

Sprachmodelle im Natural Language Processing

„Betrachtung von Sprachmodellen des Natural Language Processing in
Bezug auf Besonderheiten der menschlichen Sprache“

Franziska Meyer

Sommersemester 2020

Seminararbeit im Modul: 10-201-2333 Wissen in der modernen Gesellschaft

Seminarleitung: Prof. Hans-Gert Gräbe, Ken Pierre Kleemann

Inhaltsverzeichnis

Einleitung	3
Natural Language Processing	4
Was ist NLP?	4
Nutzung von NLP	5
Ablauf NLP	6
Sprachmodelle im NLP	8
Was sind Sprachmodelle?	8
ELIZA	8
Regelbasiertes NLP	9
Statistische Sprachmodelle	10
Neuronale Netzwerke	12
Masked Language Modeling	13
Linguistik und NLP	15
Sprachmodelle und Kommunikation	15
Semantik und Bedeutung	16
Pragmatik und Kontext	18
Fazit	20
Literaturverzeichnis	21
Eidesstaatliche Erklärung	23

Einleitung

Natural Language Processing (im Folgenden auch mit NLP abgekürzt) ist bereits seit einigen Jahren in unserem Alltag so sehr verankert, dass es kaum noch wegzudenken ist. Mit NLP wurde die Interaktivität bestimmter Prozesse, die letztlich darauf abzielen, dem Menschen zu helfen und ihm automatisierbare Tätigkeiten abzunehmen, erhöht.

Zwar wird gleich noch genauer auf NLP eingegangen, vorausnehmen möchte ich jedoch einen Punkt, der wesentlich zur Entwicklung des Leitfadens dieser Seminararbeit dient. In manchen Quellen wird das Ziel von NLP nämlich als „direkte Kommunikation zwischen Mensch und Computer auf Basis der natürlichen Sprache“ (Luber/Litzel, 2016) beschrieben. Diese Aussage mag auf den ersten Blick vielleicht gar nicht so undenkbar wirken, schließlich handelt es sich um ein Ziel, nicht um den aktuellen Stand. Im Folgenden soll deshalb nach und nach zunächst ein Verständnis für NLP entwickelt werden, sowie für die zugrundeliegenden Sprachmodelle, durch welche NLP überhaupt seine Funktion erlangt. Anhand dieses Wissens sollen obiges genanntes Ziel, sowie weitere mögliche Schwierigkeiten, welche NLP bezüglich der menschlichen Sprache zu überwinden hat, kritisch betrachtet werden, um die entsprechenden Probleme hinsichtlich ihrer Überwindbarkeit zu bewerten.

Natural Language Processing

Zu Beginn soll sich nun mit Natural Language Processing selbst auseinandergesetzt werden. In diesem ersten Abschnitt geht es somit darum, den Begriff verstehen zu lernen, den groben Ablauf von NLP zu kennen und nachvollziehen zu können.

Was ist NLP?

Natural Language Processing (dt.: natürliche Sprachverarbeitung) ist ein Teilgebiet der Linguistik, der Informatik und der künstlichen Intelligenz. Besonders wird NLP eingesetzt, um Aufgaben zu bewältigen, die in der Regel menschliches Denken erfordern würden. Somit kann man zunächst vereinfacht sagen, dass NLP für die Aufgaben, für die der entsprechende Prozess ausgelegt ist, menschliches Denken simulieren können soll. Wie bereits erwähnt, wird sogar vom Ziel von „direkter Kommunikation zwischen Mensch und Computer auf Basis der natürlichen Sprache“ (Luber/Litzel, 2016) gesprochen. Den Begriff Kommunikation zu verwenden, wenn über ein Gespräch zwischen Mensch und Maschine gesprochen wird, erscheint mir hier jedoch recht hochgegriffen. Während wir uns mit diesem Begriff und seinem Verhältnis zu NLP jedoch erst unter dem Punkt Linguistik und NLP beschäftigen wollen, sei zunächst nach und nach behandelt, wozu NLP derzeit in der Lage ist und was aufgrund dessen möglich scheint.

NLP ist also darauf ausgelegt, den Computer in die Lage zu versetzen, „natürlich gesprochene Sprache so verarbeiten zu können, dass sie verstanden, manipuliert, interpretiert“ (sas, o.J.) und darauf reagiert werden kann. Diese natürlich gesprochene Sprache bezieht sich hierbei auf die Sprache, die wir als Menschen zur Kommunikation nutzen. Erreicht werden soll eine möglichst einwandfreie Kommunikation zwischen Mensch und Computer. In der Regel muss sich der Mensch jedoch an die Möglichkeiten des Computers anpassen. Ein System, welches sprachgesteuert Antworten auf Fragen gibt, muss schließlich noch lange nicht in der Lage dazu sein, „Smalltalk“ zu führen.

Die Komplexität dieser Aufgabe zeigt sich besonders, wenn man allein unsere sehr komplexe Sprache an sich betrachtet, welche verarbeitet werden muss. Wörter können mehrere Bedeutungen haben. Problematisch wird es auch bei Sarkasmus, Ironie oder rhetorischen Fragen, da hier die Bedeutung der Wörter nicht mehr zwingend etwas mit der Bedeutung der Aussage zu tun haben muss, sondern sich auch viel aus Kontext und Stimmlage lesen lässt. Bevor wir uns jedoch damit auseinandersetzen, wie gut die Möglichkeiten für NLP stehen, die entsprechenden linguistischen Besonderheiten zu erlernen, welche unsere Sprache so

einzigartig machen, betrachten wir zunächst, auf welchem Stand NLP sich aktuell befindet, um einen Eindruck über die Grenzen und Möglichkeiten von NLP zu erhalten.

Nutzung von NLP

Um sich NLP ein wenig besser vorstellen zu können, seien einige Anwendungen aufgeführt, welche NLP nutzen.

Dazu gehören beispielsweise E-Mail Filter, da sich gezeigt hat, dass „bestimmte Wörter, sowie gewisse Satzstrukturen eine Spam Mail signalisieren“ (tableau, o.J.) nach denen E-Mails von NLP durchforstet werden.

Ein anderes bekanntes Prinzip sind auch Smart Assistants wie Alexa, Siri, der Google Assistant oder Cortana, um einige sehr große Vertreter dieser Gattung zu nennen. Auch sie sind alle in der Lage, unsere gesprochenen Wörter und teilweise auch Kontext so zu analysieren, dass sie deren Bedeutung entschlüsseln und darauf reagieren können (vgl. tableau, o.J.).

Auch Predictive Texts sind bekannte NLP Verfahren. Wenn auf einem Smartphone ein Text geschrieben wird, so schlägt die genutzte Tastatur häufig einige Folgewörter vor. Auch das ist eine Art von NLP, wenn Wörter auf ihre aufeinanderfolgende Häufigkeit vorgeschlagen werden.

