

Özkan Abdullah Akyüz

Masterarbeit im Fach

Information Systems

Vorhersage der politischen Ausrichtung von Nachrichtenartikeln mittels Machine Learning

Themasteller: Prof. Dr. Detlef Schoder

Vorgelegt in der Masterprüfung
im Studiengang Information Systems
der Wirtschafts- und Sozialwissenschaftlichen Fakultät
der Universität zu Köln

Köln, November 2020

Inhaltsverzeichnis

Inhaltsverzeichnis	I
Abbildungsverzeichnis.....	II
Tabellenverzeichnis	IV
Abkürzungsverzeichnis.....	V
1 Motivation.....	1
2 Theoretischer Hintergrund.....	2
3 Methodik.....	5
3.1 Literaturrecherche.....	6
3.1.1 Verfahren basierend auf Machine Learning.....	9
3.1.2 Graphentheorie-basierte Verfahren.....	16
3.1.3 Statistische Verfahren	19
3.1.4 Sentimentanalysen	20
3.2 Machbarkeitsstudie.....	21
3.2.1 Mediennavigatoren.....	22
3.2.2 Datensätze	25
3.2.3 Analyse des Datensatzes und der Mediennavigatoren.....	32
3.3 Implementierung.....	44
3.3.1 Target Labeling	45
3.3.2 Preprocessing & Vectorization	46
3.3.3 Modellanwendung & -bewertung	47
4 Ergebnisse.....	51
5 Diskussion.....	59
6 Fazit	62
Literaturverzeichnis	64
Eidesstaatliche Erklärung	74
Einverständniserklärung zur Plagiatsprüfung.....	75
Lebenslauf.....	76

Abbildungsverzeichnis

Abbildung 1: Methodisches Vorgehen	5
Abbildung 2: Vorgehen während der Implementierung	6
Abbildung 3: Mediennavigator von SPR.....	22
Abbildung 4: Mediennavigator von AllSides	24
Abbildung 5: Mediennavigator von Ad Fontes Media	25
Abbildung 6: Verteilung der Klassen im NELA-GT-2018 Datensatz anhand des Mediennavigators von SPR ohne Berücksichtigung der Nachrichten, die keiner politischen Klasse zugewiesen sind	35
Abbildung 7: Verteilung der Klassen im NELA-GT-2018 Datensatz anhand des Mediennavigators von AllSides ohne Berücksichtigung der Nachrichten, die keiner politischen Klasse zugewiesen sind	35
Abbildung 8: Verteilung der Klassen im NELA-GT-2018 Datensatz anhand des Mediennavigators von Ad Fontes Media ohne Berücksichtigung der Nachrichten, die keiner politischen Klasse zugewiesen sind	36
Abbildung 9: Verteilung der Klassen im Kaggle Datensatz anhand des Mediennavigators von SPR ohne Berücksichtigung der Nachrichten, die keiner politischen Klasse zugewiesen sind	37
Abbildung 10: Verteilung der Klassen im Kaggle Datensatz anhand des Mediennavigators von AllSides ohne Berücksichtigung der Nachrichten, die keiner politischen Klasse zugewiesen sind	38
Abbildung 11: Verteilung der Klassen im Kaggle Datensatz anhand des Mediennavigators von Ad Fontes Media ohne Berücksichtigung der Nachrichten, die keiner politischen Klasse zugewiesen sind	38
Abbildung 12: Anzahl der Nachrichten je Datensatz, die keiner politischen Klassen zugeordnet wurde.....	39
Abbildung 13: Anteil der in den Mediennavigatoren abgedeckten Nachrichtenquellen durch die Datensätze	40
Abbildung 14: Verteilung der Klassen im zusammengeführten Datensatz	44
Abbildung 15: Accuracy Werte der Modelle, die im <i>Full Experiment</i> angewandt worden sind.....	54
Abbildung 16: Accuracy Werte der Modelle, die im <i>Full Experiment</i> und in den <i>Ablation Studies</i> angewandt worden sind	55

Abbildung 17: Dauer (in Sekunden), die ein Modell im *Full Experiment* zum Trainieren für durchschnittlich ein *Epoch* benötigt..... 56

Abbildung 18: Verhältnis der Dauer (in Sekunden) zum Trainieren eines *Epochs* der angewandten Modelle im *Full Experiment* im Vergleich zu ihrer erzielten Accuracy 57

Tabellenverzeichnis

Tabelle 1: Übersicht der Literaturrecherche	7
Tabelle 2: Anzahl der Artikel je Quelle im NELA-GT-2018 Datensatz	26
Tabelle 3: Anzahl der Nachrichtenanbieter im Kaggle Datensatz.....	31
Tabelle 4: Anzahl der Nachrichtenanbieter je Mediennavigator	32
Tabelle 5: Anzahl der Nachrichtenanbieter im NELA-GT-2018 Datensatz je Mediennavigator	33
Tabelle 6: Anzahl der Nachrichtenanbieter im Kaggle Datensatz je Mediennavigator .	33
Tabelle 7: Verteilung der Klassen im NELA-GT-2018 Datensatz anhand des Mediennavigators von SPR	34
Tabelle 8: Verteilung der Klassen im NELA-GT-2018 Datensatz anhand des Mediennavigators von AllSides.....	34
Tabelle 9: Verteilung der Klassen im NELA-GT-2018 Datensatz anhand des Mediennavigators von Ad Fontes Media.....	34
Tabelle 10: Verteilung der Klassen im Kaggle Datensatz anhand des Mediennavigators von SPR	36
Tabelle 11: Verteilung der Klassen im Kaggle Datensatz anhand des Mediennavigators von AllSides.....	37
Tabelle 12: Verteilung der Klassen im Kaggle Datensatz anhand des Mediennavigators von Ad Fontes Media.....	37
Tabelle 13: Anzahl der Nachrichten je Nachrichtenanbieter und deren politische Richtung im zusammengeführten Datensatz	41
Tabelle 14: Verteilung der Klassen im zusammengeführten Datensatz	43
Tabelle 15: Übersicht der angewandten Modelle	49
Tabelle 16: Accuracy Werte der Modelle im <i>Full Experiment</i> und in den <i>Ablation Studies</i>	53
Tabelle 17: Konfusionsmatrix des <i>RoBERTa</i> Modells im <i>Full Experiment</i>	58
Tabelle 18: Konfusionsmatrix des <i>BERT</i> Modells im <i>Full Experiment</i>	58

Abkürzungsverzeichnis

AUC	Area under the curve
BERT	Bidirectional Encoding Representations from Transformer
CNN	Convolutional Neural Network
GRU	Gated Recurrent Unit
IBC	Ideological Book Corpus
LSTM	Long Short-term Memory
MAE	Mean Absolute Error
MBFC	Media Bias Fact Check
MVDAM	Multi-view document attention model
NLP	Natural Language Processing
RoBERTa	Robustly optimized BERT approach
RNN	Recurrent Neural Network
SPR	Swiss Policy Research
SVM	Support Vector Machine
TFIDF	Term frequency inverse document frequency

1 Motivation

Heutzutage erstellen die Medien im Internet in einem sehr hohem Tempo eine schier unglaubliche Anzahl von Inhalten (Biessmann, 2016). Dabei stellen für viele Menschen Nachrichten, die sie online finden können, eine zentrale Informationsquelle dar (Dallmann et al., 2015). Jedoch können wie alle anderen Medien die Nachrichten einem *Bias* unterliegen, welcher die nicht neutrale Berichterstattung beschreibt (Hamborg et al., 2019). Diese nicht objektive Berichterstattung kann sich in der Wortwahl, der Auswahl zu berichtenden Ereignissen oder der Gestaltung von Bildern bemerkbar machen (Hamborg et al., 2019; Iyyer et al., 2014; Park et al., 2011; Patricia Aires et al., 2019). So kann eine nicht neutrale Berichterstattung die Politik wie zum Beispiel den Ausgang von Wahlen beeinflussen (Dardis et al., 2008; Knight & Chiang, 2008).

Obwohl eine objektive Berichterstattung zum professionellen Journalismus unabdingbar ist, argumentieren einige Wissenschaftler, dass Medien bewusst voreingenommen seien (Groseclose & Milyo, 2005; Matthew Gentzkow & Jesse M. Shapiro, 2010; Park et al., 2011). Für offene Gesellschaften in Demokratien ist es von Relevanz, dass man den politischen *Bias* von Nachrichten erkennen und identifizieren kann, da Medien aufgrund ihrer Informationsfunktion die vierte Macht in Demokratien bilden (Dallmann et al., 2015; Wilke, 2012). Außerdem ist es von Bedeutung, dass Menschen andere Menschen, Medienhäuser oder ausgedrückte Meinungen in ein politisches Spektrum zuordnen können (Zhao et al., 2018). Jedoch ist es äußerst schwierig den politischen *Bias* in Nachrichten überhaupt zu erkennen (Biessmann, 2016; Iyyer et al., 2014; Misra & Basak, 2016; Park et al., 2011; Yan et al., 2017). Als mögliche Erklärung wird angeführt, dass Menschen mit unterschiedlich politischer Orientierung sich in der Wortwahl unterscheiden, wenn sie über dasselbe Thema sprechen und, dass nur Experten in der Lage seien den politischen Bias anhand dieser Differenzen festzumachen (Iyyer et al., 2014; Yan et al., 2017). Dabei greifen die Experten auf Methoden der Politikwissenschaften zurück und analysieren manuell den politischen *Bias* (Dallmann et al., 2015; Hamborg et al., 2019). In Zeiten von Big Data ist es höchst unpraktisch und ineffizient solche Analysen händisch durchzuführen (Iyyer et al., 2014), nichtdestotrotz kann die Technologie hier effektiv eingesetzt

werden um den politischen *Bias* von einer Vielzahl von Nachrichtenartikeln automatisiert zu ermitteln (Aires et al., 2019; Biessmann, 2016).

Das Ziel dieser Masterarbeit ist die Formulierung von Designprinzipien von Modellen, die den politischen *Bias* von englischsprachigen Nachrichtenartikeln automatisiert identifizieren. Diese Modelle basieren auf Methoden des *Machine Learning*. In der Masterarbeit wird ein Design Science Ansatz verfolgt, in der die Erstellung von Artefakten im Vordergrund steht (Recker, 2013).

