachvollziehen und besseren Verständnis ausführlich kommentiert.

## 6.7 PROBLEMATIK

Bei der Implementierung gab es mehrere Probleme. Zunächst stellte sich die Frage, welche Informationen in die Neuronen gegeben werden. Zunächst gab ich jedem Neuron eine Liste mit allen Gewichten mit und mein Output Ziel war ein bestimmtes Gewicht. Jedoch stellte sich dieser Ansatz als falsch heraus, da nach einem Ergebnis gesucht wurde und die Gewichte sich dementsprechend ändern sollten. Daher konnte ein Gewicht nicht das Output Ziel des Netzes darstellen.

Es konnte auch nicht funktionieren, weil jedes Neuron die Liste aller Gewichten als Input bekommen hat. Dies konnte zu keinem Ergebnis führen, weil somit die Ausgangsbedingungen für alle gleich waren. Dieses Problem ist eines der größten bei Rekurrenten Neuronales

Netzen. Es heißt „Garbage in, garbage out!“<sup>12</sup> (Donalek 2011: 27) und stellt insbesondere beim Testen eine große Schwierigkeit dar.

Ein weiteres größeres Problem stellte die Matrix dar. Da ein Neuronales meist mit großen Datensätzen und Tabellen erstellt wird, war auch die Handhabung der Matrix mit einem Input, der ID schwierig.

Das letzte größere Problem war ein allgemeines Problem des Rekurrenten Neuronales Netzes. Bei der Implementierung des Netzes muss beachtet werden, wie mit den Daten umgegangen wird. Die Daten müssen so ausgeschöpft werden, dass ein zufriedenstellendes Ergebnis erreicht werden kann. Doch gilt es zu beachten, dass die Daten nicht zu detailliert ausgewertet werden. Dies nennt man Over-fitting. Sind zu viele Eingaben vorhanden, dann rechnet das Netz zu viel aus und das Ergebnis hat nichts mehr mit den Trainingsdaten zu tun<sup>11</sup> (vgl. Papageis & Kim). Dies geschieht unter anderem auch, wenn zu viele HiddenLayer vorhanden sind.

Wenn die Daten allerdings nicht genug ausgeschöpft werden, dann geschieht das Gegenteil. Dieser Prozess wird Under-fitting genannt. Die Neuronen können aufgrund mangelnder Komplexität die Trainingsdaten nicht richtig vertreten. Dies geschieht bei einem Mangel an HiddenLayer (siehe Anhang XVI). Daher war insbesondere die Anzahl an HiddenLayer ein Problem. Nach einiger Recherche stellte sich dann heraus, dass  $\text{Math.pow}(\text{anzahlDerNeuronen}, 0.5)$  sich am Besten eignet (siehe 6.3.1.1 Layers).

Dasselbe Problem ergab sich auch bei den Neuronen des HiddenLayer. Doch funktionierte auch in diesem Fall der Ansatz mit der Wurzel aus der Anzahl der Neuronen aus denselben Gründen sehr gut.

## 7. ANNAs Oberfläche

### 7.1 PRINZIP

Da die Künstliche Intelligenz ANNAs für die Verteilung der Gewichte und die dadurch entstehenden persönlichen Präferenzen zuständig ist, wird sie eine Bibliothek / einen externen Dienste für die Weboberfläche und Datenbankbindung nutzen. Es handelt sich hierbei um atpSQL, eine Content-Management-Plattform mit Anwendungsintegration.

### 7.2 UMSETZUNG

Im Internet (<http://www.anna.is.it>) kann nun über ein Suchfeld eine Suchanfrage gestellt werden. Diese Suchanfrage kann aus Schlagworten, wie „Bonn“ oder auch aus Sätzen bestehen: „Finde Gymnasien in Bonn“, oder „Ich suche Gymnasien in Bonn“.

Der eingegebene Text wird zunächst von LanguageTools (siehe 11. Implementierung /src/org/anna/lang/...) analysiert. LanguageTool filtert den Text auf Substantive und die gefilterten Worte werden in der Grundform an die Wissenstabelle weiter gegeben. Zunächst werden die Wörter getrennt und einzeln überprüft. Aus dem Satz „Suche Gymnasien in Bonn“ analysiert LanguageTool:

- Suche = entweder Suche (Substantiv) oder suchen (Verb)
- Gymnasien = **Gymnasium** (Substantiv)
- in = in (Präposition)
- Bonn = **Bonn** (Eigenname)

Die relevanten Begriffe sind nun Gymnasium und Bonn, da sie die einzigen Substantive sind und, da Suche aus dem Schema rausfällt, weil es ein Verb sein kann (siehe Anhang XIX).

Wenn man sehen möchte, wie ANNA durch LanguageTool einen Satz versteht, so muss man nur ein „#“ hinter die Anfrage schreiben (siehe Anhang XIX). Es gibt auch ein paar Sonderfälle zu beachten. Da der gesuchte Begriff, zum Beispiel, ein zusammengesetzter Begriff sein kann, kann er einen Gedankenstrich im Namen haben (zum Beispiel: „Nordrhein-Westfalen“). Hier muss der eingegebene Text am „-“ Zeichen getrennt werden. Ansonsten wird anhand der Leerzeichen getrennt.

Die Wissenstabelle setzt sich aus knapp 500.000 Datensätzen zusammen, die von dmoz (<http://dmoz.de>) und auch destatis (<https://www.destatis.de>) zusammengestellt und durch Analysen aufbereitet wurden. Beide Anbieter boten die Daten kostenlos an und dmoz wurde bis vor kurzem auch von Google benutzt. Der Vorteil, den die Daten bieten, ist, dass sie von Menschen klassifiziert wurden und deutschsprachig sind. Jedoch sind die Daten von dmoz mittlerweile veraltet.

Ein Datensatz der Wissenstabelle besteht aus mehreren Komponenten. Er hat:

- den Begriff - den Name des Datensatzes (Es kann ein Wort sein oder mehrere Wörter subsumieren, wie zum Beispiel „Ernst-Moritz-Arndt Gymnasium“.)
- den oText - das heißt, eine Beschreibung in html
- den Link - wenn der Datensatz zu einer anderen Internetseite führt
- den Begriffsumfang - eine empirische Zahl, die beschreibt, wie viele Begriffe unter diesem Begriff zusammengefasst werden
- einen Begriffsinhalt - eine Zahl, die reziprok zum Begriffsumfang ist und die Präzision/ Genauigkeit des Begriffs beschreibt
- eine Klassifikation - eine intern angelegte Klassifikation des Begriffs
- Schlagworte - Wörter, die im Zusammenhang stehen und über die der Datensatz auch gefunden werden kann

Die nach der LanguageTool Analyse übergebenen Worte werden nun einzeln in der Wissenstabelle gesucht. Sie werden generisch im Link gesucht, das heißt, dass das gesuchte Wort im Link vorkommen muss. Aber in den Schlagworten, dem oText, und in dem Begriff werden sie nur halbgenerisch gesucht, das heißt, dass das Wort entweder am Satzanfang oder am Satzende sein muss (siehe 11. Implementierung [SearchEngine.java](#)).

Nach dem Absenden der Suchanfrage und der Analyse von LanguageTool wird zwischen drei Fällen unterschieden:

1. nicht verstanden (LanguageTool kennt Teile der Eingabe nicht und die Wörter sind nicht in der Datenbank vorhanden)
2. nicht gefunden (LanguageTool kennt zwar alle Wörter, aber in der Datenbank ist kein Ergebnis zu diesen vorhanden)
3. gefunden (LanguageTool hat den Text verstanden und etwas dazu gefunden)

Wenn die Eingabe nicht verstanden wurde, gibt ANNA die Ergebnisse ihrer Analyse aus und weist auf das unbekannte Wort hin. In Zukunft soll es dann möglich sein ANNAs Wissen zu erweitern (siehe Anhang XIX).

Wenn ANNA jedoch alles verstanden hat, aber keine Suchergebnisse in der Datenbank gefunden hat, dann gibt sie, im Gegensatz zu anderen Suchmaschinen, aus, dass sie keine Ergebnisse gefunden hat. Sie gibt nicht ähnliche Ergebnisse zurück, sodass der Nutzer nicht Zeit verliert und gezielt suchen kann. In Zukunft soll es auch hier möglich sein ANNAs Wissen zu erweitern (siehe Anhang XX).

Es gibt auch den Fall, dass ANNA Ergebnisse zu der Suche findet. Diese Ergebnisse gibt sie dann nach dem Begriffsinhalt sortiert aus, da dieser mit der Genauigkeit auch das Gewicht repräsentiert. Neben den Ergebnissen, von denen ANNA nur die ersten 128 anzeigt, damit es nicht unübersichtlich wird, gibt es auch noch die Filterfunktionen. Diese setzen sich aus den vorkommenden Klassifikationen aller gefundenen Suchergebnissen und deren Schlagworten zusammen, sodass sich durch das Auswählen von Filtern die Auswahl der Suchergebnisse anpasst und verändert. Dadurch können dann alle Suchergebnisse angesprochen werden (siehe Anhang XXI).

Zuletzt gibt es noch die Filterfunktionen, die der Nutzer vor Auswahl eines Suchergebnisses, anwenden kann. Es gibt insgesamt drei verschiedene Filter (hier sind es zum Beispiel die ersten Filter, die zur Suche „Bonn“ zur Auswahl stehen):

#### Filter (Klassen)

- atpsql|geo|ort|kreisangehörige\_gemeinde (1)
- atpsql|geo|ort|kreisfreie\_stadt|großstadt (1)
- atpsql|geo|ort|stadt (1)
- atpsql|geo|region|gemeindevverband (1)
- atpsql|geo|region|kreis (1)
- atpsql|atom|schlagwort (2)
- www|html|home|content (1652)

Zunächst gibt es die Klassen, die gefiltert werden können. Diese spiegeln die innere Struktur wieder und lassen Zusammenhänge zum Begriff erkennen.

#### Filter (Symbol)

-  icon-atom (2)
-  icon-gps (5)
-  icon-internet (1652)

Es gibt Symbole, welche die Art der Information anzeigen und beschreiben, was nach dem Auswählen geschieht. Das Atom weist auf einen Grundbegriff des Wissensspeichers hin. Die Links / Internetquellen sind mit der Weltkugel gekennzeichnet und die geographischen Daten werden durch das GPS-Symbol gekennzeichnet. Sie sollen in Zukunft eine direkte Routenplanung zulassen.

#### Filter (Schlagworte)

- Abstrakt (2)
- Afrika (2)
- Ahrweiler (2)
- Altenpflege (2)
- 

Die Schlagworte sind die Schlagworte, die die Begriffe definieren. Durch additives Anklicken werden so die Suchergebnisse teilweise drastisch verkleinert.