Übersetzungen, die einen Satz als Ganzes erfassen müssen, um eine möglichst gezielte, kontextuelle Übersetzung anzubieten, sind oft ebenfalls NLP gestützt. Eine direkte Übersetzung war vor NLP nicht immer möglich, da „unterschiedliche Satzstrukturen in verschiedenen Sprachen existieren, die zuvor unbeachtet bleiben mussten“ (tableau, o.J.), wo NLP nun gegen wirken kann.

Es zeigt sich also: NLP ist in vielen Bereichen der Mensch-Maschine Interaktion bereits vertreten und man könnte behaupten, dass so ziemlich jeder, der schon einmal ein Smartphone bedient oder das Internet benutzt hat unumgänglich mit NLP Erfahrungen gesammelt hat, nur dass es eben nicht immer wirklich auffällige Erfahrungen sind.

Ablauf NLP

Als nächstes wird der eigentliche Ablauf von NLP vorgestellt. Hierzu sei nun anzumerken: Es geht im Folgenden lediglich darum, den Ablauf grob vorzustellen und zu beschreiben. Die folgende Prozessbeschreibung ist keine exakte oder ultimative Art, um Natural Language Processing durchzuführen.

Zu Beginn, oder eher noch vor dem eigentlichen NLP, muss eine Eingabe erfolgen. Diese kann entweder schriftlich eingegeben oder mündlich vorgetragen werden. Bei letzterem muss die Eingabe dann zusätzlich noch in ein schriftliches Dokument umgewandelt werden (Speech-to-Text), wobei wir uns allerdings mit diesem Vorgang nicht weiter befassen. Auf diesem Eingabedokument wird dann unser NLP Prozess angewendet, welcher für eine Annotation des Dokumentes sorgt, sodass auf Grundlage dessen eine Ausgabe generiert werden kann (vgl. datenbanken-verstehen, o.J.).

Der Prozess an sich lässt sich nun grob in die folgenden Schritte aufteilen:

1. Tokenisierung: Der erste Schritt entspricht einer „Unterteilung in (logische) Segmente. Segmente können Abschnitte, Sätze oder einzelne Wörter sein“ (datenbanken-verstehen, o.J.). Gehen wir davon aus, dass jedes Wort ein Token ist, so würde ein Satz mit vier Wörtern entsprechend in vier Token zerlegt werden, denen in folgenden Schritten weitere Eigenschaften zugewiesen werden können.
2. Lexikalische/Morphologische Analyse: Als nächstes werden die Token auf ihre Struktur analysiert, sodass ihnen „eine Wortart zugewiesen werden kann und Wortstämme gebildet werden können“ (datenbanken-verstehen, o.J.). Zwei weitere Unteraufgaben dieses Schrittes sind „Stemming“ und „Lemmatization“. Ersteres befasst sich damit, Wörter in ihre Ursprungsform zurückzuführen. Die Wörter „springen“, „gesprungen“ und „sprang“ können allesamt auf den Infinitiv „springen“ zurückgeführt werden. Zweiteres hat die Aufgabe, Wörter näher auf ihre Bedeutung zurückzuführen. So ist „besser“ z.B. eine Steigerung, aber damit eben eine Version, von „gut“ (vgl. edureka!, 2018).
3. Syntaktische Analyse: Weiterhin wird die Satzstruktur ermittelt, welche den Token eine Funktion zuordnet, welche über Subjekt, Objekt, und Prädikat ausgedrückt werden. „Die Wörter werden so angeordnet, dass ihre Beziehung zueinander deutlich wird“ (tutorialspoint, o.J.).

4. Semantische Analyse: Den einzelnen Token wird eine genaue Bedeutung, z.B. die Bedeutung zugewiesen, die dieses Wort in einem Lexikon haben würde (vgl. tutorialspoint, o.J.), sowie ggf., mittels der Named-Entity-Recognition, ein Verfahren zur Erkennung von Eigennamen, einer Gruppe zugewiesen, die aussagt, ob das entsprechende Token z.B. eine Person, ein Ort oder eine Organisation ist (vgl. datenbanken-verstehen, o.J.).
5. Diskursanalyse: Zuletzt wird noch möglichst eine Verbindung zu vorigen oder nachfolgenden Sätzen gebildet, sodass Frage-Antwort Konstrukte erstellt werden können, oder aber vorige Sätze durch weitere Informationen an Bedeutung gewinnen (vgl. tutorialspoint, o.J.).
6. Pragmatische Analyse: Im besten Fall soll sogar eine pragmatische Analyse durchgeführt werden. Diese soll dafür Sorge tragen, dass „das, was gesagt wird, reinterpreted wird, so das sich zeigt, was eigentlich gemeint ist“ (tutorialspoint, o.J.). Man könnte sagen, dass hierdurch der Kontext besser hervorgehoben werden soll, wobei allerdings hierfür verstärkt Wissen aus der realen Welt vorliegen müsste, welches aus den Trainingskorpora jedoch nicht zwingend hervorgeht.

So lautet zumindest die Theorie. Jedoch sei hier anzumerken, dass der Vorgang mit jedem Schritt ein wenig komplexer wird, wobei zuletzt die pragmatische Analyse für eine Maschine geradezu unmöglich scheint. Erwartet wird hier, dass „Wissen aus der realen Welt“ vorliegt. Über den Umfang wird zwar keinerlei Aussage getroffen, doch damit existiert andersherum auch keine entsprechende Begrenzung. Dieses Wissen kann also theoretisch auch so umfangreich erwartet werden, wie es aus keinem Training je hervorgehen kann, da die Gesamtheit des menschlichen Wissens sich nicht nur schwierig als Trainingskorpus realisieren lässt, sondern höchstwahrscheinlich gar nicht, da besagtes Wissen ständig weiterwächst. Somit zeigt sich hier ein weiteres, im besten Falle schwer überwindbares, Problem.

Weiter soll nun betrachtet werden, wie so ein Training mit einem Sprachmodell überhaupt abläuft, um die Möglichkeiten und Komplexität von NLP noch konkreter betrachten zu können. Auch soll im kommenden hervorgehoben werden, welche Entwicklung in Richtung der betrachteten und noch zu betrachtenden Probleme NLP in den letzten Jahren durchgemacht hat.

Sprachmodelle im NLP

Wir wissen nun also, was NLP ausmacht und können beschreiben, wie die Annotation einer Eingabe von statten geht. Was wir allerdings noch nicht wissen, ist, was davor passieren muss, damit NLP die Annotation auch möglichst korrekt macht.

Was sind Sprachmodelle?