Die wissenschaftliche Arbeit umfasst insgesamt sechs Kapiteln. In Kapitel 2 werden die theoretischen Hintergründe einiger der zu implementierender Modelle, speziell aus dem Blickwinkel des *Natural Language Processing (NLP)* heraus, erläutert. Anschließend wird in Kapitel 3 das methodische Vorgehen vorgestellt. Im Fokus des vierten Kapitels steht die Präsentation der Ergebnisse der Implementierung sowie die Designprinzipien. Im Anschluss an das Kapitel 4 werden die Ergebnisse im Kapitel 5 einer kritischen Bewertung unterzogen. Abschließend wird im Kapitel 6 das Fazit dieser Arbeit gezogen.

2 Theoretischer Hintergrund

Anders als numerische Datentypen stellen Informationen, die in reiner Textform vorliegen, unstrukturierte Daten dar und lassen sich infolgedessen nicht problemlos verarbeiten (Provost & Fawcett, 2013). Eine Methode den Inhalt von Texten zu repräsentieren ist der *Bag-of-Words* Ansatz (Kowsari et al., 2019; Provost & Fawcett, 2013). Dabei wird im einfachsten Fall die Anzahl der Wörter (*term frequency*) über aller Dokumente im Textkorpus festgehalten. Durch Berücksichtigung von *n-grams* können neben einzelne Wörter auch Wortgruppen der Länge n in dem *Bag-of-Words* aufgenommen werden. Je nach Länge der Wörter werden diese *uni-* ($n=1$), *bi-* ($n=2$) oder *trigrams* ($n=3$) genannt (Kowsari et al., 2019). Der Nachteil dieser Methodik ist, dass die Anzahl der *n-grams* nicht Aufschluss darüber gibt, wie wichtig ein *n-gram* im Textkorpus ist. Die *term frequency inverse document frequency (TFIDF)* ist ein Maß, welches die *n-grams* der Dokumente im Textkorpus gewichtet. Es setzt die Häufigkeit eines *n-grams* in einem Dokument in Relation zu der Anzahl der Dokumente, in dem das *n-gram* enthalten ist (Kowsari et al., 2019). Die *TFIDF* Gewichtungen von *n-grams* können als Eingabevektoren verschiedenen Klassifikatoren wie der logistischen Regression übergeben werden.

Jedoch können Klassifikatoren wie die *Support Vector Machine (SVM)* oder Entscheidungsbäume nicht anhand der *TFIDF* Werte die semantische Information von natürlichen Texten verarbeiten (Malte & Ratadiya, 2019). Diesen Nachteil beseitigen *recurrent neural networks (RNN)* in Verbindung mit sogenannten *Word Embeddings*. *Word Embeddings* sind Methoden, die semantisch ähnliche Wörter in einer Vektorform repräsentieren (Kowsari et al., 2019). Ein weiterer Vorteil dieser Vektorrepräsentation gegenüber dem *Bag-of-Words* Ansatz ist, dass die Dimension der Vektoren viel geringer ist und mathematische Operationen zwischen den Vektoren durchgeführt werden können (Mikolov, Sutskever, et al., 2013). *Word2Vec* (Mikolov, Chen, et al., 2013), *GloVe* (Pennington et al., 2014) und *FastText* (Bojanowski et al., 2017) sind auf großen Datensätzen vortrainierte Repräsentationen von semantisch ähnlichen Wörtern und können als *embedding layer* in die Architekturen von neuronalen Netzen eingebunden werden. *RNN* sind neuronale Netze, welche besonders geeignet sind, um zeitlich strukturierte Informationen wie Zeitreihen oder Sequenzen von Text zu verarbeiten (Elman, 1990; Malte & Ratadiya, 2019). Aufgrund ihrer Architektur können *RNNs* essenzielle Informationen von langen Sequenzen nicht persistieren, da sie während des Trainings unter dem *vanishing* und *exploding gradient problem* leiden (Malte & Ratadiya, 2019). *Long Short-Term Memory (LSTM)* (Hochreiter & Schmidhuber, 1997) und *Gated Recurrent Units (GRU)* (Chung et al., 2014) sind spezielle Varianten der *RNN* Architektur, die beide vorangegangene Probleme bewältigen. In *LSTM* haben Neuronen einen Zustand (*cell state*) und drei *Gates* (*input*, *output*, *forget*). Durch die Interaktionen zwischen den *Gates* wissen die Neuronen, welche Informationen sie verwerfen und behalten sollen. Jedoch nimmt das Trainieren von *LSTM* wegen der vielen Rechenoperationen sehr viel Zeit in Anspruch. Das Trainieren von *GRU* geht durch die Verschmelzung des *input* und *forget Gates* in einem *update Gate* bei gleicher Performanz wesentlich schneller vonstatten (Malte & Ratadiya, 2019).

Encoder und *Decoder* sind *LSTM*, die die Architektur des *RNN Encoder-Decoder* Modells maßgeblich charakterisierten (Cho et al., 2014; Sutskever et al., 2014). Dieses Modell beherrscht die Zuweisung einer Sequenz unbekannter Länge zu einer weiteren Sequenz unbekannter Länge. Es findet Anwendung bei der Übersetzung von Sprachen oder bei der Beantwortung von Fragen (Malte & Ratadiya, 2019; Sutskever et al., 2014). Während der *Encoder* eine Eingabe verarbeitet ist der *Decoder* verantwortlich aus dem übergebenen Kontextvektor die Ausgabe zu generieren. Eine Erweiterung des

Modells stellen die *Attention* Mechanismen dar, welche die Eingabe nicht in einen Vektor mit fixer Länge enkodiert (Bahdanau et al., 2016). Stattdessen wird dem *Decoder* mehrere Kontextvektoren für die gesamte Eingabesequenz mitgegeben. Diese Kontextvektoren enthalten unterschiedliche Gewichtungen für die Wörter aus der Eingabe, welcher der *Decoder* bei der Generierung der Ausgabe berücksichtigt. Durch die Berücksichtigung von *Attention* in der Modellarchitektur konnten gute Resultate im Übersetzen vom Englischen ins Französische erzielt werden (Bahdanau et al., 2016).

Eine performante Architektur ist das *Transformer* Modell, welches vollständig ohne *RNN* auskommt (Vaswani et al., 2017). Sie besteht aus je sechs gestapelten *Encoder* und *Decoder* Blöcken. Jeder *Encoder* Block enthält eine Schicht bestehend aus einem *Multi-Head Attention* gefolgt von einem neuronalen Netz. Die *Decoder* Blöcke enthalten zusätzlich einen vorangestellten *Masked Multi-Head Attention*. Der *Multi-Head Attention* im *Transformer* unterscheidet sich von dem von Bahdanau et al. (2016) vorgestellten *Attention* dahingehend, dass für die Wörter in der Eingabesequenz mehrere Vektoren (*Queries*, *Keys* und *Values*) definiert werden und diese anschließend durch mehrere Matrixoperationen zu einem einzelnen Vektor zusammengefasst werden. Dadurch fasst das *Transformer* Modell für ein Wort mehrere semantische Bedeutungen in einer einzelnen Repräsentation zusammen.

Basierend auf der Architektur des *Transformer* Modells haben Devlin et al. (2019) das *Bidirectional Encoding Representations from Transformers (BERT)* Modell vorgestellt. Das *BERT* Modell stellt eine vortrainierte Architektur dar, welche durch die Implementierung einer zusätzlichen Schicht (*output layer*) für verschiedene Anwendungsfälle eingesetzt werden kann (sog. *fine-tuning*). Das Vortrainieren (sog. *pre-training*) des *BERT* Modells erfolgte anhand der Vorhersage eines maskierten Wortes in einem Satz (*masked language modeling*) und der Klassifikation eines nachfolgenden Satzes anhand eines vorangegangenen Satzes (*next sentence prediction*). In der Arbeit trainieren die Autoren zwei *BERT* Modelle vor, welche sich in der Anzahl der vortrainierten Parameter in der Architektur unterscheiden. Die Basis Architektur weist 110 Millionen Parameter auf, während die große Architektur knapp etwas mehr als dreimal so viele Parameter enthält. Auf der Grundlage des *BERT* Modells wurde das *DistilBERT* Modell abgeleitet, welches auf der geringeren Parameteranzahl die Trainingszeit erheblich verkürzt (Sanh et al., 2020). Das *Robustly optimized BERT approach (RoBERTa)* Modell hingegen trainiert das ursprüngliche

BERT Modell zum einen länger und zum anderen auf einem größeren Datensatz, sodass die Modellgüte des *BERT* Modell übertroffen wird (Liu et al., 2019).

3 Methodik

Das methodische Vorgehen dieser Masterarbeit kann in Abbildung 1 entnommen werden und lässt sich in die drei einzelnen Kategorien Literaturrecherche (Kapitel 3.1), Machbarkeitsstudie (Kapitel 3.2) und Implementierung (Kapitel 3.3) zerlegen.

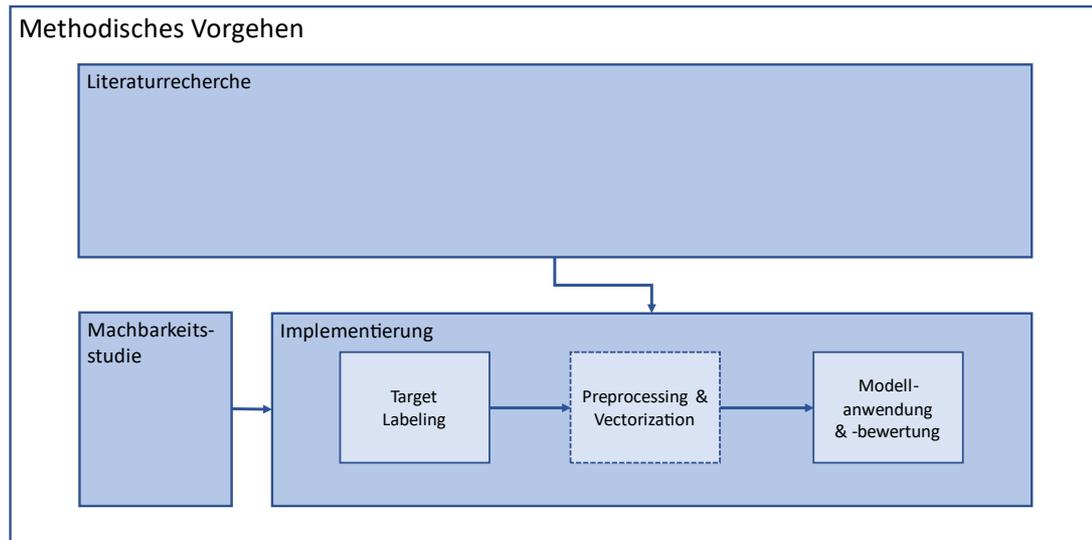


Abbildung 1: Methodisches Vorgehen

In der Literaturrecherche (Kapitel 3.1) werden die Erkenntnisse von Arbeiten, die die politische Richtung von Texten wie zum Beispiel Nachrichten oder Tweets identifizieren, synthetisiert und als Anregung für die Implementierung benutzt. Außerdem sind in der Literaturrecherche Arbeiten inkludiert, in denen Texte die Trainingsdaten darstellen und die zu klassifizierende Objekte nicht unbedingt Nachrichten sind. Zur Literaturrecherche werden keine Exklusionskriterien gesetzt, d.h. es werden auch solche Arbeiten berücksichtigt, die die politische Richtung nicht unbedingt mittels *Machine Learning* ermitteln.