Das Filtern soll das Ergebnis heterogener machen. Wenn man „Nordrhein-Westfalen“ sucht, dann bekommt man 37500 Treffen. Wenn man jedoch spezifischer sucht, zum Beispiel nach „Gymnasien in Bonn“, dann hat man nur noch 11 Ergebnisse und durch die Einschränkung MINT hat man noch drei Ergebnisse und bei Informatik gibt es nur noch ein Ergebnis.

### 7.3 GEWÜNSCHTES ERGEBNIS

Als Ergebnis hätte ich mir gewünscht, dass alle Suchergebnisse und das ausgewählte Ergebnis an die Künstliche Intelligenz übergeben und von dieser gewichtet werden. Sie sollte die Gewichte neu arrangieren und im Anschluss an den Wissensspeicher übergeben, damit diese dann bei den zugehörigen Begriffen gespeichert werden können. Dadurch sollen die Suchergebnisse bei der nächsten Suche in einer anderen - besseren - Reihenfolge auftauchen.

### 7.4 RESULTAT

Die von mir durchgeführte Implementierung veranschaulicht die innerhalb meiner Facharbeit erarbeiteten Ansätze und bietet die Möglichkeit im Internet die Suchmaschine auszuprobieren. Momentan sind im Verhältnis zu anderen Suchmaschinen relativ wenige Daten vorhanden. Jedoch lässt sich durch die Filterfunktion im Vergleich zu anderen Suchmaschinen die Suche besser eingrenzen. Außerdem kann man einen ersten Eindruck der Personalisierung gewinnen.



## 8. Zusammenfassung

### 8.1 AUFWAND

Ich habe insgesamt viel Arbeit in die Facharbeit investiert, da ich zuvor noch kein so großes Projekt realisiert habe und daher auch mein erstes grösseres Projekt mit Erfolg zu Ende bringen wollte.

Die Programmierung hat aufgrund der Größe und Komplexität des Projekts viel Zeit eingenommen. Außerdem hat das Überprüfen und die Fehlerfindung/-korrektur, da mir kein erprobten Algorithmen für meine Thematik zur Verfügung standen, ebenfalls lange gedauert.

Die größte Herausforderung stellte für mich die Struktur des Rekurrenten Neuronalen Netzes dar. Es war zunächst schwer vorstellbar, wie die mathematischen Formeln, auf die Gewichte angewandt, am Ende zu einem sinnvollen Ergebnis kommen sollen.

Da Rekurrente Neuronale Netze sehr komplex sind und auch viel mit komplexen mathematischen Formeln und Ansätzen zu tun haben, musste ich mich auch intensiv mit diesen Grundlagen beschäftigen. Das Nachvollziehen der mathematischen Grundlagen fiel mir mitunter auch am schwersten.

### 8.2 BEWERTUNG DES ERGEBNISSES

Grundsätzlich bin ich mit meinem Ergebnis zufrieden. Mein Ziel war es einen adaptiven Assistenten mithilfe eines Rekurrenten Neuronalen Netzes zu implementieren. Die Rahmenbedingungen zur endgültigen Fertigstellung sind mittlerweile gegeben. Es existiert sowohl das Neuronale Netz als auch die Website mit den Filterfunktionen und der Suchfunktion. Auch eine Datenbank gibt es schon, auch wenn diese noch nicht sehr groß und ein bisschen veraltet ist.

Allerdings muss die Anbindung des Neuronalen Netzes noch erfolgen. Beide Teile arbeiten für sich, aber leider ist die Personalisierung durch das Neuronale Netz noch nicht in Verbindung mit der Datenbank und den Suchergebnissen vorhanden.

Auch die Authentifizierung, die für die Personalisierung essentiell ist, muss noch implementiert werden. Es fehlen auch noch weitere Features von ANNA, so wie das

Hinzufügen von neuem Wissen oder, dass sich die Filter anpassen, wenn ein Filter gewählt wurde. Wenn möglich wäre es auch schon, wenn ANNA noch kürzere Laufzeiten und eine aktuellere Datenbank hat.

Auf der anderen Seite ist mein Ergebnis ein neuer Ansatz, welcher auch schon durch die Filter-Option ein Stück weit deutlich wird. Diese gibt dem Nutzer eine viel größere Möglichkeit den Assistenten effektiv zu nutzen, da er - im Gegensatz zu anderen Assistenten - nicht nur Suchergebnisse bekommt, die nach den Interessen von Lobbyisten und großen Firmen ausgerichtet sind, sondern Suchergebnisse, die sein persönliches Interesse widerspiegeln. Die Zahl der Ergebnisse reduziert sich so auch erheblich.

Ein weiterer Vorteil des Prinzip von ANNA ist, dass die Justierung der Gewichte in einem Rekurrenten Neuronalen Netz stattfindet. Es wäre auch möglich gewesen, die Gewichte durch einfaches aufaddieren zu verändern. Das heißt, dass für jedes Mal, das der Nutzer eine bestimmte Option anklickt, diese Option um einen an Gewicht zunimmt. Allerdings würden bei diesem Ansatz die anderen Optionen und das Verhalten des Nutzers nicht mit einbezogen.

Ich hatte mir das Ziel gesetzt, ANNA zu implementieren, jedoch verlor ich zunächst ein bisschen den Mut, als ich mir der Komplexität der Aufgabenstellung immer bewusster wurde. Glücklicherweise kam ich immer besser mit dem Thema zurecht und habe durch das Projekt viel Erfahrung gewonnen. Zudem hatte ich Spaß bei der Erstellung der Facharbeit und Implementierung. Schade ist lediglich, dass ich aufgrund der zu Verfügung stehenden Zeit noch nicht ganz fertig geworden bin. Allerdings bin ich sehr motiviert und gewillt, ANNA noch weiter zu implementieren und der Öffentlichkeit zu präsentieren.

## 9. Literaturverzeichnis

1. Martin, James A. (05.10.2016): *Virtual assistant faceoff: Alexa, Cortana, Google Assistant and Siri*, erhältlich auf <http://www.cio.com/article/3127736/software/virtual-assistant-faceoff-alexa-cortana-google-assistant-and-siri.html#slide1> (Letzter Zugriff: 21.02.2017, 16:06 Uhr)
2. Lezius, Wolfgang (Universität Paderborn) (01.08.1998): *Die Wortklassensysteme von Morphy (Vollständiges Klassensystem, großes und kleines Tag Set)*, erhältlich auf <http://morphy.wolfganglezius.de/content/2-download/wklassen.pdf> (Letzter Zugriff: 21.02.2017, 16:06 Uhr)
3. [wikidot.com](http://www.wikidot.com) (07.09.2014): *Using chunks*, URL: <http://wiki.languagetool.org/using-chunks> (Letzter Zugriff: 21.02.2017, 16:06 Uhr)
4. Kriesel, David (2007): *Ein kleiner Überblick über Neuronale Netze*, erhältlich auf [dkriesel.com](http://dkriesel.com) (Letzter Zugriff: 21.02.2017, 16:06 Uhr)
5. Vortrag von Prof. Dr. Jürgen Schmidhuber ([idsia.ch](http://idsia.ch)) (09.05.2016): „*Künstliche Intelligenz wird alles ändern*“, erhältlich auf <https://www.youtube.com/watch?v=rafhHIQgd2A> (Letzter Zugriff 21.02.2017, 15:58 Uhr)
6. Bachelor-Thesis von Barbara Zöllner: *Extraktion von Regeln aus neuronalen Netzen*, Juli 2014
7. Druckversion der Internetseite [www.neuronalesnetz.de](http://www.neuronalesnetz.de) (Veröffentlichung unbekannt): *Neuronale Netze Eine Einführung*, letzter Zugriff 13.02.2017, 18:42 Uhr
8. ePub von Souza, Alan: *Neural Network Programming with Java*, 2016
9. [cs.stanford.edu](http://cs.stanford.edu) (Veröffentlichung unbekannt): *Neural Networks-The perceptron*, URL: <https://cs.stanford.edu/people/eroberts/courses/soco/projects/neural-networks/Neuron/index.html> (Letzter Zugriff 28.02.2017, 17:17 Uhr)
10. Raschka, Sebastian (2013-2017): *What is the difference between a Perceptron, Adaline, and neural network model?*, URL: <https://sebastianraschka.com/faq/docs/diff-perceptron-adaline-neuralnet.html> (Letzter Zugriff: 28.02.2017, 17:36)

11. Papagelis, Anthony J. & Kim, Dong Soo (Veröffentlichung unbekannt): *Multi-Layer Perceptron*, URL: <http://www.cse.unsw.edu.au/~cs9417ml/MLP2/> (Letzter Zugriff: 01.03.2017, 19:25)
12. PowerPoint von Donalek, Ciro (04.2011): *Supervised and Unsupervised Learning*, URL: [http://www.astro.caltech.edu/~george/aybi199/Donalek\\_Classif.pdf](http://www.astro.caltech.edu/~george/aybi199/Donalek_Classif.pdf) (Letzter Zugriff: 28.02.2017, 19:06)

# 10. Anhang

## I. Wertehierarchie

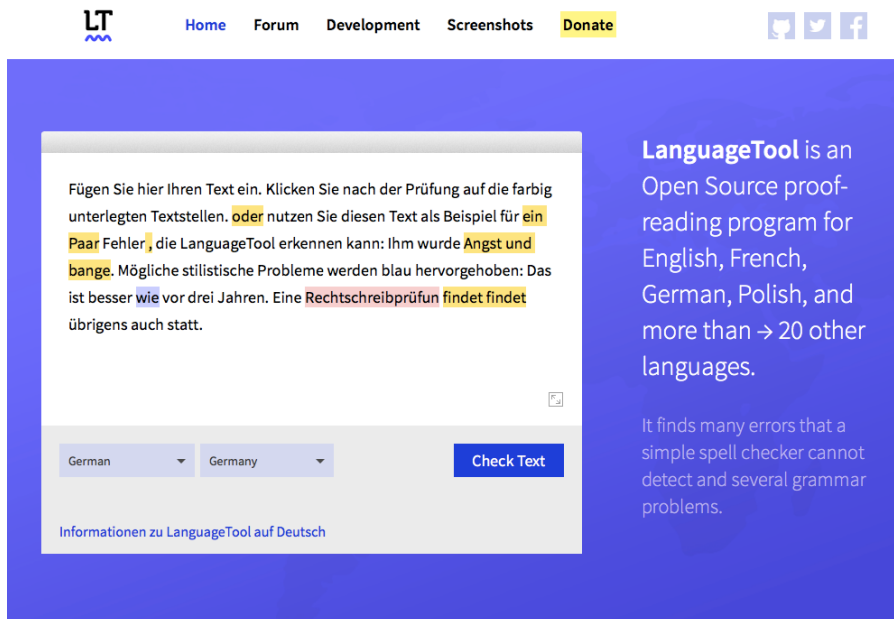
Beispiel: Ermittelte Wertehierarchie der Suchquellen

Nutzer 1	Nutzer 2
Duden	Wikipedia
Chefkoch	kochbar
Essen und Trinken	Brigitte
...	...