Zunächst muss hier noch eine Begrifflichkeit geklärt werden. Das, was in dieser Seminararbeit im Vorausgegangenen sowie im Folgenden als Sprachmodell (engl.: language model) bezeichnet wird, bezieht sich auf die theoretischen Konstrukte, welche erdacht und als Systeme umgesetzt wurden, um NLP zu trainieren. Dieser Begriff folgt sämtlichen, der in dieser Arbeit verwendeten Quellen, sollte jedoch nicht verwechselt werden mit Modellen, welche vom natürlichen Sprachgebrauch von Menschen handeln. Einige dieser Modelle werden weiter unter ebenfalls eingeführt, jedoch unter dem Namen Kommunikationsmodelle.

Ein Sprachmodell für NLP entspricht einem „Textkorpus (oder einer Sammlung von Textkorpora), mit dem trainiert wurde, eine Sprache zu analysieren, um dadurch bestimmte Regeln und Muster aus der Sprache abzuleiten“ (datenbanken-verstehen, o.J.). Wir brauchen diese Sprachmodelle, damit NLP Konstrukte unserer Sprache kennenlernt, so dass es später Eingaben auf Grundlage dessen, was es bereits kennt, analysieren kann. Wichtig ist dabei natürlich, dass die Korpora, mit denen NLP trainiert wird, mit den Inhalten, die dem Prozess letztlich als Input gegeben werden, möglichst weit übereinstimmen. Die Korpora müssen also inhaltlich sinnvoll gewählt werden.

Sprachmodelle haben sich dabei im Laufe der Zeit viel entwickelt, wodurch auch NLP immer konkreter und intelligenter wurde. Entsprechend sei diese Entwicklung hier vorgestellt, um einen Überblick auf die sich aus den Modellen entwickelten Systeme und den dadurch verbesserten Herangehensweisen an existierende Probleme zu geben.

ELIZA

Hierzu sei direkt gesagt, bei ELIZA handelt es sich nicht um ein Sprachmodell, welches sich NLP zunutze gemacht hat oder einen großartigen Korpus nutzt, allerdings ist genau dieses Programm eines der ersten Errungenschaften, wenn es um die Interaktion von Mensch und

Maschine geht und soll deshalb ebenfalls im Rahmen dieser Arbeit ein wenig Aufmerksamkeit bekommen.

ELIZA wurde 1966 entwickelt und war damit eine erste Simulation von menschlicher Interaktion. Eine natürliche Kommunikation mit einer Maschine war bereits damals das Ziel für ELIZA, auch wenn es natürlich bei weitem nicht so ausgereift war, als dass man besonders hier überhaupt von Kommunikation sprechen kann. Und am zugrunde liegenden Korpus zeigt sich hier besonders die Entwicklung: ELIZA arbeitete nämlich in einfacher Version lediglich mit einem Thesaurus, einem strukturierten Wörterbuch. Ein eingegebener Satz wird nach einem Wort durchsucht, welcher sich im Wörterbuch befindet. ELIZA stellte entsprechende Fragen zum selben oder einem übergeordneten Begriff. Auf eine Aussage wie: „Ich habe Probleme mit meinem Vater.“, könnte ELIZA reagieren mit „Erzähle mir mehr von deinem Vater.“ oder „Was kommt dir noch in den Sinn, wenn du an deinen Vater denkst.“. Falls ein entsprechender Begriff nicht gefunden wird, werden Ausweichphrasen gestellt, z.B. „Können Sie näher darauf eingehen?“, oder es sollte sich (in späteren Versionen) auf vorige Aussagen zurückbezogen werden können (vgl. Katzlberger, 2018). Obwohl ELIZA einen Psychotherapeuten simuliert und sich als solcher definitiv nicht durchgesetzt hat, war die Neuerung, mit einer Maschine ein Gespräch zu führen, dennoch ein wichtiger Meilenstein. Denn „ELIZA zielte darauf ab, die Benutzer zu täuschen, indem es sie glauben ließ, dass sie ein Gespräch mit einem echten Menschen führten.“ (Katzlberger, 2018), ein Ziel, welchem wir durch NLP schon sehr viel näher gekommen sind.

Regelbasiertes NLP

Ab den 1980ern etablierte sich dann das regelbasierte NLP, bei welchem die entsprechenden Systeme anhand von handgeschriebenen Regeln aussagekräftiger gemacht wurden, als ELIZA es war. Man ließ also NLP mit bestimmten Regeln trainieren. Dabei gab es solche Regeln, die angaben, wie korrekte Sätze konstruiert werden, oder sie beschrieben bestimmte reguläre Ausdrücke, die wir auch heute noch in der Programmierung nutzen. So konnte ganz einfach angegeben werden, dass die Sequenz „. „, also ein Punkt, gefolgt von einem Leerzeichen, einen Satz beendet. Vorstellen kann man sich das ganze also wie eine sogenannte kontextfreie Grammatik. Nach einem bestimmten Element können nur andere bestimmte Elemente gewählt werden, welche nach besagten Regeln auch gewählt werden dürfen. Beispielhaft kann dies so erklärt werden: Auf ein Wort, welches als Artikel, bzw. Bezeichner erkannt wurde, folgt oft ein Nomen, aber auch ein Adjektiv darf vorkommen, während Verben an dieser Position nicht auftreten können (vgl. Mayo, 2018).

Regelbasiertes NLP ist also eine Art Lücken füllen und tatsächlich in manchen Fällen sehr geeignet, aber nicht unbedingt generalisierbar. Für eine semantische oder gar pragmatische Analyse ist regelbasiertes NLP denkbar ungeeignet, schließlich gibt es viel zu viele Besonderheiten der menschlichen Sprache, welche von simplen Regeln nicht in Gänze erfasst werden können. Davon, dass hiermit Wissen aus der realen Welt erlernt werden könnte, ist somit ebenfalls denkbar unwahrscheinlich, wenn man bedenkt, dass für sämtliches Wissen, welches die Menschheit besitzt, eine Regel geschrieben werden müsste. Ganz davon ab lässt regelbasiertes NLP auch nicht darauf schließen, dass es in irgendeiner Weise für eine tatsächliche Kommunikation genutzt werden könnte. Auch wenn Syntax, also der korrekte Aufbau von Sätzen, durchaus zu einer gelungenen Konversation beisteuert, so ist der Inhalt, welcher hier gar nicht weiter beachtet wird, für ein tatsächliches Gespräch nicht weniger ausschlaggebend. Heute wird regelbasiertes NLP deshalb hauptsächlich noch dafür verwendet, Trainingsdaten vorzubereiten, also z.B. die Tokenisierung oder das Stemming vorzunehmen, wie wir es uns zuvor schon angesehen haben, auch wenn sie heute weniger handgeschrieben, als über If-Else Anfragen gestellt werden.