Parallel zur Literaturrecherche findet die Suche und Evaluierung nach geeigneten Datensätzen sowie diverser Mediennavigatoren statt. Letzteres wird benötigt, um die Nachrichten in ein politisches Spektrum zuzuordnen. Diese Schritte stellen die Machbarkeitsstudie (Kapitel 3.2) dar, in der abgeglichen wird, ob die im Internet verfügbaren Datensätze ausreichend Nachrichten über alle politischen Klassen hinweg enthalten.

Im Anschluss an die Machbarkeitsstudie findet die Implementierung diverser Modelle (Kapitel 3.3) statt. Die Implementierung umfasst vor der eigentlichen Modellanwendung auch das *Target Labeling* und das optionale *Preprocessing* und die *Vectorization*. Im *Target Labeling* (Kapitel 3.3.1) wird jeder Nachricht basierend auf dem Resultat der Machbarkeitsstudie eine politische Klasse zugewiesen, während im *Preprocessing & Vectorization* (Kapitel 3.3.2) die Nachrichten bereinigt werden. Das *Preprocessing* und die *Vectorization* beinhaltet mehrere erforderliche Schritte, um die reinen Nachrichten in Textform in eine Form zu bringen, die für Modelle notwendig sind, die auf dem *Bag-of-Words* Ansatz basieren. Abschließend erfolgt die Anwendung der Modelle sowie die Bewertung der Modellgüte (Kapitel 3.3.3). Die einzelnen Teilaufgaben während der Implementierung können in Abbildung 2 entnommen werden.

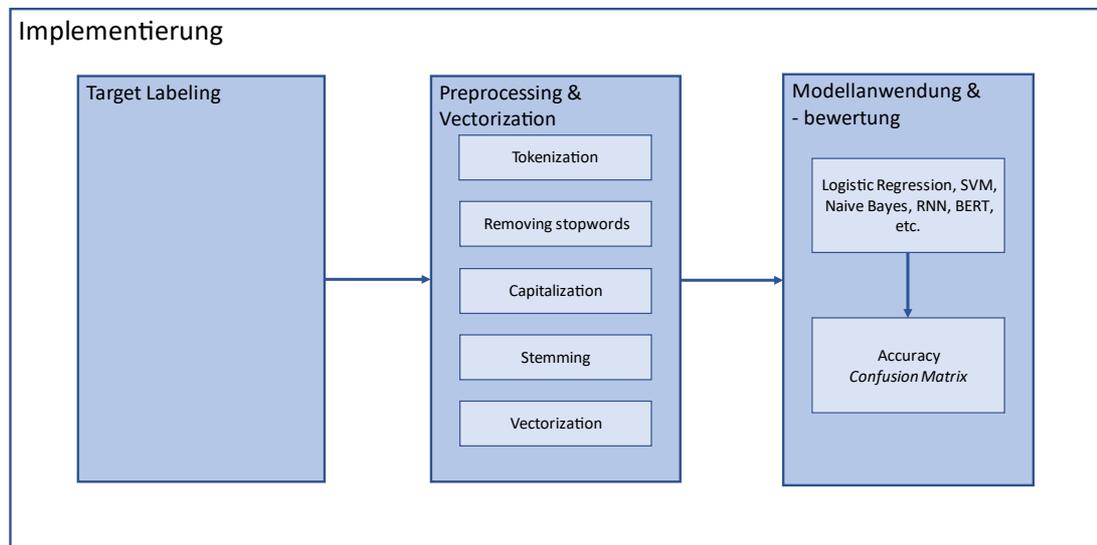


Abbildung 2: Vorgehen während der Implementierung

3.1 Literaturrecherche

Die Klassifizierung der politischen Richtung wurde in der jüngeren Vergangenheit mittels Verfahren, die auf *Machine Learning* (Kapitel 3.1.1), der Graphentheorie (Kapitel 3.1.2), statistischen Methoden (Kapitel 3.1.3) oder Sentimentanalysen (Kapitel 3.1.4) basieren, durchgeführt. Im Fokus der Arbeiten lag hauptsächlich die Klassifikation von Nachrichtenartikeln, Tweets oder politische Reden. Es sind auch Arbeiten inkludiert, in denen die eben erwähnten Texten die Trainingsdaten der Modelle bilden. Tabelle 1 zeigt eine Übersicht die in der Literaturrecherche synthetisierten Arbeiten sowie die angewandten Verfahren als auch die Art des Textes auflistet. Bei den meisten Arbeiten werden die Verfahren näher durch einen speziellen

Vorhersage der politischen Ausrichtung von Nachrichtenartikeln mittels Machine Learning

verwendeten Algorithmus oder durch die Bezeichnung der Autoren ihrer Modelle beschrieben.

Tabelle 1: Übersicht der Literaturrecherche

Paper	Verfahren	Details zum Verfahren	Domäne des Textes
(Ahmed & Xing, 2010)	Statistisch	Multi-view LDA	Blogposts
(Aires et al., 2019)	Graphentheorie	Community detection algorithm	Nachrichten
(Baly et al., 2018)	Machine Learning	SVM	Nachrichten
(Baly et al., 2019)	Machine Learning	Multi-task Ordinal Regression	Nachrichten
(Biessmann, 2016)	Machine Learning	Logistische Regression	Politische Reden
(Chun et al., 2019)	Machine Learning	SVM, Neuronale Netze, BERT	Tweets
(Cohen & Ruths, 2013)	Machine Learning	SVM	Tweets
(Conover et al., 2011)	Machine Learning	SVM	Tweets
(Dallmann et al., 2015)	Sentimentanalyse	-	Nachrichten
(Dumitru & Rebedea, 2019)	Machine Learning	Logistische Regression, SVM, Random Forest	Nachrichten
(Gangula et al., 2019)	Machine Learning	Headline Attention Model	Nachrichten
(Golbeck & Hansen, 2014)	Statistisch	-	Twitter Benutzer
(Hausladen et al., 2020)	Machine Learning	Logistische Regression, Ridge Regression, SVM	Meinung eines Richters
(Hirst et al., 2010)	Machine Learning	SVM	Politische Reden

Vorhersage der politischen Ausrichtung von Nachrichtenartikeln mittels Machine Learning

(Iyyer et al., 2014)	Machine Learning	Neuronale Netze (RNN)	Politische Reden, extrahierte Sätze
(Kim & Lee, 2014)	Graphentheorie	-	Nachrichten
(Kulkarni et al., 2018)	Machine Learning	MVDAM	Politische Reden
(Lin et al., 2008)	Statistisch	Joint Topic and Perspective Model	Politische Reden, Nachrichten
(Maynard & Funk, 2012)	Sentimentsanalyse	-	Tweets
(Misra & Basak, 2016)	Machine Learning	Neuronale Netze (LSTM)	Politische Reden, extrahierte Sätze
(Park et al., 2011)	Sentimentanalyse	Simple Sentiment Classifier	Nachrichten
(Potthast et al., 2018)	Machine Learning	-	Nachrichten
(Preoțiu-Pietro et al., 2017)	Machine Learning	Lineare Regression	Tweets
(Sapiro-Gheiler, 2018)	Machine Learning	Entscheidungsbaum, SVM, Lasso-Regression, Naive Bayes	Politische Reden
(Stukal et al., 2019)	Machine Learning	Neuronale Netze	Tweets
(Volkova et al., 2014)	Graphentheorie	Batch und Streaming Models	Tweets
(Yan et al., 2017)	Machine Learning	Logistische Regression	Tweets
(Yu et al., 2008)	Machine Learning	SVM, Naive Bayes	Politische Reden
(Zhao et al., 2018)	Machine Learning	Neuronale Netze	Nachrichten
(Zhou et al., 2011)	Graphentheorie	Label Propagation Algorithms	Nachrichten

Bemerkung: Leere Felder in der Spalte „Details zu Verfahren“ bedeuten, dass die Arbeit keine weiteren Informationen diesbezüglich spezifiziert.

3.1.1 Verfahren basierend auf Machine Learning

Im Folgendem werden eine Reihe von Forschungen präsentiert, in welcher die Autoren diverse *Machine Learning* Modelle implementiert haben.

Baly et al. (2018) trainieren in ihrer Arbeit ein auf der *SVM* basierendes Modell, um neben der Vertrauenswürdigkeit von Nachrichtenquellen deren politische Richtung gesondert vorherzusagen. Die Autoren *crawkten* insgesamt 94814 Nachrichten von über 1000 Nachrichtenquellen. Sie reichern ihre Daten mit Attributen an und gruppieren diese in die Kategorien *Article*, *Twitter*, *Wikipedia*, *URL* und *Traffic*. Die Kategorie *Article* enthält 141 Attribute und sind an die Attribute des NELA2017 Datensatzes von Horne et al. (2018) angelehnt. Die Autoren führen zwei Experimente durch. Bei dem ersten Experiment trainieren sie ihr Modell nur auf jeweils eine Kategorie von Attributen, während beim zweiten Experiment das Modell mit allen fünf Kategorien und mit vier von fünf Kategorien trainiert wird (d.h. jeweils eine Kategorie wird ausgelassen). Im ersten Experiment schneidet das Modell, welches auf der Kategorie *Article* basiert, am besten ab und erzielt eine *Accuracy* von 41,74%. Im zweiten Experiment erzielt das Modell, welches alle Attribute berücksichtigt, eine *Accuracy* von 39,87% bei einer Klassifikation der politischen Richtungen in sieben Klassen und eine *Accuracy* von 68,86% bei einer Klassifikation der politischen Richtungen in drei Klassen. In beiden Klassifikationen nimmt die *Accuracy* am meisten ab, wenn unter allen Kategorien die Kategorie *Article* ausgelassen wird.