Beispiel der Suchergebnisse bei dem Suchwort „Apfelkuchen“

Nutzer 1	Nutzer 2
Duden   Apfelkuchen   Rechtschreibung, Bedeutung, Definition <a href="http://www.duden.de/rechtschreibung/Apfelkuchen">www.duden.de/rechtschreibung/Apfelkuchen</a>	Apfelkuchen - Wikipedia <a href="https://de.wikipedia.org/wiki/Apfelkuchen">https://de.wikipedia.org/wiki/Apfelkuchen</a>
Apfelkuchen Rezepte   Chefkoch.de <a href="http://www.chefkoch.de/re/s0/apfelkuchen/Rezepte.html">www.chefkoch.de/re/s0/apfelkuchen/Rezepte.html</a>	Apfelkuchen Rezepte - kochbar.de <a href="http://www.kochbar.de/rezepte/apfelkuchen.html">www.kochbar.de/rezepte/apfelkuchen.html</a>
Apfelkuchen: die besten Rezepte - [ESSEN & TRINKEN] <a href="http://www.essen-und-trinken.de/apfelkuchen">www.essen-und-trinken.de/apfelkuchen</a>	Apfelkuchen - die besten Rezepte   BRIGITTE.de <a href="http://www.britte.de/rezepte/.../apfelkuchen---die-besten-rezepte-10626016.html">www.britte.de/rezepte/.../apfelkuchen---die-besten-rezepte-10626016.html</a>
...	...

## II. LanguageTool Website



languageTool.org

## III. Example.class von LanguageTool

```
/* LanguageTool, a natural language style checker
 * Copyright (C) 2013 Daniel Naber (http://www.danielnaber.de)
 *
 * This library is free software; you can redistribute it and/or
 * modify it under the terms of the GNU Lesser General Public
 * License as published by the Free Software Foundation; either
 * version 2.1 of the License, or (at your option) any later version.
 *
 * This library is distributed in the hope that it will be useful,
 * but WITHOUT ANY WARRANTY; without even the implied warranty of
 * MERCHANTABILITY or FITNESS FOR A PARTICULAR PURPOSE. See the GNU
 * Lesser General Public License for more details.
 *
 * You should have received a copy of the GNU Lesser General Public
 * License along with this library; if not, write to the Free Software
 * Foundation, Inc., 51 Franklin St, Fifth Floor, Boston, MA 02110-1301
 * USA
 */
package org.languagetool.clientexample;

import org.languagetool.JLanguageTool;
import org.languagetool.Language;
import org.languagetool.Languages;
import org.languagetool.rules.RuleMatch;

import java.io.IOException;
import java.util.List;

/**
 * A simple interactive test to see if the languages are available and don't crash.
 */
public class Example {

    public static void main(String[] args) throws IOException {
        List<Language> realLanguages = Languages.get();
        System.out.println("This example will test a short string with all languages known to
LanguageTool.");
        System.out.println("It's just a test to make sure there's at least no crash.");
        System.out.println("Using LanguageTool " + JLanguageTool.VERSION + " (" +
JLanguageTool.BUILD_DATE + ")");
        System.out.println("Supported languages: " + realLanguages.size());
        for (Language language : realLanguages) {
```

```

    JLanguageTool langTool = new JLanguageTool(language);
    String input = "And the the";
    List<RuleMatch> result = langTool.check(input);
    System.out.println("Checking '" + input + "' with " + language + ":");
    for (RuleMatch ruleMatch : result) {
        System.out.println("    " + ruleMatch);
    }
}
}
}
}
}

```

## IV. Ergebnis von Example.class

```

This example will test a short string with all languages known to LanguageTool.
It's just a test to make sure there's at least no crash.
Using LanguageTool 3.6 (2016-12-28 09:41)
Supported languages: 4
Checking 'And the the' with German (Germany):
    GERMAN_SPELLER_RULE:0-3:Möglicher Rechtschreibfehler gefunden
    GERMAN_WORD_REPEAT_RULE:4-11:Möglicher Tippfehler: ein Wort wird wiederholt
Checking 'And the the' with German:
    GERMAN_WORD_REPEAT_RULE:4-11:Möglicher Tippfehler: ein Wort wird wiederholt
Checking 'And the the' with German (Austria):
    GERMAN_SPELLER_RULE:0-3:Möglicher Rechtschreibfehler gefunden
    GERMAN_WORD_REPEAT_RULE:4-11:Möglicher Tippfehler: ein Wort wird wiederholt
Checking 'And the the' with German (Swiss):
    GERMAN_SPELLER_RULE:0-3:Möglicher Rechtschreibfehler gefunden
    GERMAN_WORD_REPEAT_RULE:4-11:Möglicher Tippfehler: ein Wort wird wiederholt

```

## V. Modifizierung von Example.class

```

/* LanguageTool, a natural language style checker
 * Copyright (C) 2013 Daniel Naber (http://www.danielnaber.de)
 *
 * This library is free software; you can redistribute it and/or
 * modify it under the terms of the GNU Lesser General Public
 * License as published by the Free Software Foundation; either
 * version 2.1 of the License, or (at your option) any later version.
 *
 * This library is distributed in the hope that it will be useful,
 * but WITHOUT ANY WARRANTY; without even the implied warranty of
 * MERCHANTABILITY or FITNESS FOR A PARTICULAR PURPOSE. See the GNU
 * Lesser General Public License for more details.
 *
 * You should have received a copy of the GNU Lesser General Public
 * License along with this library; if not, write to the Free Software
 * Foundation, Inc., 51 Franklin St, Fifth Floor, Boston, MA 02110-1301
 * USA
 */

import org.languagetool.AnalyzedSentence;
import org.languagetool.AnalyzedToken;
import org.languagetool.AnalyzedTokenReadings;
import org.languagetool.JLanguageTool;
import org.languagetool.Language;
import org.languagetool.Languages;
import org.languagetool.rules.RuleMatch;

import java.io.IOException;
import java.util.List;

/**
 * A simple interactive test to see if the languages are available and don't
 * crash.
 */
public class Example {

```

```

public static void main(String[] args) throws IOException {
    String testtext = "Ich heie Anais Siebers. Wer bist du? Frage Google!";

    // Lade deutsche Sprache gemaess
    // ~/META-INF/org/language-tool/language-module.properties
    Language deutsch = getGermanLanguage();
    if (deutsch == null) {
        System.out.println("Deutsch als Sprache nicht verfgbar");
        System.exit(3);
    }

    JLanguageTool langTool = new JLanguageTool(deutsch);

    List<String> testsaetze = Semantik.satzTrennung(testtext);

    for (String testsatz : testsaetze) {
        List<RuleMatch> result = langTool.check(testsatz);
        System.out.println(
            "Checking '" + testsatz + "' with " + deutsch + ":");
        for (RuleMatch ruleMatch : result) {
            String[] fromToPos = ruleMatch.toString().split(":")[1]
                .split("-");
            System.out.println("    '"
                + testsatz.substring(Integer.parseInt(fromToPos[0]),
                    Integer.parseInt(fromToPos[1]))
                + "' ==> " + ruleMatch);
        }

        System.out.println(langTool.getAnalyzedSentence(testsatz)
            .getAnnotations().contains("."));

        List<AnalyzedSentence> analyse = langTool.analyzeText(testsatz);

        for (AnalyzedSentence sentence : analyse) {
            // for (String lemma : sentence.getLemmaSet()) {
            // System.out.println("LemmaSet:");
            // System.out.println(lemma);
            // }
            // Let's see what we could get ;- )

            System.out.println("SATZ = " + testsatz);
            String[] orgText = (" " + testsatz).split(" ");
            String[] analysedText = sentence.toString().split(" ");
            if (orgText.length != analysedText.length) {
                System.out.println("LENGTH MISMATCH! " + orgText.length
                    + " != " + analysedText.length);
                continue;
            }
            for (int w = 0; w < orgText.length; w++) {
                System.out.println("WORT ==> " + orgText[w] + " ==> "
                    + analysedText[w]);
            }
            // for (AnalyzedTokenReadings aTokens : sentence.getTokens()) {
            // System.out.print("\tTOKEN="+ aTokens.getToken());
            // for (AnalyzedToken aToken : aTokens.getReadings()) {
            // System.out.println("\t\t" + aToken.getLemma());
            // }
            // }
        } // **/
    }

}

public static Language getGermanLanguage() {
    // Lade verfuegbare Sprachen gemaess
    // ~/META-INF/org/language-tool/language-module.properties
    List<Language> sprachen = Languages.get();
    for (Language sprache : sprachen) {
        if (sprache.getLocale().toString().equals("de"))
            return sprache; // deutsch gefunden
    }
    return null; // kein deutsch verfgbar
}
}

```



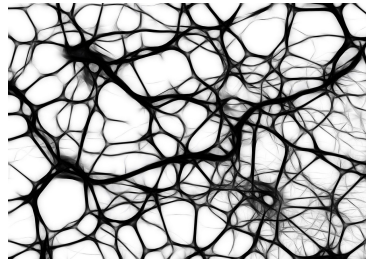
## VI. Ergebnis der modifizierten Example.class

```
Checking 'Ich heiÙe Anais Siebers.' with German (Germany):
'Anais' ==> GERMAN_SPELLER_RULE:10-15:Möglicher Rechtschreibfehler gefunden
'Siebers.' ==> GERMAN_SPELLER_RULE:16-24:Möglicher Rechtschreibfehler gefunden
true
SATZ = Ich heiÙe Anais Siebers.
WORT ==> ==> <S>
WORT ==> Ich ==> Ich[Ich/SUB:AKK:PLU:NEU,Ich/SUB:AKK:SIN:NEU,Ich/SUB:DAT:PLU:NEU,Ich/SUB:DAT:SIN:NEU,Ich/SUB:GEN:PLU:NEU,Ich/SUB:GEN:SIN:NEU,Ich/SUB:NOM:PLU:NEU,Ich/SUB:NOM:SIN:NEU,ich/PRO:PER:NC
WORT ==> heiÙe ==> heiÙe[heiÙen/VER:1:SIN:KJ1:NON,heiÙen/VER:1:SIN:PRÄ:NON,heiÙen/VER:3:SIN:KJ1:NON,heiÙen/VER:IMP:SIN:NON,heiÙ/ADJ:AKK:PLU:FEM:GRU:SOL,heiÙ/ADJ:AKK:PLU:MAS:GRU:SOL,heiÙ/ADJ:AKK:F
WORT ==> Anais ==> Anais[Anais/null,0]
WORT ==> Siebers. ==> Siebers[Siebers/null,0].[</S>./PKT,<P/>,0]
Checking 'Wer bist du?' with German (Germany):
false
SATZ = Wer bist du?
WORT ==> ==> <S>
WORT ==> Wer ==> Wer[wer/PRO:RIN:NOM:FEM,wer/PRO:RIN:NOM:MAS,0]
WORT ==> bist ==> bist[setn/VER:2:SIN:PRÄ:NON,sein/VER:AUX:2:SIN:PRÄ,0]
WORT ==> du? ==> du[ich/PRO:PER:NOM:SIN:ALG:2:B,0]?[</S>?/PKT,<P/>,0]
Checking 'Frage Google!' with German (Germany):
'Google' ==> DE_CASE:6-12:AuÙer am Satzanfang werden nur Nomen und Eigennamen großgeschrieben
false
SATZ = Frage Google!
WORT ==> ==> <S>
WORT ==> Frage ==> Frage[Frage/SUB:AKK:SIN:FEM,Frage/SUB:DAT:SIN:FEM,Frage/SUB:GEN:SIN:FEM,Frage/SUB:NOM:SIN:FEM,fragen/VER:1:SIN:KJ1:SFT,fragen/VER:1:SIN:PRÄ:SFT,fragen/VER:3:SIN:KJ1:SFT,fragen/
WORT ==> Google! ==> Google[googeln/VER:1:SIN:KJ1:SFT,googeln/VER:1:SIN:PRÄ:SFT,googeln/VER:3:SIN:KJ1:SFT,googeln/VER:IMP:SIN,0]![</S>!/PKT,<P/>,0]
```