Statistische Sprachmodelle

Die nächste große Entwicklung von Sprachmodellen bezieht sich auf statistische Sprachmodelle, welche Wahrscheinlichkeiten nutzen, um nachfolgende Wörter vorherzusagen. Es entspricht somit quasi einer Markow-Kette, die dafür verwendet werden, Wahrscheinlichkeiten für das Eintreten zukünftiger Ereignisse anzugeben. Hier werden besonders oft die sogenannten n-Gramme genutzt: n-Gramme werden eingesetzt, um Wahrscheinlichkeiten für ein Wort nach n-1 vorigen Wörtern zu ermitteln. Angesichts der vorhin genannten Beispiele lassen sich n-Gramme sehr gut an Predictive Texts beschreiben. Auf Basis eines Wortes w wird der in den Korpora am häufigsten auftretende Nachfolger x des von w vorgeschlagen (vgl. Jurafsky, 2012/wikipedia, 2020).

$$P(w_1, \dots, w_m) = \prod_{i=1}^m P(w_i | w_1, \dots, w_{i-1}) \approx \prod_{i=1}^m P(w_i | w_{i-(n-1)}, \dots, w_{i-1})$$

Quelle: en.wikipedia.org/wiki/Language_model

Das n in n -Gramm bezieht sich dabei auf eine natürliche Zahl. Allerdings sind natürlich nicht unbedingt alle n 's wirklich zielführend dafür, NLP zu trainieren. Unigramme, bei denen $n = 1$ ist, sind dabei zum Beispiel nicht sonderlich nützlich für Predictive Texts, wie sich anhand des folgenden generierten Texts mittels eines Unigrams zeigt:

that, or, limited, the

Quelle: web.stanford.edu/class/cs124/lec/language modeling.pdf

Unigramme betrachten nämlich immer nur die Häufigkeit, mit der ein Wort selbst in einem Korpus vorkommt. Ein Vorschlag in Bezug auf den Vorgänger wird damit nicht gegeben. Sie sind eher dafür nützlich, die Häufigkeit eines Wortes innerhalb eines Sprachgebrauchs zu erfassen, unter Betrachtung der Predictive Texts würde aber kein sinnvoller Satz zu Stande kommen. Also betrachten wir hierfür lieber Bi- ($n = 2$) oder Trigramme ($n = 3$), welche ebenfalls die Wahrscheinlichkeiten eines Wortes hinsichtlich seines Vorkommens vorschlagen, allerdings anhand seiner Wahrscheinlichkeit nach seiner $n-1$ Vorgängerwörter. n -Gramme basieren damit auf einem statistischen Prinzip. Wenn man die Angaben der folgenden Grafik betrachtet, so zeigen sich direkt logische Wortfolgen, welche man sich in Kombination zu benutzen vorstellen kann:

outside, new, car, parking, lot, of, the, agreement, reached

Quelle: web.stanford.edu/class/cs124/lec/language modeling.pdf

Allerdings gibt es auch einige Einschränkungen bei n -Grammen: „Je größer n gewählt wird, umso genauer ist zwar die Vorhersage, erfordert jedoch auch entsprechende Rechenleistung“ (Rizvi, 2019), die an irgendeinem Punkt nicht mehr zwingend gegeben sein könnte. Außerdem muss natürlich betrachtet werden, „dass ein Wort, welches im Textkorpus, mit dem das Programm trainiert wurde, nicht vorkommt, stets eine Wahrscheinlichkeit von 0 hat“ (Rizvi, 2019), egal ob es existiert oder nicht, welches durchaus ein wichtiges Problem darstellt, besonders da ein Sprachmodell damit an dem Punkt „veraltet“ scheint, sobald Neologismen, also Wortneuschöpfungen, in der realen Welt konstruiert und etabliert sind. Im Grunde genommen können wir also auch mit n -Grammen nicht davon ausgehen, dass sämtliches Wissen, welches erwartet werden könnte, zur Verfügung steht, wenn wir noch nicht einmal davon ausgehen können, dass alle uns bekannten oder zu erwarteten Wörter zur Verfügung stehen. Weiter kann man hier auch noch keine großartige Möglichkeit betrachten, dass ein Computer eine tatsächliche Konversation führen könnte, wenn lediglich die wahrscheinlichsten Sätze, bzw. die wahrscheinlichsten Antworten auf vorausgegangene Fragen oder Appelle gewählt werden. So ist nach wie vor eine große Einschränkung gegeben. Einerseits kann man somit natürlich davon ausgehen, dass wahrscheinliche und damit richtige Antworten auf

Faktenfragen gefunden werden. Kommunikation basiert jedoch nicht auf einem festen Antwortpattern, sodass diese durch statistisches NLP nicht erreicht wird, wenn jedoch mittlerweile die Konstruktion, bzw. Neuordnung von Sätzen möglich scheint, deren Inhalte und Syntax aus den Trainingskorpora hervorgehen.

Neuronale Netzwerke

Seit einigen Jahren zeigt sich nun allerdings eine weitere sehr effektive Methode, um NLP zu trainieren. Deep Learning, welches sich an der Vorstellung des Lernprozesses des menschlichen Gehirns orientiert, und die dafür verwendeten neuronalen Netzwerke, werden seitdem für das Training genutzt. Das Besondere ist hier, dass die Maschine hierbei in der Lage ist, ohne menschliches Eingreifen weiter zu lernen.

Deep Learning liegt ein künstliches neuronales Netzwerk zugrunde, welches aus Knotenpunkten besteht, die künstliche Neuronen repräsentieren und aus Verbindungen mit Gewichtungen, die für die Entscheidungsprozesse von Nöten sind. Je höher eine Gewichtung, desto wahrscheinlicher wird ein Input im entsprechend folgenden Knoten dieser Verbindung erwartet. Je nachdem, wie präzise die letzte Ausgabe war und wie sehr andere Ergebnisse mit ihr korrelieren, kann eine Gewichtung angepasst werden, wodurch der eigentliche Lernprozess eintritt. Mehr Zwischenschichten an Neuronen bedeuten hierbei, dass umso komplexere Sachverhalte von einem neuronalen Netzwerk verstanden werden (vgl. Luber/Litzel, 2017).

Dadurch lassen sich die Neuronale Netzwerke besonders für zwei Verfahren nutzen:

- 1) Clustering: Es werden nicht annotierte Daten nach Ähnlichkeiten gruppiert. Es wird dadurch gelernt, ähnliche Daten zu finden, aber eben auch Auffälligkeiten in größtenteils ähnlichen Daten zu betrachten (vgl. pathmind, o.J.).
- 2) Klassifikation: Ein bereits annotierter Datensatz wird analysiert, so dass zwischen den gegebenen Labels der Annotation und den eigentlichen Daten bestimmte Verbindungen gefasst werden können (vgl. pathmind, o.J.).