In ihrer nachfolgenden Arbeit trainieren Baly et al. (2019) ein Modell auf der gleichen Datengrundlage wie in ihrer vorherigen Arbeit (Baly et al., 2018). Ihr Modell ist diesmal nun in der Lage die politische Richtung und die Vertrauenswürdigkeit gemeinsam mittels einer *Multi-task Ordinal Regression* zu klassifizieren. Darüber hinaus verwenden die Autoren vier Hilfsaufgaben (sog. *auxiliary tasks*), um die sieben Klassen der politischen Richtung gröber zusammenzufassen. So wurden die sieben Klassen etwa zu fünf, drei oder zwei Klassen vereint. Die Hilfsaufgaben wurden sowohl gesondert als auch gemeinsam bei der Anwendung des Klassifikators angewandt. Der *Mean Absolute Error (MAE)* beträgt bei der Klassifikation der politischen Richtung ohne jegliche Hilfsaufgaben bei 1,798 und verbessert sich durch die Hinzunahme der Hilfsaufgabe, die die Vertrauenswürdigkeit der Nachrichtenquelle darstellt, auf einen *MAE* von 1,584. Der beste *MAE* in Höhe von 1,475 wird erzielt, wenn neben der Darstellung der Vertrauenswürdigkeit der

Nachrichtenquellen die Zuordnung der politischen Klassen sowohl in drei als auch in fünf Klassen gemeinsam zusammengefasst wird.

Mit der Identifikation der politischen Richtung von Twittern Konten haben sich Preoțiu-Pietro et al. (2017), Conover et al. (2011), Stukal et al. (2019) und Chun et al. (2019) auseinander gesetzt.

Preoțiu-Pietro et al. (2017) entwickeln ein Modell, das neben der politischen Aktivität die politische Ausrichtung von Benutzern anhand ihrer Tweets in sieben Klassen vorhersagt. In einem ersten Experiment zeigen die Autoren, dass Cluster basierend auf *Word2Vec* Vektoren die höchste Vorhersagegenauigkeit liefern, um die politische Ausrichtung von Benutzer festzustellen. Ferner wird in einem weiteren Experiment eine logistische Regression trainiert, welches eine *Accuracy* von 22,2% erzielt. Diese *Accuracy* ist um 2,6 Prozentpunkte höher als die des *Baseline* Modells.

Mit der Klassifikation der politischen Orientierung von Twitter Nutzern befassen sich ebenfalls Conover et al. (2011). Sie implementieren zwei *SVM*, welche die politische Richtung eines Nutzers zum einen anhand seines Tweets und zum anderen anhand der benutzten Hashtags vorhersagen. Die *SVM* verzeichnen eine *Accuracy* von 79,2% bzw. 90,8%. Außerdem konstruieren die Autoren wie Volkova et al. (2014) einen Graphen, welches die Beziehung zwischen den Nutzern untereinander abbildet. So sind zwei Nutzer in dem Graphen mit einer Kante verbunden, falls diese sich in ihren Tweets erwähnen oder den Tweet des anderen Nutzers teilen. Auf Grundlage des Graphen wenden die Autoren einen Algorithmus an, die die politische Richtung eines Nutzers anhand der politischen Gesinnung seiner Nachbarn ermittelt. Obwohl dieser Algorithmus eine um fünf Prozentpunkte bessere *Accuracy* erzielt, bedarf es die Notwendigkeit mit jedem neuen Nutzer den Graphen zu aktualisieren.

Stukal et al. (2019) beschäftigen sich mit der Identifikation der politischen Richtung von Twitter Nutzern, die von Bots geführt werden. Die Autoren implementieren diverse künstliche neuronale Netze, die sich in ihrer Anzahl an Schichten und Knoten innerhalb der Schichten unterscheiden. Die Aufgabe der neuronalen Netze ist es einen identifizierten Bot entweder zur russischen Regierung, zur russischen Opposition, zur ukrainischen Regierung oder zu keiner Regierung nahestehend zu klassifizieren. Das beste künstliche neuronale Netz ist charakterisiert durch drei Schichten mit über 300000 Parametern, welcher auf dem Testdaten für die einzelnen Klassen *Precision* und *Recall* Werte von mindestens 91% erzielt. Auf zwei

vom Modell nicht gesehenen Datensätzen klassifiziert das Modell nur 6 von 238 und 3 von 162 Instanzen falsch.

Chun et al. (2019) implementieren in ihrer Arbeit mehrere Modelle, um vorherzusagen, ob ein Tweet von einem russischen Troll verfasst worden ist. In diesem Fall versuchen nachfolgende Klassifikatoren die politische Gesinnung des Trolls zu ermitteln. Ihr Datensatz besteht aus etwas weniger als drei Millionen Tweets von 2848 russischen Trollen und ebenfalls drei Millionen zufälligen Tweets, die von normalen Twitter Konten stammen. Für diese beiden Klassifikationsaufgaben implementieren die Autoren verschiedene Modelle, wie zum Beispiel eine *SVM*, ein künstliches neuronales Netz, zwei *Convolutional Neural Networks (CNN)* und das von Devlin et al. (2019) entwickelte *BERT* Modell. Alle fünf Modelle erzielen bei der Identifikation der politischen Richtung der Trolle eine *Accuracy* zwischen 84% und 89%. In dieser Klassifikation schneidet das *BERT* Modell mit einer *Accuracy* von 89% am besten ab. Darüber hinaus erzielt das *BERT* Modell auch eine *Accuracy* von 99% bei der Klassifikation, ob ein Tweet von einem Troll verfasst worden ist, während die übrigen Modelle *Accuracy* Werte zwischen 56% und 84% verzeichnen.

Die Identifikation der Parteizugehörigkeit von Politikern ist Gegenstand der Arbeiten von Biessmann (2016), Yu et al. (2008), Sapiro-Gheiler (2018) und Hirst et al. (2010) gewesen.

Biessmann (2016) implementiert eine logistische Regression und zeigt, dass es möglich ist die Parteizugehörigkeit von deutschen Politikern vorherzusagen. Der Autor trainiert sein Modell mit den Transkripten der Reden im Bundestag und beschränkt sich dabei auf die 17. und 18. Legislaturperiode. Dabei testet er sein Modell neben den Reden im Bundestag zusätzlich auf den aus Parteiprogrammen stammenden Sätzen. In der 17. Legislaturperiode werden fünf Klassen (jeweils die Parteien CDU/CSU, FDP, Grüne, Linke und SPD) vorhergesagt, während es in der nachfolgenden Legislaturperiode nur vier Klassen sind (CDU/CSU, Grüne, Linke und SPD). Das Modell erzielt eine *Precision* von 66% und 64% auf den Testdaten der Reden im Bundestag für die 17. und 18. Legislaturperiode. Für die Klassifikation der Sätze aus dem Parteiprogramm erzielt das Modell (bezogen auf die 17. Legislaturperiode) eine *Precision* von 35%, welche sich jedoch auf 88% verbessert, wenn man die Sätze zu Kategorien wie Wirtschaft, politisches System, Außenpolitik oder Freiheit und Demokratie zusammenfasst.

Yu et al. (2008) implementieren mehrere Modelle, um auch die Parteizugehörigkeit von Mitgliedern im US-Senat und Repräsentantenhaus vorherzusagen. Im Vordergrund ihrer Arbeit steht jedoch dabei die Zeit- und Personenabhängigkeit der Klassifikatoren. Es werden drei *SVM* und zwei Naive Bayes Klassifikatoren angewandt, die auf verschiedenen *Bag-of-Words* Ansätzen wie der *TFIDF* basieren. Die Personenabhängigkeit der Klassifikatoren wird in einem Experiment untersucht, in welchem die Reden im Senat als Trainingsdaten und die Reden des Repräsentantenhauses als Testdaten dienen. In einem weiteren Schritt wurden die Trainings- und Testdaten miteinander vertauscht. Das Experiment zeigt, dass die Modelle besser abschneiden, wenn sie auf den Reden des Repräsentantenhauses als auf den Reden des Senats trainiert wurden. Beispielsweise erzielt die *SVM*, die auf dem *TFIDF* Ansatz beruht, eine *Accuracy* von 81%. Diese personenunabhängigen Modelle wurden in einem weiteren Experiment auf ihre Zeitunabhängigkeit überprüft. Dabei wurden die Modelle auf Reden des Senats getestet, die in dem Zeitraum von 1989 bis 2006 gehalten wurden. Die Ergebnisse zeigen, dass die Modellgüte der Klassifikatoren zeitabhängig ist. Die *Accuracy* Werte für die *SVM* variieren beispielsweise von 39,4% bis 87%.

Sapiro-Gheiler (2018) implementiert einen Entscheidungsbaum, einen Naive Bayes Klassifikator, eine *SVM* und eine Lasso-Regression, um ebenfalls die Parteizugehörigkeit von US-Politikern anhand ihrer Reden im Kongress zu ermitteln. Die Modelle weisen alle eine *Accuracy* von mindestens 72% nach, wobei die Lasso-Regression mit 98,32% am besten abschneidet gefolgt von der *SVM* mit einer *Accuracy* von ca. 90%.

Auch Hirst et al. (2010) weisen in ihrer Arbeit die Parteizugehörigkeit bzw. die politische Ideologie von kanadischen Politikern anhand ihrer Reden im Parlament nach. Die implementierte *SVM* erzielt *Accuracy* Werte von mindestens 80% für Reden von Politikern des 36. Parlaments. In einem weiteren Experiment weisen die Autoren *Accuracy* Werte zwischen 57% und 89% für Reden von Politikern des 39. Parlaments nach. Da in beiden Parlamenten die Regierung und Opposition aus unterschiedlichen Parteien gebildet wurde, nahmen die Autoren an, dass ihr Klassifikator eher die Regierungszugehörigkeit vorhersagt, anstatt die Parteizugehörigkeit. Diese These belegten die Autoren, indem sie zeigten, dass die Reden der Politiker nicht durch ihre ideologische Überzeugung gekennzeichnet sind, sondern durch Verteidigen der Parteifreunde und Kritisieren der Mitglieder anderer Parteien.