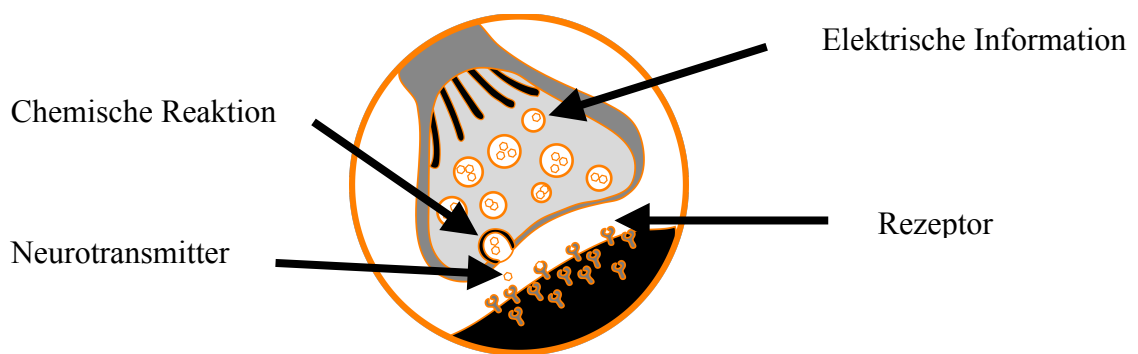
## VII. Wortklassen System - SGML-Abkürzungen

ABK Abkürzung	NOM Nominativ
ADJ Adjektiv	NON nicht-schwach
ADV Adverb	PA1 Partizip 1
AKK Akkusativ	PA2 Partizip 2
ART Artikel	PER personal
ATT attributiv	PLU Plural
AUX Hilfsverb	POS possessiv
BEG begleitend	PRÄ Präsens
B/S begleitend oder stellvertretend	PRD prädikativ
CAU kausal	PR1 proportional
COU Land	PRO Pronomen
DAT Dativ	PRP Präposition
DEF bestimmt	PRT Präteritum, Imperfekt
DEM Demonstrativpronomen	REF reflexiv
EIG Eigenname	REL relativ
EIZ erweiterter Infinitiv mit zu	RIN relativ oder interrogativ
FEM femininum	SFT schwach
GEB Gebiet	SIN Singular
GEN Genitiv	SOL alleinstehend
GEO geographischer Eigenname	STD Stadt
GRU Grundform	STV stellvertretend
IND unbestimmt	SUB Substantiv
INF Infinitiv	SUP Superlativ
INJ Interjektion	SZ Satzzeichen
IMP Imperativ	SZE Satzendezeichen
INR Interrogativpronomen	SZK Komma
KJ1 Konjunktiv 1	SZT Satztrennzeichen
KJ2 Konjunktiv 2	TMP temporal
KMP Kompositum	UNT unterordnend
KOM Komparativ	VER Verb
KON Konjunktion	VGL vergleichend
LOK lokal	VOR Vorname
MAS maskulinum	WAT Gewässer
MOD modal	ZAL Zahlwort
MOU Gebirge	ZAN Zahl beziehungsweise Ziffernfolge
NAC Nachname	ZUS Verbusatz
NEB nebenordnend	1 1. Person
NEG Negationspartikel	2 2. Person
NEU neutrum	3 3. Person
NIL Wortform nicht gefunden	A höflich
NOA ohne Artikel	B vertraut
NOG ohne Genus	

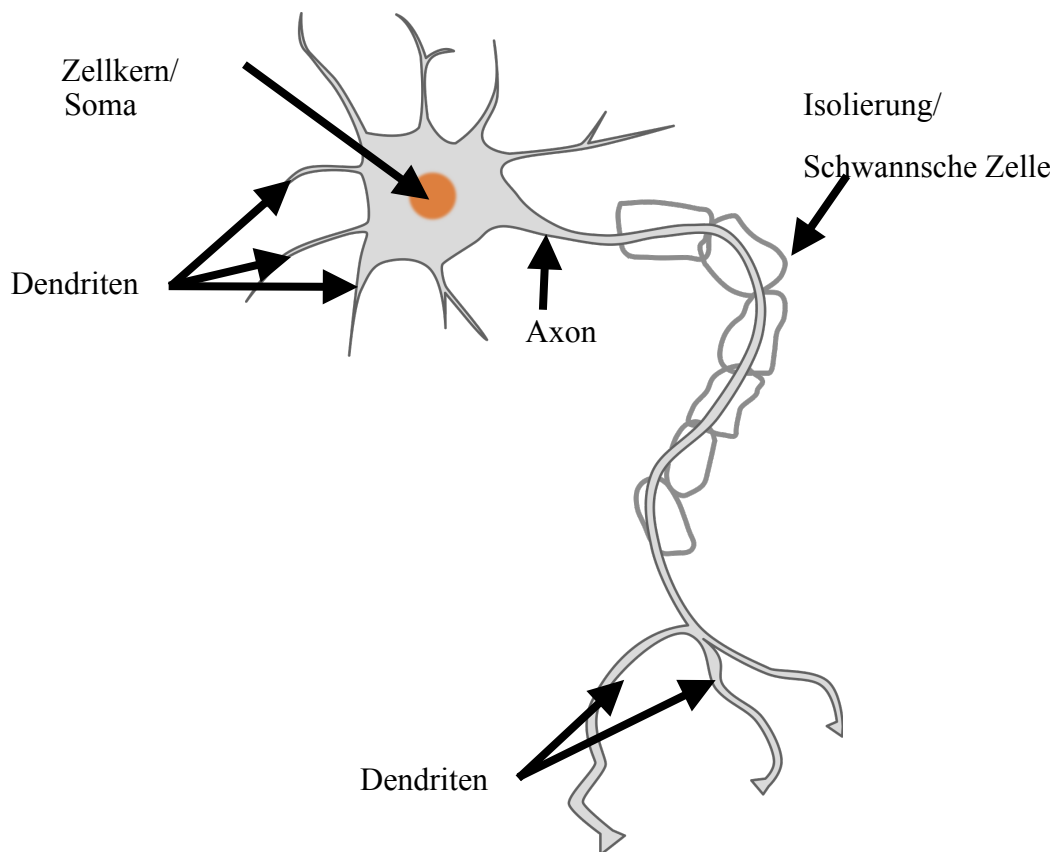
## VIII. Neuronales Netz (Gehirn)



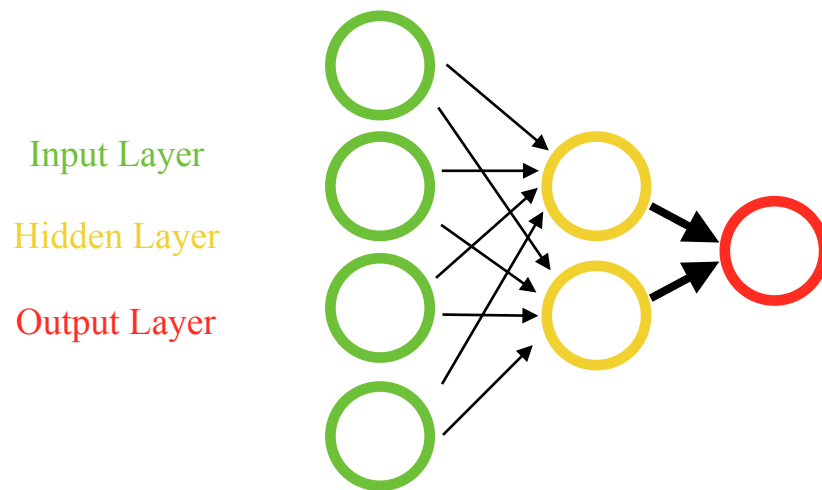
## IX. Chemische Synapse



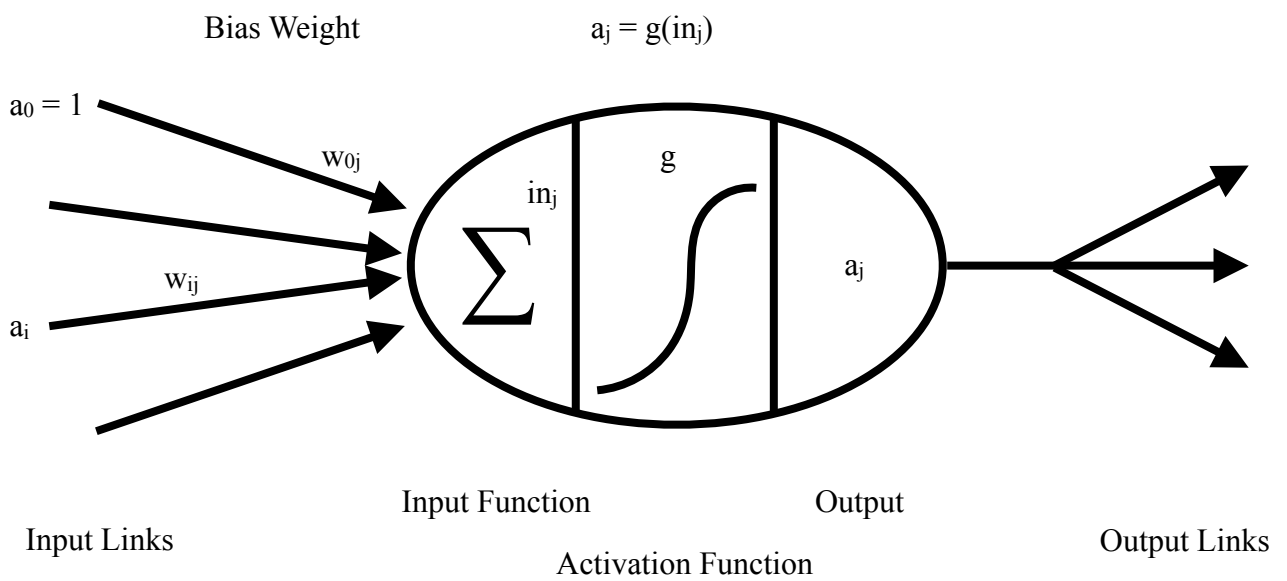
## X. Neuron



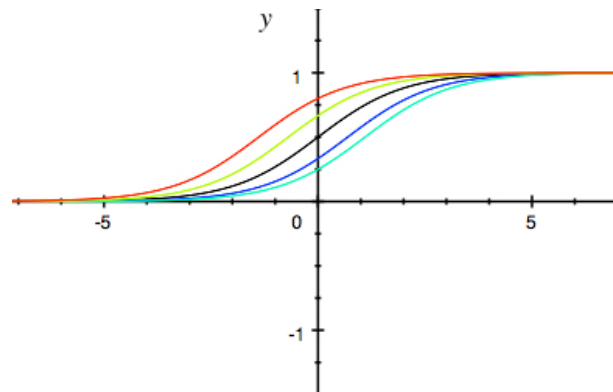
## XI. Layer des Künstlichen Neuronales Netzes