Das Klassifikationsverfahren wird beispielsweise dafür eingesetzt, um Spam E-Mails vorzeitig zu erkennen. Gerade in Spam E-Mails zeigen sich besonders häufig bestimmte Inhalte und auch oft etwas unnatürliche oder falsche Grammatik, welche durch den fortlaufenden Lernprozess durch Neuronale Netzwerke immer besser erkannt werden, je mehr Spam E-Mails erhalten und auch nicht aus dem Spamordner entfernt werden. Funktioniert der Lernprozess mit neuronalen Netzwerken wie erwartet, ließe sich ebenfalls erwarten, dass

Konversationen mit NLP Systemen tatsächlich bedeutungsvoller würden. Je nachdem welche Satzstruktur und welche Wörter in einer Anfrage verwendet wurden, könnten diese mit genügend Knotenpunkten andere Antworten hervorbringen. Doch auch hierfür müsste eine gewisse Komplexität des Systems erreicht werden, wobei eine allumfassende Netzstruktur, die eine Vielzahl an Antworten auf eine rein theoretisch unendliche Menge an Anfragen, noch immer unvorstellbar ist.

Masked Language Modeling

Seit 2018 existiert noch ein weiteres vielversprechendes Verfahren für das Training von NLP. Dieses ist in der Lage, einen gesamten Input, also „ein gesamtes Dokument auf einmal zu erfassen, wodurch sich ein sogenanntes bidirektionales Verhalten ergibt“ (Horev, 2018). So können Wahrscheinlichkeiten vom Auftreten bestimmter Wörter nicht nur auf Vorgänger, sondern auch auf Nachfolger von Wörtern bezogen werden.

Dieses Prinzip wurde als Masked Language Modeling (MLM) bezeichnet, das dahinterstehende System nennt sich BERT. BERT wurde also darauf trainiert, bestimmte, „maskierte“ Wortlücken richtig zu befüllen, einfach aus dem Kontext heraus, in dem vorangegangene, als auch nachfolgende Wörter stehen. BERT ist somit dazu in der Lage, das folgende Beispiel mit sinnvollen Begriffen zu füllen (vgl. Devlin/Chang, 2018).

Input: The man went to the [MASK]₁ . He bought a [MASK]₂ of milk .
Labels: [MASK]₁ = store; [MASK]₂ = gallon

Quelle: ai.googleblog.com/2018/11/open-sourcing-bert-state-of-art-pre.html

Damit zeigt sich auch die Möglichkeit, dass BERT in der Lage dazu ist, Sätze durch ihren Kontext als aufeinander folgend zu klassifizieren, genannt Next Sentence Prediction (NSP).

Sentence A = The man went to the store.
Sentence B = He bought a gallon of milk.
Label = IsNextSentence

Sentence A = The man went to the store.
Sentence B = Penguins are flightless.
Label = NotNextSentence

Quelle: ai.googleblog.com/2018/11/open-sourcing-bert-state-of-art-pre.html

Durch diese Möglichkeiten zeigt BERT neue Möglichkeiten, menschliche Sprache besonders kontextuell besser verstehen zu können, wodurch auch NLP in Zukunft nur umso bessere Kommunikationsmöglichkeiten bieten kann. Betrachten wir also, was uns das System BERT nun bietet, hinsichtlich der vorigen Einschränkungen. Zunächst können wir nun davon ausgehen, dass durch das bidirektionale Verhalten eine besondere Analyse des Kontexts zustande kommt. Somit scheint MLM nun tatsächlich sämtliche sechs Schritte vom

theoretischen Ablauf von NLP abarbeiten zu können. Weiter sei nun aber die Frage, ob die pragmatische Analyse auch andere Probleme mit aus der Welt schafft. Ich denke trotz allem, dass wir trotz BERT noch immer nicht von einer Kommunikation an sich reden können. Zu diesem Aspekt sei aber unter Linguistik und NLP mehr gesagt. Selbst wenn eine Art von Unterhaltung zustande kommen würde, so können wir definitiv auch immer noch nicht von einem allumfassenden Wissen ausgehen, welches in besagtem System verankert ist. Letztlich ist alles, was der Computer weiß oder womit er lernt, seine Wissensbasis. Nach wie vor ist der Computer darauf angewiesen, mit den Informationen zu arbeiten, welche der Mensch ihm zu Beginn zur Analyse gegeben hat. Es gibt zwar Wissensbasen, welche sich aus bekannten Fakten neue Fakten ableiten können, jedoch befähigt dies eine Maschine nicht dazu, sich komplett neue oder innovative Konstrukte zu erdenken. Eine künstliche Intelligenz scheint zum aktuellen Stand somit hauptsächlich in der Lage, bereits bekanntes neu anzuordnen, um eine Erzeugung von Wissen oder Konstrukten zu simulieren.

Wir können soweit also davon ausgehen, dass die Sprachmodelle und -systeme durchaus dazu in der Lage sind, durch umfassende Analysen unsere menschliche Sprache so zu erfassen, dass sie diese „verstehen“. Auch die Nutzung oder Ausgabe von Sprache, bzw. das Finden einer passenden Antwort ist möglich. Jedoch möchte ich mich im nächsten Abschnitt noch mit einigen Konstrukten aus der Linguistik befassen und weshalb mir diese auch in naher Zukunft für NLP wie unüberwindbare Hürden erscheinen.

Linguistik und NLP

Sowohl Natural Language Processing, als auch die dafür benötigten Modelle zum Trainieren sind uns nun geläufig. Im Folgenden wird sich nun noch damit beschäftigt, auf welche Probleme wir stoßen, wenn wir versuchen, tatsächlich mit einem NLP System zu kommunizieren und was hierfür aus der Linguistik näher betrachtet werden kann, um diese komplexen Probleme auch als solche zu verstehen.

Sprachmodelle und Kommunikation

Eines der, meiner Meinung nach, wichtigsten Probleme, welches bislang immer wieder in dieser Arbeit aufgetaucht ist, lässt sich mit einem Begriff nennen. Dieser lautet Kommunikation. Kommunikation lässt sich beschreiben als „Verständigung und Austausch, die zwischen Menschen in Form von Sprache oder Zeichen abläuft“ (abiweb, o.J.). Dabei wird in der Regel eine Information übertragen. Diese Information kann im Grunde alles sein, was ein Mensch einem anderen ausdrücken möchte. Nach dem Kommunikationsmodell Organon-Modell nach Bühler hat Sprache drei Funktionen. Diese sind als Darstellung (Bezug auf einen Sachverhalt), Ausdruck (Sender drückt etwas über sich selbst aus) und Appell (Sender fordert Empfänger zu etwas auf) bezeichnet (vgl. studienkreis, o.J.). Der wichtigste Punkt, welcher nun aber betrachtet werden soll, ist, dass es in allen drei Funktionen um einen menschlichen Sender und einen menschlichen Empfänger geht. Beziehen wir das Organon-Modell nun auf NLP, so könnten man zunächst annehmen, dass hier stets ein Mensch einen Appell an ein System richtet. Etwa, um den Namen der Hauptstadt von Frankreich zu erfahren, oder um einen vorhandenen Satz in die englische Sprache zu übersetzen. Und das ist auch nicht komplett falsch, allerdings handelt es sich hier nun nicht mehr um einen Appell, wie er im Organon-Modell beschrieben ist.