Der Fokus der Arbeiten von Dumitru & Rebedea (2019) und Potthast et al. (2018) liegt in der Klassifikation von überparteilichen Nachrichten.

Dumitru & Rebedea (2019) verfolgen in ihrer Arbeit das Ziel neben Falschmeldungen (sog. *Fake News*) überparteiliche Nachrichten in Echtzeit zu klassifizieren. Dafür implementieren sie eine logistische Regression, eine *SVM* und einen *Random Forest*, die alle *Accuracy* Werte von mindestens 90% auf den Trainingsdaten aufweisen. Die Modelle werden auf ungesehenen Nachrichten angewandt, die aus den Links von Tweets entnommen wurden. Die *Accuracy* der Modelle nimmt auf realen Nachrichten stark ab und beträgt nur noch zwischen 5,92% und 34,81%.

Potthast et al. (2018) beschäftigen sich auch in ihrer Arbeit mit der Vorhersage von überparteilichen Nachrichten. Sie implementieren ein Klassifikator, welcher auf dem Schreibstil der Autoren basiert und testen diesen gegen einen Themen-basierten und einen Naive Bayes Klassifikator. In dem ersten Experiment überprüfen die Autoren, ob die Modelle neutrale, linke oder rechte Nachrichten identifizieren können. Das Modell der Autoren erzielt eine *Accuracy* von 60% und schneidet um 4 Prozentpunkte schlechter ab als das beste Modell. In einem weiteren Experiment untersuchen die Autoren, ob die Modelle generell zwischen neutrale und überparteiliche Nachrichten unterscheiden können. Diesmal schneidet ihr Modell mit einer *Accuracy* von 75% am besten ab und der Vorsprung zum zweitbesten Modell beträgt wieder vier Prozentpunkte.

In ihren Arbeiten zeigen Yan et al. (2017) und Cohen & Ruths (2013), dass die *Accuracy* ihrer Modelle stark von der Domäne der Trainings- und Testdaten abhängt.

Yan et al. (2017) gehen der Frage nach, ob ein Klassifikator in der Lage ist die politische Ideologie von Texten einer Domäne vorherzusagen, auf denen er nicht trainiert worden ist. In ihrem Experiment trainieren die Autoren auf drei unterschiedlichen Trainingsdatensätzen eine logistische Regression, die auf dem *Bag-of-Words* Ansatz basiert. Der erste Datensatz enthält Reden von liberalen und konservativen US-Politikern, während der zweite Datensatz Nachrichten von der liberalen Webseite [salon.com](https://www.salon.com/)¹ und der konservativen Webseite townhall.com²

¹ <https://www.salon.com/>

² <https://townhall.com/>

beinhaltet. Texte aus dem liberalen Wiki RationalWiki³ und dem konservativen Wiki Conservapedia⁴ bilden den dritten Datensatz. Die Autoren stellen fest, dass ihr Modell *Accuracy* Werte von 82,99%, 91,93% und 84,93% (für die obigen Datensätze in gleicher Reihenfolge) erzielt, wenn die Testdaten aus der gleichen Domäne stammen wie die Trainingsdaten. Ihr Modell ist jedoch nicht in der Lage die politische Richtung der Daten aus anderen Domänen vorherzusagen. So liegt zum Beispiel die *Accuracy* des auf den Reden von US-Politikern trainiertem Modell bei dem aus den Wikis stammenden Testdaten nur bei 49,4%.

Cohen & Ruths (2013) zeigen in ihrer Arbeit, dass vorherige Arbeiten, die die politische Richtung von Twitter Nutzern anhand ihrer Tweets ermitteln, viel zu hohe *Accuracy* Werte nachweisen. Sie bemängeln die Datengrundlage vorheriger Arbeiten, weil die Tweets nicht nach verschiedenen Personengruppen differenziert worden sind. Daher legen die Autoren in ihrer Arbeit drei Datensätze zugrunde, die Tweets von homogenen Personengruppen beinhalten. Der erste Datensatz besteht aus Tweets von liberalen und konservativen US-Politikern. Der zweite Datensatz enthält Tweets von Twitter Nutzern, die ihre politische Richtung offenlegen. Der letzte Datensatz beinhaltet Tweets von Nutzern mit unbekannter politischer Richtung. Die Autoren implementieren eine *SVM* und weisen für den ersten Datensatz eine *Accuracy* von 91%, für den zweiten Datensatz eine *Accuracy* von 84% und für den dritten Datensatz eine *Accuracy* von 68% nach. In einer weiteren Klassifikation begutachten die Autoren die Modellgüte des *SVM*, wenn dieser die politische Richtung von Benutzer vorhersagen soll, auf denen er nicht trainiert worden ist. Sie stellen fest, dass die *Accuracy* nicht nur stark abnimmt, sondern auch zum Teil schlechtere Ergebnisse liefert als ein Klassifikator, der die politische Richtung zufällig bestimmt.

In den Arbeiten von Iyyer et al. (2014), Misra & Basak (2016) und Zhao et al. (2018) wird bei der Klassifikation auf die Architektur des *RNN* und im Speziellen auf die *LSTM* zurückgegriffen.

Iyyer et al. (2014) klassifizieren die politische Ideologie mithilfe eines *RNN* und benutzen dabei *Word2Vec* Vektoren, um die semantische Bedeutung von Wörtern abzubilden. Die Autoren trainieren ihr Modell auf dem *Convote* (Thomas et al., 2006) Datensatz und dem *Ideological Book Corpus (IBC)* (Gross et al., 2013), um die

³ https://rationalwiki.org/wiki/Main_Page

⁴ https://www.conservapedia.com/Main_Page

politischen Reden von US-Politikern (*Convote*) und die aus den Magazinen und Büchern extrahierten Sätze (*IBC*) als liberal oder konservativ zu klassifizieren. Das Modell wird gegen diverse logistische Regressionen getestet, die entweder auf dem *Bag-of-Words* Ansatz oder *Word Embeddings* basieren. Die Resultate zeigen, dass auf beiden Datensätzen das *RNN* eine höhere *Accuracy* von mindestens drei Prozentpunkten erzielt. Außerdem zeigt das Ergebnis, dass die auf dem *Word Embeddings* basierende logistische Regression im *IBC* Datensatz besser abschneidet als die übrigen logistische Regressionen.

Misra & Basak (2016) verwenden ebenfalls wie Iyyer et al. (2014) den *IBC* Datensatz, um mithilfe einer *LSTM* die politische Richtung der Reden als konservativ oder liberal zu klassifizieren. Sie bewerten ihre Modellgüte im Vergleich zu Iyyer et al. (2014) anhand des F1 Wertes und der *area under the curve (AUC)*. Ihr Modell wird mit einem Naive Bayes Klassifikator verglichen. Auf dem *IBC* Datensatz erzielt der *LSTM* einen F1-Wert von 57%, während hingegen der Naive Bayes Klassifikator nur einen F1-Wert von ca. 30% nachweisen kann. Der F1-Wert für den *LSTM* beträgt bei der Klassifikation von politischen Reden ungefähr 72%, wobei diesmal der Naive Bayes Klassifikator einen etwas geringeren F1-Wert von 70% erzielt.

Zhao et al. (2018) trainieren zwei auf neuronalen Netzen basierende Klassifikatoren und versuchen Nachrichtenartikeln als neutrale, liberale und konservative Nachricht zu klassifizieren. In der Klassifikation kommen sowohl der *CNN* als auch der *LSTM* jeweils zum Einsatz, welche sich die semantische Bedeutung der Nachrichten zunutze machen. Bei der Anwendung der Modelle werden den neuronalen Netzen bis zu 4 Schichten hinzugefügt. Beide Klassifikatoren erzielen *Accuracy* Werte von mindestens 80%, wobei der *CNN* mit vier Schichten die höchste *Accuracy* von 83,04% und der *CNN* mit zwei Schichten die niedrigste *Accuracy* von 80,66% erzielt.

Sowohl Kulkarni et al. (2018) als auch Gangula et al. (2019) ermitteln den politischen *Bias* von Nachrichten, in dem sie in ihren Modellen die *Attention* Mechanismen berücksichtigen (Bahdanau et al., 2016).

Kulkarni et al. (2018) entwickeln in ihrer Arbeit das *Multi-view document attention model (MVDAM)*, um die politische Richtung von *gecrawlten* Nachrichtenartikeln aus 59 Quellen in die Kategorien *Left*, *Center* und *Right* zu klassifizieren. Ihr *MVDAM* modelliert neben den linguistischen Informationen des Titels und des Textkörpers zusätzlich die Hyperlinkstruktur des Artikels in einem zweischichtigen *Feed-Forward*

Network. In ihrem Experiment wird die Bewertung des *MVDAM* anhand der *Precision*, *Recall* und *F1* Werte vorgenommen. Die Autoren vergleichen ihr Modell mit den einzelnen Komponenten des *MVDAM* sowie mit einer logistischen Regression. Das Experiment zeigt, dass das *MVDAM* der logistischen Regression und seinen einzelnen Komponenten überlegen ist. Ferner wird gezeigt, dass unter allen Komponenten, die alleinige Modellierung des Textkörpers die höchsten *Precision*, *Recall* und *F1* Werte erzielt.

Gangula et al. (2019) entwickeln in ihrer Arbeit das *Headline Attention Model*, um die Partei zu ermitteln, die eine Nachricht in ihrer Berichterstattung favorisiert. Die Autoren nehmen an, dass sich schon in der Nachrichtenüberschrift die favorisierte Partei ableiten lässt. Für ihr Experiment sammeln die Autoren 1329 Nachrichten. Die Modellgüte des *Headline Attention Model* wird gegen einen Naive Bayes Klassifikator, eine *SVM* und diversen künstlichen neuronalen Netzen verglichen. Ihr *Headline Attention Model* erzielt eine *Accuracy* von 89,54% und übertrifft das zweitbeste Modell um 4,22 Prozentpunkte.