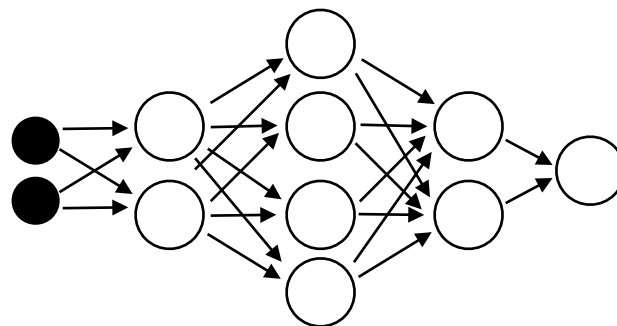
## XII. Aufbau eines technischen/künstlichen Neurons<sup>6</sup>(Zöllner 2014: 11, Abb. 3



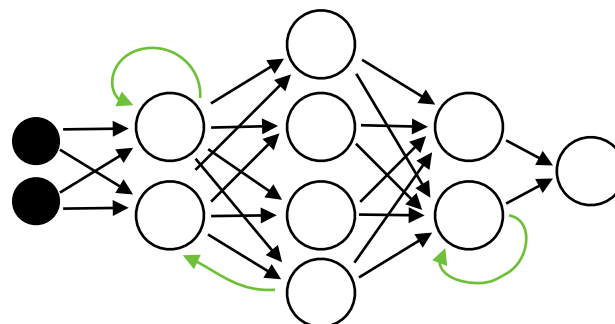
**XIII. Sigmoide Funktion Graph (schwarz = ursprüngliche Funktion, bunt = Flexibilität der Funktion)**



**XIV. Vergleich: Neuronales Netz - Rekurrentes Neuronales Netz**

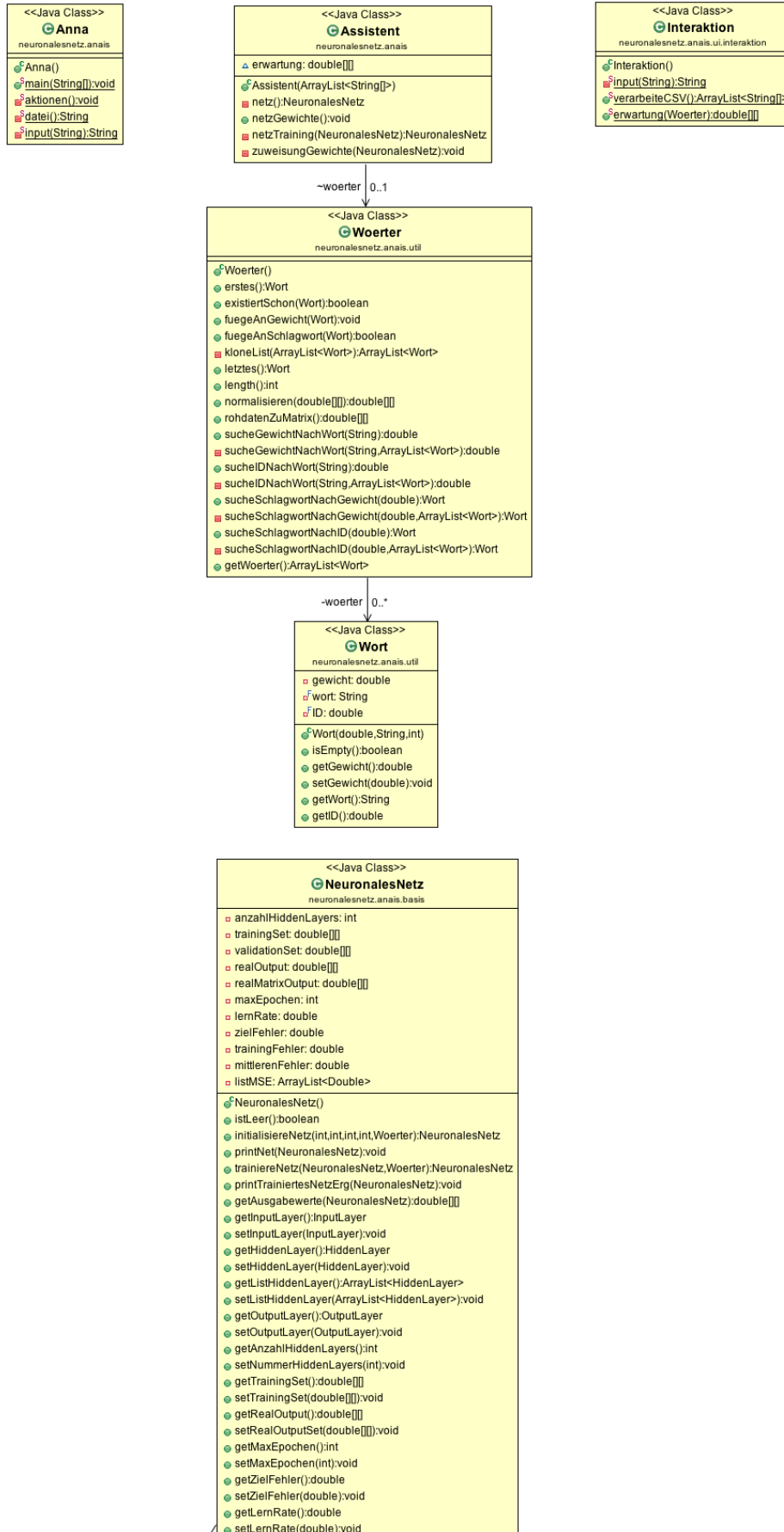


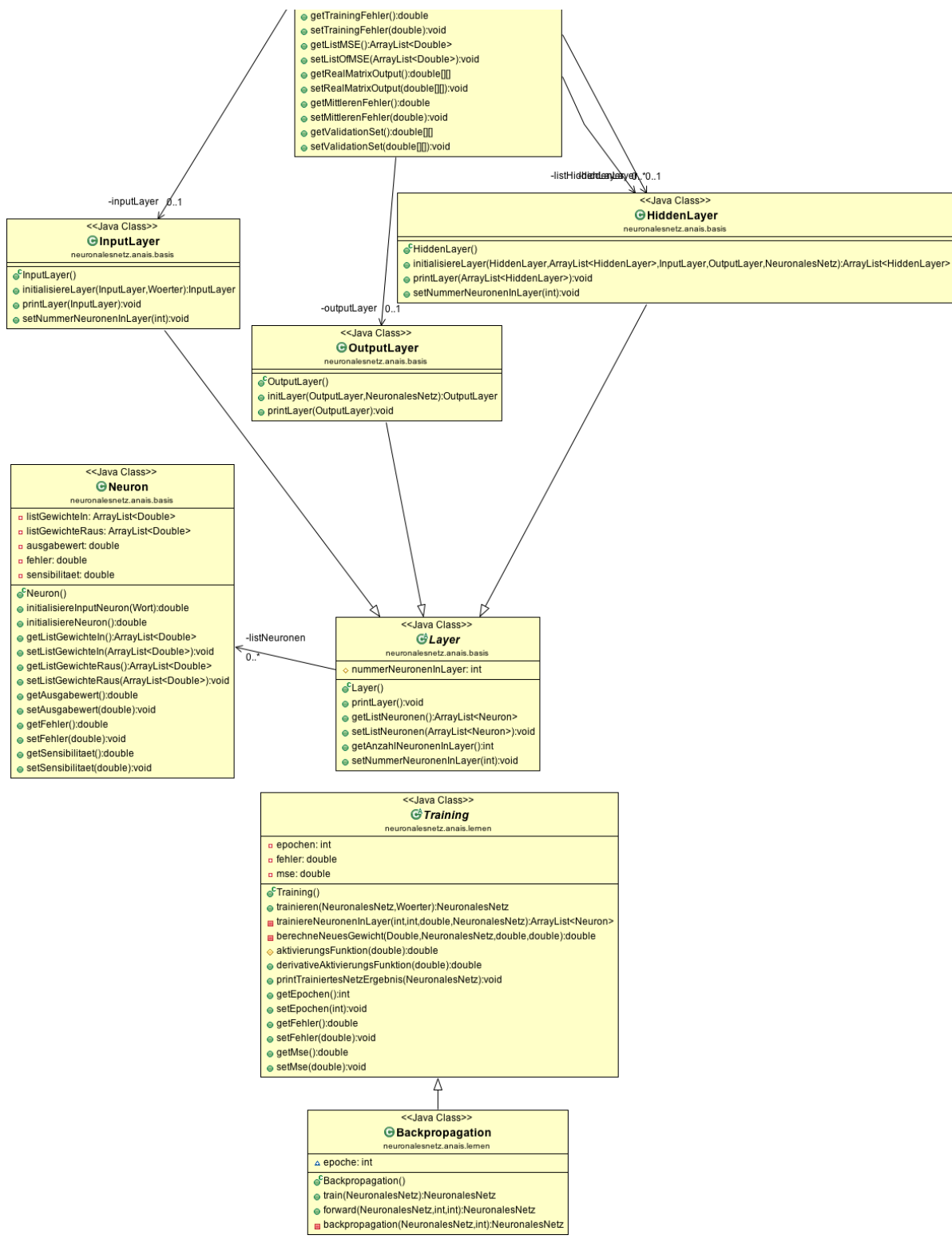
Neuronales Netz



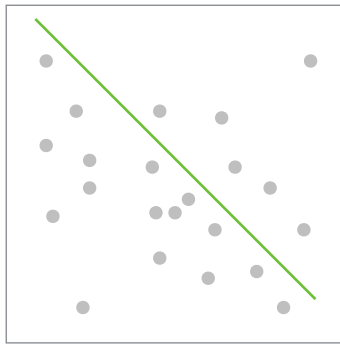
Rekurrentes Neuronales Netz

# XV. UML-Diagram

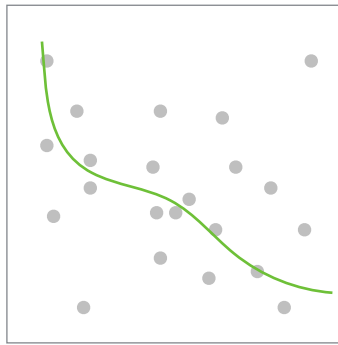




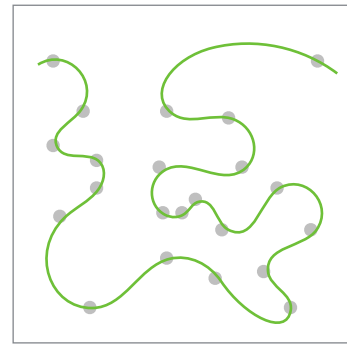
## XVI. Under-/Over-fitting data



Under-fitting



gutes fitting



Over-fitting

## XVII. Beispiel „Ich suche ein Haus in Bonn.“

ursprüngliche Tabelle aus der Datenbank:

Asche SeniorenHilfens24	<a href="http://www.pflegebonn.de/">http://www.pflegebonn.de/</a>	Bonnia Pflegen bietet Informationen zur genutzigen Seniorenbetreuung zu Hause durch Kräfte aus Polen und stellt ihre Philosophie und ihr Leistungsspektrum v	dmoz.Gesundheit.Pflege,Altenpflege,Produkte,Dienstleistungen
Kern, Dr. med. Stephan	<a href="http://kern-bonn.de/">http://kern-bonn.de/</a>	Die Facharzt- und Hausarztpraxis mit Schwerpunkt Diabetologie informiert über ihr Diagnose- und Behandlungsspektrum sowie über ihre Fortbildungen in Bonn	dmoz.Gesundheit,Praxen,Arzte,Internisten,Diabetologen
Offen! Tee Haus Michael Schwarzreiter	<a href="http://www.offen-tee.com/">http://www.offen-tee.com/</a>	Kindertees/Kabare Tees/Das Jugendzentrum in Bonn bietet Hausaufgabenbetreuung, Sport und verschiedene Kurse wie z.B. Samba oder Töpfern an	dmoz.Kids.and,Teens,Leute,Gesellschaft,Jugendtreffs,-zentren,Nordrhein-Westfalen
August Macke - Fotopostcard	<a href="http://www.augustmacke.com/">http://www.augustmacke.com/</a>	August Macke: Konzeption und Grafik, August Mackes Merchandise, Ausstellung im August Macke Haus in Bonn	dmoz.Kultur,Kunstgeschichte,Künstler,M,Macke,„August
Naturpark Rheinland	<a href="http://www.naturpark-rheinland.de/">http://www.naturpark-rheinland.de/</a>	Der Naturpark liegt direkt vor der Haustür der Städte Köln und Bonn. Auf über 1000 km² bietet das Naherholungsgebiet der Region eine einzigartige Landschaft	dmoz.Nordrhein-Westfalen,Regionen,Köln,Rhein,Bucht,Natur,Umwelt,Gesellschaft,Naturschutzgebiete
Katzenchutz Bonn/Weiss-Sieg e.V.	<a href="http://www.katzenchutz-bonn.de/">http://www.katzenchutz-bonn.de/</a>	Der Verein stellt Katzen vor, die ein neues Zuhause suchen	dmoz.Nordrhein-Westfalen,Städte,Gemeinden,B,Bonn,Gesellschaft,Tiere,Tierschutz,Tierarten,Säugetiere,Katzen,Hauskatzen
Kulturpark Bonn und Intermzzo	<a href="http://www.kulturpark-bonn.de/">http://www.kulturpark-bonn.de/</a>	Haus der Springmaus - Das Kabarett - Theater im Bonn Enderich, Kabarett, Improvisation, Comedy und Musik	dmoz.Nordrhein-Westfalen,Städte,Gemeinden,B,Bonn,Kultur
Haus Dmz	<a href="http://www.hausdmz.de/">http://www.hausdmz.de/</a>	Kunsthalle Bonn, im DMZ-Haus ist ein Kunstmuseum. Ein interaktiver Künstler Treffpunkt, Galerie, Atelier, Haus der Musik und Literatur	dmoz.Nordrhein-Westfalen,Städte,Gemeinden,B,Bonn,Kultur
Bender Hausverwaltung	<a href="http://www.hausverwaltung-bender.de/">http://www.hausverwaltung-bender.de/</a>	Bender Hausverwaltung und Immobilien in Bonn, Wohnungs- und Mietverwaltung im privaten sowie im geschäftlichen Bereich	dmoz.Nordrhein-Westfalen,Städte,Gemeinden,B,Bonn,Wirtschaft,Immobilien
Bonnia GmbH	<a href="http://www.bonnia.de/">http://www.bonnia.de/</a>	Ladengeschäft und Online-Shop für Designgeschichte	dmoz.Nordrhein-Westfalen,Städte,Gemeinden,B,Bonn,Wirtschaft,Möbel,Inneneinrichtung,Online-Shops,Haus,Garten,Fußböden,Teppiche,Teppichböden
Hausgeräte Dienst Lebrato	<a href="http://lebrato.de/">http://lebrato.de/</a>	Verkauf von Neu- und Gebrauchtgeräten von der Waschmaschine bis zum Toaster sowie Reparaturdienst oder Fabrikate	dmoz.Nordrhein-Westfalen,Städte,Gemeinden,B,Bonn,Wirtschaft,Unterhaltungs-,Haushaltselektronik
Bang und Olufsen Weltweit Odorous G	<a href="http://bosoerose.bang-olufsen.de/bom/">http://bosoerose.bang-olufsen.de/bom/</a>	Der Händler steht auf der Nebelstr. 79 und Hill-Garten vor	dmoz.Nordrhein-Westfalen,Städte,Gemeinden,B,Bonn,Wirtschaft,Unterhaltungs-,Haushaltselektronik
Kessels Espresso Studio, Inhaber Frank	<a href="http://www.kessels-espresso-studio.de/">http://www.kessels-espresso-studio.de/</a>	Bietet in seinen Geschäftsräumen in der Fischstraße in Bonn Espresso Maschinen verschiedener Hersteller an sowie ausgesuchte Café-Spezialitäten und dn	dmoz.Nordrhein-Westfalen,Städte,Gemeinden,B,Bonn,Wirtschaft,Unterhaltungs-,Haushaltselektronik
F & H Elektronik Bonn GmbH	<a href="http://www.f-h-elektronik-bonn.de/">http://www.f-h-elektronik-bonn.de/</a>	Bietet elektronische Bauelemente, Audio-Video-Zubehör, Tablet-Zubehör, Computer-Zubehör und Lötstation für Heim und Party an. Informationen über das	dmoz.Nordrhein-Westfalen,Städte,Gemeinden,B,Bonn,Wirtschaft,Unterhaltungs-,Haushaltselektronik
S+H Hausgeräte	<a href="http://www.sh-hausgeraete.de/">http://www.sh-hausgeraete.de/</a>	Kauf und Verkauf gebrauchter und neuer Hausgeräte mit Filialen in Bonn und Köln	dmoz.Nordrhein-Westfalen,Städte,Gemeinden,B,Bonn,Wirtschaft,Unterhaltungs-,Haushaltselektronik,K,Köln
Enzi Die Hausmeister GmbH	<a href="http://www.enzi-oh.de/">http://www.enzi-oh.de/</a>	Bietet neben einem Hausmeisterservice auch Dienstleistungen, Treppenhausreinigung und Wiederaufbau in Raum Köln/Bonn an	dmoz.Nordrhein-Westfalen,Städte,Gemeinden,K,Köln,Wirtschaft,Dienstleistungen,Hausmeisterservice
English Shop, Alexander McWhirney	<a href="http://www.english-shop.de/">http://www.english-shop.de/</a>	Lebermittel, alkoholfreie und alkoholhaltige Getränke, Geschenkartikel, Haushaltsgegenstände und Magazine aus England und Amerika. [D-50667 Köln, 53111	dmoz.Online-Shops,Spezialitäten
Sunderschnigang	<a href="http://www.buero-sunderschnigang.de/">http://www.buero-sunderschnigang.de/</a>	Er prüft die Rechnung sowie die Wirtschaftlichkeit und Ordnungsmäßigkeit der Haushalts- und Wirtschaftsprüfung. Über den Umfang nach Bonn, die Rechtsgrundlagen, mit	dmoz.Wirtschaft,Bauwesen,Bauausführung,Installationen,Wohnungsbau,Fertigbau,Holzbaueisen,Tafelbaueisen
Lehner Hobz&GmbH	<a href="http://www.lehner-holzhaus.de/">http://www.lehner-holzhaus.de/</a>	Das Unternehmen baut individuell konstruierte Wohnhäuser in Holztafelbauweise. Ausführungen auch als zertifiziertes Passivhaus möglich. Eine Bilderseite	dmoz.Wirtschaft,Bauwesen,ErneuerbareEnergie,Planung,Architektur,Wohnungsbau,Nordrhein-Westfalen,Städte,Gemeinden,Bonn
SB Medienleistungen GmbH	<a href="http://www.sbl.de/">http://www.sbl.de/</a>	Das Studienangebotsunternehmen hat sich auf Absatzleistungen, Hausarbeiten, Hausarbeiten, Hausarbeiten, Hausarbeiten, Hausarbeiten, Hausarbeiten, Hausarbeiten	dmoz.Wirtschaft,Frantgewerbe,Versicherungen,Sachversicherungen
UBN Lehrerbundesse Niedersachsen	<a href="http://www.ubn.de/">http://www.ubn.de/</a>	Man ist spezialisiert auf Hausarbeiten, Glas- und Unfallversicherungen und bezieht sich auf den Personenkreis der pädagogischen Berufe und den öffentlichen Dienst	dmoz.Wirtschaft,Frantgewerbe,Versicherungen,Sachversicherungen
Studierendenparlament der Rheinischen Film	<a href="http://www.sp.uni-bonn.de/">http://www.sp.uni-bonn.de/</a>	Das Bonner Studierendenparlament wählt den Bonner ASB und beschließt den Haushalt der Bonner Studierendenschaft. Es wird jedes Jahr neu gewählt.	dmoz.Wissen,Bildung,Hochschulen,Nordrhein-Westfalen,Rheinische,Friedrich-Wilhelms-Universität,Bonn

zum Testen genutzte csv-Datei mit Schlagworten:

**Schlagworte:** Test.csv unter ./src/neuronalesnetz/anais/ui/interaktion/Test.csv

dmoz,Gesellschaft,Menschen,Behinderte,Körperlich,Behinderte,Produkte,Dienstleistungen  
dmoz,Gesundheit,Frauen,Schwangerschaft,Geburt,Hebammen,Praxen,Nordrhein-Westfalen  
dmoz,Gesundheit,Pflege,Altenpflege,Produkte,Dienstleistungen  
dmoz,Gesundheit,Pflege,Altenpflege,Produkte,Dienstleistungen  
dmoz,Gesundheit,Praxen,Ärzte,Internisten,Diabetologen  
dmoz,Kids.and,Teens,Leute,Gesellschaft,Jugendtreffs,-zentren,Nordrhein-Westfalen  
dmoz,Kultur,Kunstgeschichte,Künstler,M,Macke,„August  
dmoz,Nordrhein-Westfalen,Regionen,Köln,Rhein,Bucht,Natur,Umwelt,Gesellschaft,Naturschutzgebiete  
dmoz,Nordrhein-  
Westfalen,Städte,Gemeinden,B,Bonn,Gesellschaft,Tiere,Tierschutz,Tierarten,Säugetiere,Katzen,Hauskatzen  
dmoz,Nordrhein-Westfalen,Städte,Gemeinden,B,Bonn,Kultur  
dmoz,Nordrhein-Westfalen,Städte,Gemeinden,B,Bonn,Kultur  
dmoz,Nordrhein-Westfalen,Städte,Gemeinden,B,Bonn,Wirtschaft,Immobilien  
dmoz,Nordrhein-Westfalen,Städte,Gemeinden,B,Bonn,Wirtschaft,Möbel,Inneneinrichtung,Online-  
Shops,Haus,Garten,Fußböden,Teppiche,Teppichböden  
dmoz,Nordrhein-Westfalen,Städte,Gemeinden,B,Bonn,Wirtschaft,Unterhaltungs-,Haushaltselektronik  
dmoz,Nordrhein-Westfalen,Städte,Gemeinden,B,Bonn,Wirtschaft,Unterhaltungs-,Haushaltselektronik  
dmoz,Nordrhein-Westfalen,Städte,Gemeinden,B,Bonn,Wirtschaft,Unterhaltungs-,Haushaltselektronik  
dmoz,Nordrhein-Westfalen,Städte,Gemeinden,B,Bonn,Wirtschaft,Unterhaltungs-,Haushaltselektronik  
dmoz,Nordrhein-Westfalen,Städte,Gemeinden,B,Bonn,Wirtschaft,Unterhaltungs-,Haushaltselektronik  
dmoz,Nordrhein-Westfalen,Städte,Gemeinden,K,Köln,Wirtschaft,Dienstleistungen,Hausmeisterservice

dmoz,Online-Shopse,Spezialitäten  
dmoz,Staat,Bundesregierung,-ministerien,-behörden  
dmoz,Wirtschaft,Bauwesen,Bauausführung,Installation,Wohnungsbau,Fertigbau,Holzbauweise,Tafelbauweise  
dmoz,Wirtschaft,Bauwesen,Entwurf,Planung,Architektur,Wohnungsbau,Nordrhein-Westfalen,Städte,Gemeinden,Bonn  
dmoz,Wirtschaft,Finanzgewerbe,Versicherungen,Sachversicherungen  
dmoz,Wissen,Bildung,Hochschulen,Nordrhein-Westfalen,Rheinische,Friedrich-Wilhelms-Universität,Bonn

## XVIII. Ausgabe der Konsole

Output : 0.9614035634641493

Erwartung: Bonn

ErwartungsID: 41.0

```
### INPUT LAYER ###
```

```
Neuron #1:
```

```
Input Weights:
```

```
[0.1]
```

```
[...]
```

```
Neuron #87:
```

```
Input Weights:
```

```
[0.9797978431773668]
```

```
### HIDDEN LAYER ###
```

```
Hidden Layer #1
```

```
Neuron #1
```

```
Input Weights:
```

```
[]
```

```
Output Weights:
```

```
[0.032698665460607734, 0.39611578149017346, 0.751671753146264, 0.37035863203127983, 0.7297878075717085,  
0.5611390703285437, 0.28126248395038345, 0.05893295647439689, 0.4248009490014856, 0.4155866749724335]
```

```
Neuron #1
```

```
Input Weights:
```

```
[0.4194561013086415, 0.6161477728925726, 0.30361150948061777, 0.07380355188177556, 0.3468919695623428,  
0.6265291249802273, 0.21525132433031824, 0.2110902730407226, 0.9507134259738167, 0.8736473069586513,  
0.1523501930812784, 0.7585242991078784, 0.6470122860795319, 0.9688526755936648, 0.30244160577300727,  
0.4405914729333641, 0.03492332564547462, 0.005411670967131688, 0.5612074173002394, 0.004960503287716134,  
0.9811849279519067, 0.9506147019515809, 0.33448601489459884, 0.6031722821251524, 0.3514004369662883,  
0.2349135221841573, 0.05019765502520712, 0.5178746317144849, 0.7592597184800471, 0.28998328439736565,  
0.024674120793001797, 0.7244689263846901, 0.27332471310160456, 0.13404894930113165, 0.5953336406903721,  
0.8030169730765166, 0.8845799090764186, 0.6017421754479969, 0.765097955647665, 0.28951488257139657,  
0.29902445566254865, 0.6793448758170602, 0.7656581057939327, 0.9473559360586534, 0.7134260698629494,  
0.7182027364572309, 0.6485379550032695, 0.934624041675998, 0.30502103998041574, 0.9952479408353867,  
0.1537312430850144, 0.48985333152757393, 0.455318210546357, 0.3721378569496666, 0.8401188812999266,  
0.3044623396425219, 0.5707797781511299, 0.8498587049525684, 0.37325094736389697, 0.7484799474047041,  
0.9400172970931278, 0.41064846547263767, 0.36012254717489556, 0.8992240525715368, 0.8069453923258951,  
0.24839963173955304, 0.8160807385421677, 0.9927940441788392, 0.8000448739879583, 0.7895814559655427,  
0.6316716061737087, 0.07861411092596327, 0.22363029783752797, 0.9626441768140553, 0.3706008944007868,  
0.2623655081748032, 0.7388392677545789, 0.35565626870367384, 0.17104194923254046, 0.23707719194994048,  
0.5528338678446975, 0.2103445958144835, 0.434666679545177, 0.6434579363730408, 0.6765379156598456,  
0.2410999397507667, 0.8327493202384395, 0.8133371677636139]
```

```
Output Weights:
```

```
[0.6701757168879097, 0.5880959200707789, 0.6879985132458867, 0.6338223368657057, 0.4747788080735885,  
0.1042553059070217, 0.6984729359942864, 0.7974307710887498, 0.38104664210830674, 0.8007917180155568]
```

```
[...]
```



Hidden Layer #81

Neuron #1

Input Weights:

[]

Output Weights:

[0.991447542745231]

[...]

Neuron #1

Input Weights:

[0.7048895894074656, 0.19023409667044167, 0.2759127438019605, 0.994251087270441, 0.0061816071185670785, 0.30351446160877316, 0.7443962891240556, 0.9171363187743229, 0.4208654671967874, 0.40564728991204624]

Output Weights:

[0.41672082592211845]

Mittlere quadratische Abweichung: 14434.189708841188

Number of epochs: 1000

0.0	0.01	0.02	0.03	0.04	0.05	0.06	0.07	0.08	0.09	0.1	0.11	0.12	
	0.13	0.14	0.15	0.16	0.17	0.18	0.19	0.2	0.21	0.22	0.23	0.24	
	0.25	0.26	0.27	0.28	0.29	0.3	0.31	0.32	0.33	0.34	0.35	0.36	
	0.37	0.38	0.39	0.4	0.41	0.42	0.43	0.44	0.45	0.46	0.47	0.48	
	0.49	0.5	0.51	0.52	0.53	0.54	0.55	0.56	0.57	0.58	0.59	0.6	
	0.61	0.62	0.63	0.64	0.65	0.66	0.67	0.68	0.69	0.7	0.71	0.72	
	0.73	0.74	0.75	0.76	0.77	0.78	0.79	0.8	0.81	0.82	0.83	0.84	
	0.85	0.86	NET OUTPUT: 0.9768197156130297					REAL OUTPUT: 41.0		ERROR:			

-40.02318028438697

Output: 0.9816595199928599

Erwartung: Bonn & 41.0 , Ergebnis: 0.9816595199928599

## XIX. Text nicht verstanden

ANNA suche Gymnasium in Bonn Q

**Uups**, das habe ich leider nicht verstanden ...

Folgendes habe ich verstanden:

```

suche [
  suchen/VER: 1: SIN: KJ1: SFT
  suchen/VER: 1: SIN: PRÄ: SFT
  suchen/VER: 3: SIN: KJ1: SFT
  suchen/VER: IMP: SIN: SFT
  suchen/VER: 1: SIN: KJ1: SFT
  suchen/VER: 1: SIN: PRÄ: SFT
  suchen/VER: 3: SIN: KJ1: SFT
  suchen/VER: IMP: SIN: SFT
  0
]

Gymnasium [
  Gymnasium/nuLL
  0
]

in [
  in/PRP: LOK: TMP: MOD: DAT: AKK
  0
]

Bonn [
  Bonn/EIG: AKK: SIN: NEU: ART: STD
  Bonn/EIG: AKK: SIN: NEU: NOA: STD
  Bonn/EIG: DAT: SIN: NEU: ART: STD
  Bonn/EIG: DAT: SIN: NEU: NOA: STD
  Bonn/EIG: GEN: SIN: NEU: ART: STD
  Bonn/EIG: NOM: SIN: NEU: ART: STD
  Bonn/EIG: NOM: SIN: NEU: NOA: STD
  0
]

```

Das Wort 'Gymnasium' kenne ich leider nicht.

Möglichkeiten:

Ich ...

- ... habe mich vertippt/versprochen.
- ... möchte dir den Begriff beibringen.
- ... möchte eine Datenquelle hinzufügen.
- ... brauche Hilfe.

Leider steht diese Funktion noch nicht zur Verfügung.

ANNA

suche Gymnasium in Bonn #

Ups, das habe ich leider nicht verstanden ...