Dass Mensch und Maschine nach dem Organon-Modell nicht miteinander kommunizieren, zeigt sich wie folgt: Eine Maschine (und damit auch NLP) ist letztlich nicht dazu in der Lage zu kommunizieren, wie es in der Linguistik verstanden wird, sondern lediglich Konstrukte unserer Sprache zu lernen. Das ist soweit auch sehr viel wert und für eine Vielzahl von Anwendungen auch definitiv ausreichend. Von einer Kommunikation, wie Menschen sie untereinander führen, kann meiner Meinung nach somit allerdings nicht geredet werden, eher von einer sehr umfassenden Interaktion. Manche Elemente einer Kommunikation lassen sich nicht durch eine Maschine realisieren. Letztlich kann eine Antwort, die wir einem Personal Assistent stellen, nicht wirklich auf einen Ausdruck analysiert werden. Über sich selbst drückt ein Assistent

letztlich nichts aus, während eine Antwort von einem Menschen Aufschluss darüber geben würde, ob er eine Uhr trägt, bzw. eine in Sicht ist, wenn er die Uhrzeit weiß, wenn wir nach dieser fragen, oder ob dieser einen bestimmten Fakt weiß, wenn er z.B. auf Anfrage die Hauptstadt von Frankreich nennen kann. Unter Betracht weiterer Kommunikationsmodelle, können sogar noch weniger Elemente einer Kommunikation der Interaktion mit einem Computer zugewiesen werden.

Betrachten wir hierzu noch das Kommunikationsquadrat, auch als Vier-Seiten-Modell bekannt. Nach Schulz von Thun, welcher dieses Modell erstellt hat, ist jede Nachricht nicht nur dreiseitig sondern vierseitig. Es wird auch von den vier Ohren des Empfängers der Nachricht geredet. Drei dieser Seiten sind uns schon aus dem Organon-Modell bekannt: die Sachebene, die Selbstaussage und ein Appell. Schulz von Thun fügte dem ganzen jedoch noch eine vierte Seite hinzu: der Beziehungsaspekt. Eine Nachricht hat somit stets das Potenzial, eine Aussage darüber zu machen, oder ein Verständnis dafür aufzubauen, wie Sender und Empfänger zueinander stehen (vgl. schulz-von-thun, o.J.). So kommuniziert ein Mensch in der Regel anders, je nachdem wer seine Nachrichten empfängt. Gegenüber den eigenen Eltern, dem Vorgesetzten oder dem Freundeskreis wird also stets ein anders Verhalten erwartet und akzeptiert werden. Ein Computer, ganz egal ob er nun NLP nutzt oder nicht, wird genau dies nicht tun. Egal, wie formal oder salopp eine Anfrage an z.B. einen Smart Assistent gerichtet wird, die Antwort wird trotz allem neutral ausfallen. Auch Beleidigungen, welche einen echten Menschen möglicherweise kränken würden, werden einen Smart Assistent nicht zum Schweigen bringen, oder seine Nachricht anders ausfallen lassen, sondern jede folgende Aussage auf eine Beleidigung würde noch immer mit gleicher Neutralität beantwortet werden.

Somit kann eine tatsächliche Kommunikation mit einer Maschine nach aktuellem Stand meiner Meinung nach noch lange nicht erreicht werden. Denn bislang zeigt sich lediglich eine (teils sehr ausgeprägte) Interaktion, welche anhand einiger Punkte jedoch als solche erkannt werden kann.

Semantik und Bedeutung

Im Folgenden sei der Bedeutung des Wortes „Bedeutung“ etwas Beachtung geschenkt. Was genau heißt es eigentlich, eine Bedeutung zu kennen? Für NLP scheint eine Bedeutung unabdingbar, um die einzelnen Token klassifizieren und erkennen zu können.

Betrachtet man diese Fragestellung nun aus linguistischer Perspektive, so betrachtet man ganz schnell die Frage der Semantik, die Lehre der Bedeutung. In der Semantik gibt es

wiederum eine Vielzahl an Begriffen, die uns Bedeutung in unterschiedlichen Kontexten erklärt. Um jedoch die Problematik, die sich hier zeigt, etwas einzugrenzen, möchte ich mich nur mit einem Konstrukt befassen. In der Semantik unterscheiden wir bei der Bedeutung eines Wortes zwischen der Denotation und der Konnotation. Denotation ist hierbei die Bedeutung eines Wortes, so wie sie in einem allgemeinen Lexikon gefunden werden kann. Anders ausgedrückt, die Denotation eines Wortes ist der tatsächliche Inhalt, den dieses Wort beschreibt. Die Konnotation hingegen beschreibt eher eine unterliegende Bedeutung, die oft an einen bestimmten Kontext gebunden ist und nicht in jedem kulturellen, ethischen, religiösen, etc. Umfeld verstanden werden muss (vgl. Kortmann, 2005). Um es an einem Beispiel zu verdeutlichen: Der Begriff „Schwein“ wird im Brockhaus unter folgender Definition aufgeführt: „Familie der Paarhufer [...] mit 16 Arten und einer Vielzahl von Unterarten in fünf Gattungen, die in den verschiedensten Biotopen Eurasiens und Afrikas verbreitet sind.“ (Brockhaus, o.J.). Dies entspricht somit der Denotation. In unserem kulturellen Kontext wird unter „Schwein“ jedoch auch ein abwertender Begriff verstanden, welcher der sprichwörtlichen Schmutzigkeit des Schweins entspringt (vgl. Dudenredaktion, o.J.). Obwohl auch dieses Verständnis beispielsweise im Duden als unterliegender Punkt verankert ist, ist diese Bedeutung dennoch keine Denotation. Die Möglichkeit besteht hier, dass die Aussage, einen Menschen als Schwein zu bezeichnen, nicht als diese verstanden wird, da in manchen kulturellen Kontexten eine solche Überlieferung nicht existieren könnte. Die Wahrscheinlichkeit, dass das Tier Schwein jedoch keine Bedeutung hat, ist sehr gering.