Als einzige Forschung implementieren Hausladen et al. (2020) diverse Klassifikatoren, um zu ermitteln, ob ein US-amerikanischer Richter eine liberale oder konservative Entscheidung getroffen hat. Unter den Klassifikatoren befinden sich eine logistische Regression, ein *Ridge* Klassifikator und eine *SVM*, die unter anderem auf den schriftlichen Meinungen der Richter trainiert wurden. Die Klassifikatoren werden sowohl auf allen richterlichen Meinungen als auch auf einer Teilmenge der Meinungen, die sich zum Beispiel nur auf Kriminaldelikte beziehen, angewandt. Der *Ridge* Klassifikator erzielt eine *Accuracy* von 62% auf der Gesamtheit aller richterlichen Meinungen, während die *Accuracy* um sieben Prozentpunkte abnimmt, wenn der *Ridge* Klassifikator nur auf Meinungen bzgl. Kriminaldelikten angewandt wird. Die *Accuracy* der *SVM* ohne Einschränkungen auf die richterlichen Meinungen beträgt 67%.

3.1.2 Graphentheorie-basierte Verfahren

Die nachfolgenden Arbeiten in diesem Abschnitt beschäftigen sich mit Ermittlung der politischen Richtung durch die Modellierung von Zusammenhängen mittels Graphen.

Zhou et al. (2011) sagen die politische Richtung von Nachrichten mithilfe von Graphen vorher. In dem Graph modellieren die Autoren die Interaktionen von

Benutzern der Seite digg.com⁵, einem Nachrichtenportal der Nachrichten aus verschiedenen Quellen zusammenträgt. Die Autoren machen sich zu Nutze, dass Benutzer Nachrichten positiv bewerten können und aus diesem Grund nehmen sie an, dass liberalen Benutzern liberale Nachrichten gefallen und konservative Nachrichten bei konservativen Benutzern Anklang finden. Ausgehend von dieser Annahme werden drei Modelle angewandt, die die politische Richtung von Nachrichten mit einer *Accuracy* von 95,4% bis 96,3% vorhersagen.

Auch die von Aires et al. (2019) verwendete Clusteranalyse basiert auf der Repräsentation von Nachrichten in einem Graphen. Anstelle von Benutzerinteraktionen repräsentieren Kanten im Graphen Hyperlinks in der Nachricht, die zu anderen Links verweisen. Für die Erstellung des Graphen werden zu den fünf politischen Richtungen je vier Nachrichtenquellen als Anfangsknoten hinzugefügt. Bei der Zuordnung der politischen Richtung zu den Nachrichtenquellen sowie der Anzahl der Quellen orientieren sich die Autoren am *Media Bias Fact Check (MBFC)*⁶. Die Autoren führen ihr Modell zweimal aus, wobei sie in der zweiten Ausführung Knoten ohne Zuordnung in eine politische Klasse aus dem Graphen entfernen. Beide Experimente zeigen, dass ihr Modell fünf Cluster identifiziert und, dass jeweils ein Cluster überwiegend Nachrichten aus dem linken politischen Spektrum enthält. In beiden Experimenten weisen die Autoren auch Cluster nach, die Nachrichtenquellen aus allen politischen Richtungen enthalten.

Durch das Verhalten von Twitter Benutzern versuchen Kim & Lee (2014) die politische Richtung von Nachrichten zu klassifizieren. Die Nachrichten werden in die Klassen liberal, konservativ oder moderat zugeordnet. Sie konstruieren einen Graphen bestehend aus Politikern, deren politische Überzeugung anhand ihrer Parteizugehörigkeit abgeleitet wird. Kontinuierlich werden dem Graphen normale Nutzer hinzugefügt, die über die zehn meistgelesenen Nachrichten Tweets absetzen. Die Autoren nehmen an, dass Nutzer mit der politischen Richtung einer Nachricht übereinstimmen, wenn sie diesen in einem Tweet, ohne ihre eigene Meinung wiederzugeben, teilen. Ferner wird angenommen, dass wenn Nutzer B den Tweet von Nutzer A Tweet teilt, so teilt auch Nutzer B die Ansichten bzw. die politische Orientierung von Nutzer A. Dieser Annahme wird verstärkt, wenn beide Nutzer ihre

⁵ <https://digg.com/>

⁶ <https://mediabiasfactcheck.com/>

Tweets mit anderen Nutzern teilen. In einem nächsten Schritt werden sehr aktive Nutzer (*highly active users*) identifiziert. Das sind jene, die sowohl ihre Tweets mit anderen Nutzern teilen als auch Nutzer, deren Tweets von anderen Nutzern geteilt werden. Neben den sehr aktiven Nutzern gibt es noch die kaum aktiven Nutzer (*rarely active users*). Die politische Orientierung von kaum aktiven Nutzern kann anhand einer Sentimentanalyse ihrer Tweets ermittelt werden. Währenddessen wird die politische Richtung von sehr aktiven Nutzern dadurch bestimmt, wie häufig sie mit liberalen, konservativen oder moderaten Benutzern interagieren. Die politische Orientierung eines Nutzers kann nun Aufschluss über die politische Richtung einer Nachricht geben. Falls ein Benutzer ein Tweet über eine Nachricht absetzt, der entweder keinen oder einen positiven Inhalt enthält, so stimmt die politische Richtung der Nachricht mit dem des Benutzers überein. Setzt der Benutzer ein Tweet mit einem negativen Inhalt ab, so stimmt die politische Richtung der Nachricht mit dem des Benutzers nicht überein. Die politische Richtung einer Nachricht hängt endgültig davon ab, ob mehr zustimmende oder ablehnende Tweets über diese Nachricht abgesetzt wurden. Das Modell wird angewandt, um die politische Richtung von Nachrichten anhand von 22875 Twitter Nutzern zu ermitteln. Das Modell kann die politische Richtung von 248 der 280 Nachrichten korrekt klassifizieren und erzielt eine um 27,3 Prozentpunkte höhere *Accuracy* als eine *SVM*.

Volkova et al. (2014) entwickeln ein dynamisches Modell (*Streaming Model*), welches in Echtzeit die politische Überzeugung eines Twitter Nutzers möglichst ressourcenschonend vorhersagt. In einem Graphen werden die Nutzer als Knoten und die Beziehungen unter den Nutzern als Kanten dargestellt. Die Gesamtheit der Nutzer, die mit einem Nutzer in einem Zusammenhang stehen, wird als seine Nachbarschaft definiert. Die Autoren vergleichen ihr *Streaming Model* gegen ein statisches Modell (*Batch Model*), welche nur eine fixe Anzahl von Tweets eines Nutzers berücksichtigt. Das *Batch Model* berücksichtigt in einer weiteren Variante die Anzahl der Nachbarn und ihre Tweets. Das *Streaming Model* unterscheidet sich vom *Batch Model* dahingehend, dass die Klassifikation permanent aktualisiert wird, sobald ein Nutzer und/oder seine Nachbarn neue Tweets verfassen. Das *Batch Model* sagt die politische Richtung besser voraus, wenn ein Nutzer vergleichsweise viele als wenige Tweets geschrieben hat. Des Weiteren ist die *Accuracy* höher, wenn die Nachbarn eines Nutzers mitberücksichtigt werden. Ferner stellen die Autoren fest, dass das Modell bessere Ergebnisse liefert, wenn die Nachbarn eines Nutzers genauso viele Tweets

verfasst haben wie ein Nutzer, der keine Beziehungen zu anderen Nutzern pflegt. Das *Streaming Model* der Autoren erweist sich als eine effektivere Methode als das *Batch Model*. Innerhalb weniger Wochen erzielt das *Streaming Model* eine *Accuracy* von 75% bzw. 90%. Je mehr Tweets vorhanden sind, desto schneller erreicht das *Streaming Model* eine hundertprozentige *Accuracy*.

3.1.3 Statistische Verfahren

In diesem Abschnitt beschäftigen sich die Arbeiten mit der Ermittlung der politischen Ausrichtung basierend auf statistischen Methoden.

Lin et al. (2008) legen in ihrem statistischen Modell die Annahme zugrunde, dass die Auswahl von thematischen Wörtern, denen eine ideologische Richtung nachgewiesen werden kann, abhängig von der ideologischen Richtung des Autors ist. In ihren beiden Experimenten weisen die Autoren erfolgreich nach, dass ihr Modell in der Lage ist die politische Orientierung von liberalen und konservativen US-Präsidentschaftskandidaten anhand deren Reden zu ermitteln.

Ahmed & Xing (2010) führen eine ähnliche Arbeit wie Lin et al. (2008) durch. Die Autoren legen in ihrem entwickelten Modell zugrunde, dass das verwendete Vokabular eines Autors beim Schreiben einer Nachricht sowohl vom Thema selbst als auch vom eigenen Standpunkt diesbezüglich abhängt. In ihrer Arbeit verwenden die Autoren ihr Modell unter anderem, um die politische Richtung von Blogposts zu ermitteln und vergleichen ihr Modell gegen ein *SVM*. Es wurden die zwei Datensätze Blog-1 und Blog-2 erstellt. Blog-1 enthält 2040 Blogbeiträgen aus einem rechten und zwei linken Blog/-s. In Blog-2 befinden sich 13246 Blogbeiträge von je drei linken und rechten Blogs. Ihr Modell schneidet im Blog-1 Datensatz durchgehend ab mindestens 10 Themen besser als die *SVM* und erzielt *Accuracy* Werte zwischen 90% und 95%. Im Blog-2 Datensatz liegt die *Accuracy* des *SVM* bei ungefähr 69% und unterlegt bei knapp 25 Themen dem Modell der Autoren.

Golbeck & Hansen (2014) ermitteln einen P-Wert, der es ihnen ermöglicht die politische Orientierung von Personen angeben zu können, die die Twitter Konten von Nachrichtenanbieter oder Think Tanks folgen. Die Autoren nehmen an, dass die politische Orientierung von Twitter Nutzern mit Twitter Nutzern übereinstimmt, denen sie folgen. Für die Berechnung des P-Werts identifizieren sie zuerst die Twitter Konten von US-Politiker, deren P-Wert bekannt ist. Für die Abgeordneten des Kongresses wird eine Liste von Personen erstellt, denen die Abgeordnete folgen. Der P-Wert eines

Kongressabgeordneten ist der Durchschnitt aller P-Werte der Politiker, denen gefolgt wird. Der Durchschnitt der P-Werte aller Politiker, die einen Nachrichtenanbieter oder einen Think Tank folgen, bestimmt den P-Wert des gefolgten Twitter Kontos. Die politische Richtung eines Twitter Nutzers lässt sich anhand des P-Wertes der gefolgten Twitter Nutzer ableiten.