Folgendes habe ich verstanden:

```

suche [
  suchen/VER: 1:SIN:KJ1:SFT
  suchen/VER: 1:SIN:PRÄ:SFT
  suchen/VER: 3:SIN:KJ1:SFT
  suchen/VER: IMP:SIN:SFT
  suchen/VER: 1:SIN:KJ1:SFT
  suchen/VER: 1:SIN:PRÄ:SFT
  suchen/VER: 3:SIN:KJ1:SFT
  suchen/VER: IMP:SIN:SFT
  0
]

Gymnasium [
  Gymnasium/SUB: AKK: SIN: NEU
  Gymnasium/SUB: DAT: SIN: NEU
  Gymnasium/SUB: NOM: SIN: NEU
  B-NP|NPS
]

in [
  in/PRP: LOK: TMP: MOD: DAT: AKK
  0
]

Bonn [
  Bonn/EIG: AKK: SIN: NEU: ART: STD
  Bonn/EIG: AKK: SIN: NEU: NOA: STD
  Bonn/EIG: DAT: SIN: NEU: ART: STD
  Bonn/EIG: DAT: SIN: NEU: NOA: STD
  Bonn/EIG: GEN: SIN: NEU: ART: STD
  Bonn/EIG: NOM: SIN: NEU: ART: STD
  Bonn/EIG: NOM: SIN: NEU: NOA: STD
  0
]

#[
  #/
  0
]

```

Das Wort '#' kenne ich leider nicht.

Möglichkeiten:

Ich ...

- ... habe mich vertippt/versprochen.
- ... möchte dir den Begriff beibringen.
- ... möchte eine Datenquelle hinzufügen.
- ... brauche Hilfe.

Leider steht diese Funktion noch nicht zur Verfügung.

## XX.keine Suchergebnisse

ANNA

Suche Zauberer in England

Suchergebnisse

*keine Suchergebnisse gefunden*

Filter (Klassen)  
*keine Daten vorhanden*

Filter (Symbol)  
*keine Daten vorhanden*

Filter (Schlagworte)  
*keine Daten vorhanden*

## XXI. Suchergebnisse gefunden

ANNA suche Gymnasium in Bonn

Suchergebnisse

[Beethoven-Gymnasium Bonn | dmoz](#)  
Die Internetpräsenz des Beethoven-Gymnasiums in Bonn  
<http://beethoven-gymnasium.de>

[Ruderverein am Beethoven-Gymnasiu...](#)  
Der Verein stellt sich mit Aktuellem, Terminen, Fotos, Forum und Archiv vor.  
<http://grv.missura.de>

[Kardinal-Frings-Gymnasium | dmoz](#)  
Das Gymnasium des Erzbistums Köln bietet umfangreiche Informationen über Fächerangebot, Termine und Projekte.  
<http://kardinal-frings-gymnasium.de>

[Hardtberg-Gymnasium | dmoz](#)  
Das Gymnasium mit deutsch-französischem Zweig wird vorgestellt.  
<http://joomla35.hardtberg-gymnasium.de>

[Friedrich-Ebert-Gymnasium | dmoz](#)  
Das Friedrich-Ebert-Gymnasium in Bonn informiert zur Schule, zu den verschiedenen Bildungsgängen und zum allgemeinem Schulleben.  
<http://feg-bonn.de>

[Nicolaus-Cusanus-Gymnasium \(NCG\) |...](#)  
Das bilinguale Gymnasium wird mit Schulprofil und Schulprogramm vorgestellt. Außerdem Informationen zu den Themen Schüleraustausch, Sport, Theater-AG und Schullalltag.  
<http://ncg-bonn.de>

[Gymnasial-Ruderverein am Beethoven-...](#)  
Informationen über Vorstand, Historie und die aktuellen Geschehnisse im Verein. Es gibt aktuelle Statistiken und ein Forum.  
<http://grv.missura.de>

Filter (Klassen)  
 www-html-home-content (11)

Filter (Symbol)  
 icon-internet (11)

Filter (Schlagworte)  
 Hürth (1)  
 Informatik (1)  
 Jugend (1)  
 Musik (1)  
 Rhein-Erft-Kreis (1)  
 Rudern (1)  
 Wassersport (1)  
 Abschlussjahrgang (2)  
 Sport (2)  
 Wissen (2)  
 MINT (3)  
 Schule (3)  
 Bonn (7)  
 Bildung (8)  
 Gemeinden (8)  
 Stadt (8)  
 Nordrhein-Westfalen (9)

## XXII. Filter

ANNA suche Gymnasium in Bonn

Suchergebnisse

[Beethoven-Gymnasium Bonn | dmoz](#)  
Die Internetpräsenz des Beethoven-Gymnasiums in Bonn  
<http://beethoven-gymnasium.de>

[Ruderverein am Beethoven-Gymnasiu...](#)  
Der Verein stellt sich mit Aktuellem, Terminen, Fotos, Forum und Archiv vor.  
<http://grv.missura.de>

[Kardinal-Frings-Gymnasium | dmoz](#)  
Das Gymnasium des Erzbistums Köln bietet umfangreiche Informationen über Fächerangebot, Termine und Projekte.  
<http://kardinal-frings-gymnasium.de>

[Hardtberg-Gymnasium | dmoz](#)  
Das Gymnasium mit deutsch-französischem Zweig wird vorgestellt.  
<http://joomla35.hardtberg-gymnasium.de>

[Friedrich-Ebert-Gymnasium | dmoz](#)  
Das Friedrich-Ebert-Gymnasium in Bonn informiert zur Schule, zu den verschiedenen Bildungsgängen und zum allgemeinem Schulleben.  
<http://feg-bonn.de>

[Nicolaus-Cusanus-Gymnasium \(NCG\) |...](#)  
Das bilinguale Gymnasium wird mit Schulprofil und Schulprogramm vorgestellt. Außerdem Informationen zu den Themen Schüleraustausch, Sport, Theater-AG und Schullalltag.  
<http://ncg-bonn.de>

[Gymnasial-Ruderverein am Beethoven-...](#)  
Informationen über Vorstand, Historie und die aktuellen Geschehnisse im Verein. Es gibt aktuelle Statistiken und ein Forum.  
<http://grv.missura.de>

Filter (Klassen)  
 www-html-home-content (11)

Filter (Symbol)  
 icon-internet (11)

Filter (Schlagworte)  
 Hürth (1)  
 Informatik (1)  
 Jugend (1)  
 Musik (1)  
 Rhein-Erft-Kreis (1)  
 Rudern (1)  
 Wassersport (1)  
 Abschlussjahrgang (2)  
 Sport (2)  
 Wissen (2)  
 MINT (3)  
 Schule (3)  
 Bonn (7)  
 Bildung (8)  
 Gemeinden (8)  
 Stadt (8)  
 Nordrhein-Westfalen (9)



## Suchergebnisse

[Beethoven-Gymnasium Bonn | dmoz](#)

Die Internetpräsenz des Beethoven-Gymnasiums in Bonn  
<http://beethoven-gymnasium.de>

[Friedrich-Ebert-Gymnasium | dmoz](#)

Das Friedrich-Ebert-Gymnasium in Bonn informiert zur Schule, zu den verschiedenen Bildungsgängen und zum allgemeinem Schulleben.  
<http://feg-bonn.de>

[Ernst-Moritz-Arndt-Gymnasium | dmoz](#)

Die Webseite des Ernst-Moritz-Arndt-Gymnasiums in Bonn mit umfangreichen Informationen zur Schule und Unterrichtsprojekten.  
<http://ema-bonn.de>

## Filter (Klassen)

 www-html-home-content (11)

## Filter (Symbol)

 icon-Internet (11)

## Filter (Schlagworte)

 Hürth (1) Informatik (1) Jugend (1) Musik (1) Rhein-Erft-Kreis (1) Rudern (1) Wassersport (1) Abschlussjahrgang (2) Sport (2) Wissen (2) MINT (3) Schule (3) Bonn (7) Bildung (8) Gemeinden (8) Stadt (8) Nordrhein-Westfalen (9)

## Suchergebnisse

[Ernst-Moritz-Arndt-Gymnasium | dmoz](#)

Die Webseite des Ernst-Moritz-Arndt-Gymnasiums in Bonn mit umfangreichen Informationen zur Schule und Unterrichtsprojekten.  
<http://ema-bonn.de>

## Filter (Klassen)

 www-html-home-content (11)

## Filter (Symbol)

 icon-Internet (11)

## Filter (Schlagworte)

 Hürth (1) Informatik (1) Jugend (1) Musik (1) Rhein-Erft-Kreis (1) Rudern (1) Wassersport (1) Abschlussjahrgang (2) Sport (2) Wissen (2) MINT (3) Schule (3) Bonn (7) Bildung (8) Gemeinden (8) Stadt (8) Nordrhein-Westfalen (9)

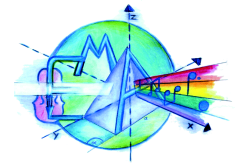
# 11. Implementierung

Aufgrund des Umfangs der Implementierung findet sich diese nur auf der CD-ROM.

# 12. Beratungsprotokolle zur Facharbeit

## Ernst-Moritz-Arndt-Gymnasium

Gymnasium der Stadt Bonn  
mit mathematisch-naturwissenschaftlichem  
und musikischem Schwerpunkt



## Beratungsprotokolle zur Facharbeit

Fach: Informatik

Schüler/in: Anais Siebers

Lehrer: Herr van den Boom

Datum	Gegenstand der Beratung	Arbeitsziel zur nächsten Beratung	Unterschrift des Lehrers	Unterschrift des Schülers
04.11.2016	Idee für die Facharbeit: - Entwicklung eines Programms zur Berechnung von .svgs und zweidimensionalen Animationen	- genauere Beschreibung des Vorhabens - Skizze der Umsetzung		
24.11.2016	Neue Idee für die Facharbeit: - Sprachgesteuerter KI-basierter adaptiver Assistent - Besprechung der Skizze des Assistenten	- Recherche: - zur Implementierung - inhaltlich - Inhaltsverzeichnis anlegen - Beginn mit der Implementierung		
21.12.2016	- Festlegung des Themas: Entwicklung eines persönlichen KI-basierten adaptiven Assistenten			
20.01.2017	- Besprechung des Inhaltsverzeichnis - möglicherweise Überschreitung der 12 Seiten	- Beginn mit der Facharbeit - Fortsetzen der Implementierung		

Das ausgefüllte Protokollblatt bitte am Ende der Facharbeit einfügen. (Im Anschluss an die Erklärung zur Eigenständigkeit.)

## 13. Selbstständigkeitserklärung

Hiermit erkläre ich, dass ich die vorliegende Arbeit selbstständig und ohne fremde Hilfe verfasst und keine anderen als die im Literaturverzeichnis angegebenen Hilfsmittel verwendet habe.

Auch die verwendeten Grafiken, Tabellen und Bilder sind, wenn sie nicht durch einen Verweis/eine Quelle markiert wurden, ausschließlich von mir selbst erstellt worden.

Insbesondere versichere ich, dass ich alle wörtlichen und sinngemäßen Übernahmen aus anderen Werken als solche gekennzeichnet habe.

---

Ort, Datum

---

Unterschrift