Ein anderes Beispiel sei das Wort „Schloss“. Nach dem Duden findet sich hier die Bedeutung „Vorrichtung zum Verschließen, Zuschließen mithilfe eines Schlüssels“ (Dudenredaktion, o.J.), sowie „Meist mehrflügeliges Wohngebäude des Adels“ (Dudenredaktion, o.J.) für entsprechendes Wort. Auch im Brockhaus werden beide Definitionen separat aufgeführt. Es handelt sich in diesem Fall um ein so genanntes Homonym, eigentlich zwei Wörter, die aus unterschiedlichsten Gründen gleich geschrieben und ausgesprochen werden. Im Falle von „Schloss“ gibt es also zwei eindeutige Denotationen. Von welcher Bedeutung gesprochen wird, schließt sich also nicht aus „Schloss“, oder „das Schloss“, da in diesem Fall sogar der Artikel gleich ist („Schild“ ist ebenfalls ein Homonym, aber: der Schild, das Schild). Nicht einmal „Das Schloss ist groß.“, muss viel aussagen, da es sich um ein gigantisches Gebäude oder aber um ein verhältnismäßig großes Vorhängeschloss handeln kann.

Was bedeutet das nun für NLP? Selbst für Menschen ist eine Eindeutigkeit eines Begriffes nicht gegeben. Sowohl Denotationen als auch Konnotationen sind uns als Menschen geläufig, Homonyme sorgen dabei sogar für eine noch größere Anzahl an Denotationen, die uns bekannt sind. Das Verständnis für die einzelnen Begriffe ist da, letztlich ist in beiden Beispielen der Kontext für das Verständnis, welche Denotation oder Konnotation gemeint ist, aber auch für einen Menschen unabdingbar. Doch im NLP Prozess zeigt sich hier ein Problem. Eine

pragmatische Analyse ist nicht zwingend Teil von NLP. Ein wirkliches Verständnis von Kontext scheint vor der Entwicklung von MLM nicht umfassend bis gar nicht vertreten. Die Wahrscheinlichkeit besteht also, dass Denotationen von Homonymen miteinander vertauscht werden könnten. Je nachdem, wie die Trainingsdaten strukturiert sind, bzw. um was für Trainingsdaten es sich handelt, besteht auch die Möglichkeit, dass Konnotationen für Denotationen und andersherum gehalten werden, wodurch vielleicht noch immer die einzelnen Bedeutungen der Wörter verstanden werden würden, der Kontext allerdings außer Acht gelassen würde, wodurch der Satz nicht mehr zwingend Sinn ergibt. Das Problem scheint sich also besonders im Bereich der Pragmatik fortzuführen.

Pragmatik und Kontext

Eine immer wiederkehrende Aussage bis hierhin bezog sich darauf, wie unabdingbar wichtig doch Kontext ist. Die erste wichtige Frage, die es dafür zu beantworten gilt, sei entsprechend, was denn eigentlich Kontext ist. Um diese Frage zu beantworten, betrachten wir einen weiteren linguistischen Begriff: Pragmatik. Die Pragmatik untersucht, wie in einer Kommunikation mehr als nur die semantische Bedeutung von Wörtern übertragen wird. Man könnte sagen, „Pragmatik befasst sich nicht mit dem, was gesagt wird, sondern mit dem, was gemeint ist“ (Bieswanger/Becker, 2008). Während sich Semantik mit der Bedeutung eines Wortes X befasst, so ist die Pragmatik damit beschäftigt herauszufinden, was es bedeutet, dass der Sprecher X benutzt und betrachtet diese Frage wiederum in Hinblick darauf, in welchem Kontext X verwendet wird (vgl. Kortmann, 2005).

Für NLP wäre nun wichtig, Pragmatik zu verstehen, um Kontext betrachten zu können, um sowohl die korrekte Bedeutung eines Wortes zu erschließen, als auch um vorangegangene Informationen in das bekannte Wissen mit einzubeziehen. Hierbei sollten sich einige Regeln einfacher umsetzbar zeigen, wie z.B., dass ein Subjekt, welches zuvor mit einem Nomen beschrieben wurde, in folgenden Sätzen mit einem Personalpronomen ersetzt werden kann. Seit der Entwicklung von MLM, scheint generell das Verständnis von Kontext nicht mehr allzu fern von der Umsetzbarkeit entfernt.

Dennoch kann, auch wenn Pragmatik damit recht umfassend verstanden wird, diese kaum auf einem menschlichen Niveau genutzt werden. Letztlich handelt Pragmatik, wie schon erwähnt, davon, was eigentlich gemeint ist. Doch auch wenn BERT beispielsweise den Kontext greifen kann und Sätze auf ihre aufeinanderfolgende Abfolge prüfen kann, so können auch „Insider“ oder logische Verknüpfungen, die den Kontext im Allgemeinen noch weiter übersteigen, von Menschen verstanden werden, von BERT höchstwahrscheinlich seltener. Betrachtet sei zum

Beispiel der Satz: „Kannst du mir sagen, wie spät es ist?“. Eine logisch nachvollziehbare Antwort wäre nun eine Uhrzeit. Ein Personal Assistent würde genau diese auch angeben. Eine entsprechende Antwort eines Menschen könnte etwas ungenau sein. „So ungefähr halb sechs.“, wäre aber dennoch eine logische Antwort auf diese Frage. Die Antwort „Naja, die Post war schon da.“, würde BERT wahrscheinlich eher als unpassend auf die zuvor formulierte Frage einstufen. Ein Mensch hingegen könnte in der Lage sein, die ungefähre Uhrzeit anhand dieser Aussage auszumachen, sofern dieser weiß, um welche Uhrzeit die Post im Normalfall bei ihm vorbei kommt. An diesem Beispiel lässt sich also sehen, dass Pragmatik sich nicht nur auf Kontext beziehen muss, sondern auch hier ein – teilweise sehr exklusives – Wissen aus der realen Welt verfügbar sein muss, um jegliche Aussagen zu verstehen, wozu die Möglichkeit nun dennoch noch nicht unabdingbar gegeben scheint, wie wir auch schon früher in dieser Seminararbeit betrachtet haben.

Fazit

Auch wenn sich ein gewisser NLP-kritischer Faden durch diese Arbeit gezogen hat, so sei letztlich nochmal zu bemerken, dass die Möglichkeiten, welche NLP bietet und auf welchem fortgeschrittenen Stand dieses Verfahren, nach nicht einmal einhundert Jahren Entwicklung, ist, wirklich überragend sind. Meiner Meinung nach wohnt NLP ein unglaubliches Potenzial inne, welches auch schon entsprechend genutzt wird und sicher auch in Zukunft immer präziser werden wird.