3.1.4 Sentimentanalysen

Die nachfolgenden Arbeiten klassifizieren die politische Richtung auf der Grundlage von Sentimentanalysen.

Mit der Klassifizierung von Tweets haben sich Maynard & Funk (2012) in ihrer Arbeit beschäftigt. Dafür analysieren sie das Sentiment von vier Millionen Tweets, die von englischen Twitter Benutzern verfasst worden sind. Ihr Modell versucht die Tweets in die Klassen *ProCon*, *AntiCon*, *ProLab*, *AntiLab*, *ProLib*, und *AntiLib* zu zuordnen. Die Klassen stellen dar, ob die Benutzer die *Labour Party*, die *Liberal Democrats* oder die *Conservative Party* befürworten oder ablehnen. Für ihr Experiment verwenden die Autoren 1000 Tweets und ihr Modell identifiziert 143 Tweets als solche, die einen Sentiment gegenüber einer Partei beinhaltet. Eine manuelle Überprüfung der Autoren zeigt, dass sie 31 dieser Tweets nicht zu den sechs Klassen zuordnen würden. Die Autoren weisen für ihr Modell dennoch eine *Precision* von 62,2% nach.

Einen Klassifikator basierend auf einer Sentimentanalyse entwickeln ebenfalls Park et al. (2011). Die Autoren untersuchen, ob der Kommentar und die politische Einstellung des Kommentators Auskunft über die politische Richtung eines Nachrichtenartikels geben kann. Die Datengrundlage bilden Nachrichten und Kommentare, die auf der südkoreanischen Webseite *Naver News*⁷ erschienen sind. Ihren Klassifikator nennen die Autoren *Simple Sentiment Classifier*, welches die Klassen *Conservative*, *Liberal* und *Vague* vorhergesagt. Die Autoren modifizieren ihr Modell hinsichtlich der Klassifikation, ob das Modell für die Vorhersage der politischen Richtung das Sentiment eines einzelnen Kommentators (*single commentar-based prediction*) oder mehrerer Kommentatoren (*multi commentator-based prediction*) analysieren soll. Außerdem werden für die Klassifikation die Nachrichtenartikeln in einem *General Set* und *Popular Set* unterteilt. Im letzteren

⁷ <https://news.naver.com/>

befinden sich die meistgelesenen Nachrichten der letzten sechs Monaten. Ihr Modell erzielt bei der *single commentar-based prediction* eine *Accuracy* von 67% im *General Set* und respektive 66% im *Popular Set*. Durch die Berücksichtigung von mehreren Kommentatoren in der Sentimentanalyse verbessert sich die *Accuracy* des Modells im *Popular Set* von etwa 45% auf bis zu 70%, wenn das Sentiment von zwölf Kommentatoren analysiert wird.

Dallmann et al. (2015) zeigen in ihrer Arbeit anhand einer Sentimentanalyse, dass die Nachrichtenanbieter die FAZ⁸, der Spiegel⁹, die taz¹⁰ und die Zeit¹¹ während der 17. Legislaturperiode durchgehend positiv über die CSU und SPD und negativ über die FDP berichten. Zusätzlich vergleichen die Autoren die Ähnlichkeit des verwendeten Vokabulars zwischen den Medien und den Parteiprogrammen deutscher Parteien und stellen fest, dass das Vokabular der FAZ am ähnlichsten mit den Parteiprogrammen der CDU und der FDP ist, während das Vokabular des Spiegels am meisten dem Parteiprogramm der SPD ähnelt. Unter allen Nachrichtenquellen ist das Vokabular mit dem Parteiprogramm der Grünen am unähnlichsten.

3.2 Machbarkeitsstudie

Die Grundlage für die Implementierung der Modelle in dieser Arbeit bilden zum einen Datensätze und zum anderen auf fundierten Methoden basierende Mediennavigatoren, um den Nachrichten in den Datensätzen eine politische Klasse zuzuweisen. Außerdem wird evaluiert, ob die in dem Datensatz vorhanden Nachrichten über alle potenziellen Klassen hinweg ausgewogen verteilt und, ob ggf. ein *Crawlen* von weiteren Nachrichten notwendig ist. Ferner wird mitangeben, wie viele Nachrichten und Nachrichtenanbieter die Mediennavigatoren abdecken.

Hierfür werden in den kommenden beiden Unterkapiteln drei Mediennavigatoren (Kapitel 3.2.1) und zwei öffentlich verfügbare Datensätze (Kapitel 3.2.2) vorgestellt. Im letzten Unterkapitel (Kapitel 3.2.3) werden die Verteilungen der politischen Klassen in den Datensätze betrachtet.

⁸ <https://www.faz.net/aktuell/>

⁹ <https://www.spiegel.de/>

¹⁰ <https://taz.de/>

¹¹ <https://www.zeit.de/politik/index>

3.2.1 Mediennavigatoren

Die Swiss Policy Research (SPR)¹² ist eine unabhängige gemeinnützige Organisation aus der Schweiz, die die Propaganda in internationalen Medien untersucht. Sie haben je ein Mediennavigator für die deutsch- und englischsprachige Medienlandschaft veröffentlicht. Der Mediennavigator für die englischsprachige Medienlandschaft (siehe Abbildung 3) ist eine 2x2-Matrix, welche auf der x-Achse die Beziehung zur Macht (*relationship to power*) und die politische Ausrichtung der Nachrichtenanbieter auf der y-Achse abbildet. Die politische Ausrichtung wird in die drei Kategorien *Liberal*, *Center* und *Conservative* unterteilt (Swiss Propaganda Research, 2019). Die SPR legt für ihre beiden veröffentlichten Mediennavigatoren nicht die Methode offen, wie sie die Nachrichtenanbieter in die Matrix verortet haben.



Abbildung 3: Mediennavigator von SPR

AllSides¹³ ist eine Gruppe von Menschen mit unterschiedlicher politischer Orientierung, die sich zum Ziel genommen haben Filterblasen zu reduzieren, indem sie zu einem Ereignis Nachrichten von diversen Anbietern präsentieren. Ihr

¹² <https://swprs.org/contact/>

¹³ <https://www.allsides.com/about>

Mediennavigator (siehe Abbildung 4) teilt englischsprachige Nachrichtenanbieter in die fünf politischen Klassen *Left*, *Lean Left*, *Center*, *Lean Right* und *Right* ein (*Media Bias Ratings AllSides*, 2020). Um die Anbieter in eine politische Richtung zu verorten, wenden die Autoren bis zu fünf verschiedene und unterschiedlich gewichtete Methoden an (*How AllSides Rates Media Bias*, 2016). In der *Blind Bias Survey* lesen Menschen aus dem gesamten politischen Spektrum eine Nachricht ohne den Anbieter zu kennen und verorten anschließend diese zu einer politischen Ausrichtung. In der *Third-Party-Analysis* werden unabhängige Forschungen oder Umfragen berücksichtigt. Der *Editorial Review* ist ähnlich aufgebaut wie die *Blind Bias Survey* und unterscheidet sich dahingehend, dass die Leser aus der eigenen Redaktion stammen. Zusätzlich können Leser im *Community Feedback* auf der Webseite selbst eine Klassifizierung des Nachrichtenanbieters vornehmen und mitteilen, ob sie die bestehende Klassifizierung zustimmen oder zurückweisen. Bei einer *Independent Research* analysiert ein Mitglied die Nachrichten eines Anbieters und berücksichtigt dabei Berichte anderer Quellen über die politische Ausrichtung jenes Anbieters. AllSides listet für jeden Nachrichtenanbieter auf, welche Methoden angewandt worden sind und gibt darüber hinaus zusätzlich an, wie zuverlässig ihre Klassifizierung ist. Auf ihrer Webseite befinden sich mehr Nachrichtenquellen, denen eine politische Richtung zugewiesen worden ist, als in Abbildung 4 dargestellt ist.



Abbildung 4: Mediennavigator von AllSides

Vanessa Otero gründete 2018 Ad Fontes Media¹⁴ mit dem Ziel, um bei Menschen und Pädagogen eine Sensibilität beim Nachrichtenkonsum zu schaffen (Ad Fontes Media, o. J.-a). Ihr Mediennavigator bildet die politische Ausrichtung der Medienanbieter auf der x-Achse ab, während die y-Achse die Vertrauenswürdigkeit der Anbieter widerspiegelt (siehe Abbildung 5). Die politische Ausrichtung wird in den sieben Ausprägungen *Most Extreme Left*, *Hyper-Partisan Left*, *Skews Left*, *Neutral or Balanced Bias*, *Skews Right*, *Hyper-Partisan Right* und *Most Extreme Right* angegeben. Zwanzig Analysten mit unterschiedlichen politischen Ansichten haben mehr als 1800 Nachrichten und Fernsehbeiträge von über 100 Medienhäuser und

¹⁴ <https://www.adfontesmedia.com>

Vorhersage der politischen Ausrichtung von Nachrichtenartikeln mittels Machine Learning

Fernsehsender begutachtet. Pro Nachrichtenanbieter wurden mindestens sieben Nachrichten analysiert, während bei großen Nachrichtenanbieter bis zu 80 Nachrichten genauer untersucht worden sind. Um die politische Ausrichtung eines Anbieters zu ermitteln, haben die die Analysten unter anderem die Wortwahl, Bilder und deren Präsentation und den Wahrheitsgehalt von Sätzen genauer betrachtet (Ad Fontes Media, o. J.-b).