In mancher Hinsicht wird, wie im Rahmen dieser Arbeit hoffentlich hervorgegangen ist, jedoch etwas übertrieben, was NLP alles können soll, wenn es dabei auch nur um sehr hochgegriffene Formulierungen geht. Letztlich ist es ja auch nicht verkehrt, Ziele zu verfolgen, eine vollständige Erfüllung dieser kann ich persönlich jedoch in naher Zukunft nicht sehen. Trotz allem, basiert das Wissen, welches eine Maschine erhält, auf dem, was ihr vom Menschen gegeben wird. Das Wissen einer Maschine ist also auf gegebenes Wissen, oder dessen Abstraktion beschränkt. Ganz davon ab, dass das gesamte Wissen der Menschheit nur schwerlich einer Maschine in Gänze zugänglich gemacht werden kann, da dieses ja stetig weiter anwächst, wie eben schon in der Arbeit betrachtet.

Somit sei nun also noch festzuhalten, dass es trotz unglaublicher Entwicklungen und Möglichkeiten, welche NLP bietet, auch immer noch Grenzen gibt und zumindest in nächster Zeit auch geben wird, welche Computer eben nicht dazu befähigen, eine tatsächliche Kommunikation als solche mit einem Menschen zu führen. Es zeigt sich eindeutig, dass unsere menschliche Sprache als exklusiv angesehen werden kann und auch ein sehr gelungener Versuch, ihre Nutzung zu simulieren, dennoch eine Simulation bleibt und als solche erkennbar ist, womit eine Kommunikation zwischen Mensch und Maschine noch immer ein Ziel ist, welches in der Zukunft liegt. Noch gibt es einfach zu viele Hürden zu überbrücken, die sich durch unser aktuelles Wissen noch nicht umsetzen lassen.

Literaturverzeichnis

KORTMANN, B. (2005): *English Linguistics: Essentials*. 1. Aufl. Berlin: Cornelsen Verlag.

BIESWANGER M. & BECKER, A. (2008): *Introduction to English Linguistics*. 2. Aufl. Tübingen: Narr Francke Attempo Verlag GmbH + Co. KG

LUBER, S. & LITZEL, N. (2016, September 01): *Was ist Natural Language Processing?* Abgerufen 27.08.2020, von <https://www.bigdata-insider.de/was-ist-natural-language-processing-a-590102/>

SHETTY, B. (2018, November 24): *Natural Language Processing for Machine Learning*. Abgerufen 30.08.2020 von <https://towardsdatascience.com/natural-language-processing-nlp-for-machine-learning-d44498845d5b>

MAYO, M. (2018, Oktober): *The Main Approaches to Natural Language Processing Tasks*. Abgerufen 29.08.2020 von <https://www.kdnuggets.com/2018/10/main-approaches-natural-language-processing-tasks.html>

RIZVI, M. (2019, August 08): *A Comprehensive Guide to Build your own Language Model in Python!*. Abgerufen 30.08.2020 von <https://www.analyticsvidhya.com/blog/2019/08/comprehensive-guide-language-model-nlp-python-code/>

JURAFSKY, D. (2012): *Language Modeling: Introduction to N-grams*. Abgerufen 29.08.2020 von <https://web.stanford.edu/class/cs124/lec/languagemodeling.pdf>

DELVIN, J. & CHANG, M. (2018, November 02): *Open Sourcing BERT: State-of-the-Art Pre-training for Natural Language Processing*. Abgerufen 29.08.2020 von <https://ai.googleblog.com/2018/11/open-sourcing-bert-state-of-art-pre.html>

HOREV, R. (2018, November 10): *BERT Explained: State of the art language model for NLP*. Abgerufen 29.08.2020 von <https://towardsdatascience.com/bert-explained-state-of-the-art-language-model-for-nlp-f8b21a9b6270>

KATZLBERGER, M. (2018, August 31): *ELIZA, der erste Chatbot der Welt. Jetzt ausprobieren!* Abgerufen 30.08.2020 von <https://katzlberger.ai/2018/08/31/eliza-der-erste-chatbot/>

NATURAL LANGUAGE PROCESSING. (2020). In *Wikipedia, die freie Enzyklopädie*. Abgerufen von https://en.wikipedia.org/wiki/Natural_language_processing

LANGUAGE MODEL. (2020). In *Wikipedia, die freie Enzyklopädie*. Abgerufen von https://en.wikipedia.org/wiki/Language_model

EDUREKA!. (2018, Oktober 16). Natural Language Processing In 10 Minutes | NLP Tutorial For Beginners | NLP Training | Edureka. Abgerufen 29.09.2020, von

https://www.youtube.com/watch?v=5ctbvKAMQO4&t=378s&ab&ab_channel=edureka%21

SAS. (o.J.): *Natural Language Processing. Was es ist und was man darüber wissen sollte.*

Abgerufen 29.09.2020, von https://www.sas.com/de_de/insights/analytics/what-is-natural-language-processing-nlp.html

TABLEAU. (o.J.): *8 common examples of natural language processing and their impact on communication.* Abgerufen 29.08.2020 von <https://www.tableau.com/learn/articles/natural-language-processing-examples>

DATENBANKEN-VERSTEHEN. (o.J.): *Natural Language Processing.* Abgerufen 29.08.2020 von <https://www.datenbanken-verstehen.de/lexikon/natural-language-processing/>

TUTORIALSPPOINT. (o.J.): *AI – Natural Language Processing.* Abgerufen 29.08.2020 von https://www.tutorialspoint.com/artificial_intelligence/artificial_intelligence_natural_language_processing.htm

PATHMIND. (o.J.): *A Beginner's Guide to Neural Networks and Deep Learning.* Abgerufen 29.08.2020 von <https://wiki.pathmind.com/neural-network>

DUDENREDAKTION. (o.J.): „Schwein“ auf Duden online. Abgerufen 29.08.2020 von <https://www.duden.de/rechtschreibung/Schwein>

DUDENREDAKTION. (o.J.): „Schloss“ auf Duden online. Abgerufen 29.08.2020 von <https://www.duden.de/rechtschreibung/Schloss>

BROCKHAUS. (o.J.): „Schwein“ auf Brockhaus. Abgerufen am 16.09.2020 von <https://brockhaus.de/ecs/enzy/article/schweine>

STUDIENKREIS. (o.J.): *Das Organonmodell von Karl Bühler einfach erklärt.* Abgerufen 30.08.2020 von <https://www.studienkreis.de/deutsch/organonmodell-karl-buehler/>

ABIWEB. (o.J.): Kommunikation – Was ist das eigentlich? Abgerufen 30.08.2020 von <https://www.abiweb.de/deutsch-methoden/thematisches-vorwissen/kommunikation/kommunikation-definition.html>

SCHULZ-VON-THUN. (o.J.): das Kommunikationsquadrat. Abgerufen 30.08.2020 von <https://www.schulz-von-thun.de/die-modelle/das-kommunikationsquadrat>