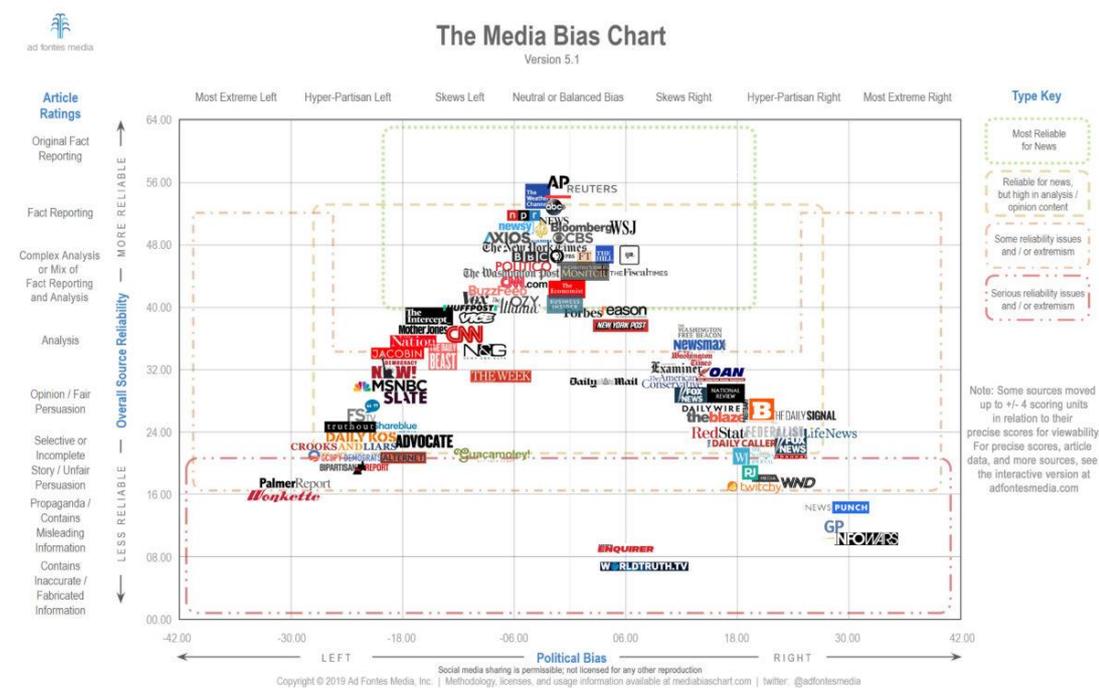


Abbildung 5: Mediennavigator von Ad Fontes Media

3.2.2 Datensätze

Basierend auf ihrer vorherigen Arbeit haben Norregaard et al. (2019) den NELA-GT-2018 Datensatz erstellt. Dieser Datensatz enthält 713534 Nachrichten von 194 verschiedenen Nachrichtenanbieter aus dem Jahre 2018. Der NELA2017 Datensatz, welche Horne et al. (2018) in ihrer vorherigen Arbeit erstellt haben, enthält nur rund ein Fünftel der Nachrichten und deckt nur 92 Nachrichtenanbieter ab. Eine weitere Besonderheit des NELA-GT-2018 Datensatzes ist, dass die Autoren die Nachrichtenanbieter mit weiteren Attributen beschreiben. So wird unter anderem anhand des Mediennavigators von AllSides die politische Ausrichtung der Nachrichtenanbieter mitangegeben. In der Masterarbeit werden 997 Beobachtungen aus dem Datensatz entfernt und nicht weiter berücksichtigt, da sie NULL-Werte enthalten. Die Anzahl der Nachrichten je Anbieter sowie die Zuordnung der

Vorhersage der politischen Ausrichtung von Nachrichtenartikeln mittels Machine Learning

Nachrichtenanbieter in eine politische Klasse anhand der in Kapitel 3.2.1 vorgestellten Mediennavigatoren kann in Tabelle 2 entnommen werden.

Tabelle 2: Anzahl der Artikel je Quelle im NELA-GT-2018 Datensatz

Quelle	# Artikeln	SPR	All Sides	Ad Fontes Media
21stCenturyWire	322	Liberal	-	-
ABC News	2808	Center	Lean Left	Neutral
Activist Post	1797	Conservative	-	-
Addicting Info	429	-	-	-
Al Jazeera	4517	-	Center	-
Alternet	4816	Liberal	Left	Hyper-Partisan Left
AMERICAblog News	42	-	-	-
Anonymous Conservative	616	-	-	-
BBC	16409	Center	Center	Neutral
BBC UK	5501	-	-	-
Bearing Arms	1193	-	-	-
Bipartisan Report	4060	-	-	Hyper-Partisan Left
Birmingham Mail	9241	-	-	-
Breitbart	1877	Conservative	Right	Hyper-Partisan Right
Business Insider	445	-	Center	Neutral
Buzzfeed	1661	Liberal	Lean Left	Skews Left
CBS News	5396	Center	Lean Left	Neutral
Channel 4 UK	888	-	-	-
Chicago Sun-Times	2113	-	Lean Left	-
CNBC	2426	-	Center	-
CNN	8202	Liberal	Lean Left	Neutral
CNS News	5263	-	Right	-
Common Dreams	27	-	-	-
Conservative Home	2248	-	-	-
Conservative Tribune	2353	-	-	-
Counter Current News	23	-	Lean Left	-
Crikey	391	-	-	-
Crooks and Liars	2465	-	-	-
Daily Beast	6634	Liberal	Left	Skews Left
Daily Kos	994	-	Left	Hyper-Partisan Left

Vorhersage der politischen Ausrichtung von Nachrichtenartikeln mittels Machine Learning

Daily Mail	3596	Conservative	Right	Neutral
Daily Signal	308	-	Right	Hyper-Partisan Right
Daily Stormer	1377	-	-	-
DC Gazette	185	-	-	-
Delaware Liberal	1125	-	-	-
Democracy 21	24	-	-	-
Dick Morris Blog	400	-	Lean Right	-
Drudge Report	18884	Conservative	Lean Right	-
Evening Standard	17637	-	-	-
Faking News	220	-	-	-
Feministing Blog	23	-	-	-
FiveThirtyEight	556	-	Center	-
Foreign Policy	701	-	Center	Neutral
Fort Russ	1090	-	-	-
Fortune	7630	-	-	-
Forward Progressives	142	-	-	-
Fox News	3106	Conservative	Lean Right	Skews Right
France24	1732	-	-	-
Freedom Daily	36	-	-	-
Freedom Outpost	321	-	-	-
Freedom-Bunker	2228	-	-	-
FrontPage Magazine	892	-	Right	-
FT Westminster Blog	10	-	-	-
Fusion	141	-	-	-
GlobalResearch	30	-	-	-
Glossy News	61	-	-	-
Hit and Run	3440	-	-	-
Hot Air	4642	-	Lean Right	-
Hullabaloo Blog	954	-	-	-
HumansAreFree	426	-	-	-
Humor Times	282	-	-	-
Informnapalm	32	-	-	-
Infowars	2518	-	Right	Most Extreme Right
Instapundit	15479	-	-	-
Intellectual Conservative	377	-	Lean Right	-
Intellihub	334	-	-	-
Interpreter Mag	28	-	-	-
Investors Business Daily	730	-	Lean Right	-
iPolitics	4253	-	-	-

Vorhersage der politischen Ausrichtung von Nachrichtenartikeln mittels Machine Learning

JewWorldOrder	1493	-	-	-
LabourList	430	-	-	-
LewRockwell	1275	Conservative	-	-
Liberal Democrat Voice	573	-	-	-
Live Action	1054	-	Lean Right	-
Losercom	10	-	-	-
Mail	8461	-	-	-
Media Matters for America	2316	-	Left	-
Mercury News	4826	-	-	-
Mint Press News	1707	Liberal	-	-
MotherJones	1128	Liberal	Left	Skews Left
MSNBC	6604	Liberal	Left	Hyper-Partisan Left
National Review	5129	Conservative	Right	Skews Right
Natural News	4187	-	-	-
New York Daily News	2042	-	Left	-
New York Post	25401	Conservative	Right	Neutral
New Yorker	265	-	Left	Skews Left
News Biscuit	1666	-	-	-
News Busters	3240	-	-	-
Newsnet Scotland	35	-	-	-
Newswars	4275	-	-	-
Newsweek	9411	-	Lean Left	-
NODISINFO	29	-	-	-
NPR	5515	Center	Center	Neutral
oann	14267	Conservative	-	Skews Right
Observer	541	-	Center	-
OSCE	636	-	-	-
Palmer Report	3539	-	-	Hyper-Partisan Left
Pamela Geller Report	367	-	-	-
PBS	1113	-	Center	-
Pink News UK	1645	-	-	-
Politicalite	737	-	-	-
Politico	628	-	Lean Left	Neutral
Politics UK	137	-	Lean Left	-
Politicscoulk	341	-	-	-
Politicus USA	4018	-	Left	-
Powerline Blog	894	-	-	-
Pravada Report	601	-	-	-
Prepare For Change	11	-	-	-

Vorhersage der politischen Ausrichtung von Nachrichtenartikeln mittels Machine Learning

Prison Planet	2253	-	-	-
Raw Story	3719	-	Left	-
Real Clear Politics	7206	-	Center	-
Real News Right Now	13	-	-	-
RedState	4808	-	Right	Hyper-Partisan Right
Reuters	3929	-	Center	Neutral
rferl	2318	-	-	-
RightWingWatch	1118	-	Right	-
RT	4286	-	-	-
Russia-Insider	1030	-	-	-
Salon	1702	-	Left	-
ScrappleFace	61	-	-	-
Shadow Proof	260	-	-	-
Shareblue	2134	-	-	Hyper-Partisan Left
SkyNewsPolitics	826	-	-	-
SkyNewsUS	995	-	-	-
Slate	514	Liberal	Left	Hyper-Partisan Left
Slugger OToole	303	-	-	-
sott.net	9319	-	-	-
Spiegel	4171	-	-	-
Sputnik	30367	-	-	-
Talking Points Memo	5841	-	-	Hyper-Partisan Left
Tass	6160	-	-	-
Telesur TV	860	-	-	-
The American Conservative	439	-	Lean Right	Skews Right
The Atlantic	1757	-	Lean Left	Skews Left
The Beaverton	854	-	-	-
The Borowitz Report	123	-	-	-
The Chaser	131	-	-	-
The Conservative	2120	-	-	-
Tree House				
The D.C. Clothesline	654	-	-	-
The Daily Blog	457	-	-	-
The Daily Caller	11549	Conservative	Right	-
The Daily Echo	3329	-	-	-
The Daily Express	1585	-	-	-
The Daily Mirror	13202	-	-	-
The Daily Record	6981	-	-	-
The Daily Star	219	-	-	-
The Denver Post	4501	-	-	-

Vorhersage der politischen Ausrichtung von Nachrichtenartikeln mittels Machine Learning

The Duran	959	-	-	-
The Fiscal Times	461	-	Lean Right	-
The Gateway Pundit	5667	-	Right	-
The Guardian	2195	Liberal	Lean Left	-
The Guardian UK	16946	-	-	-
The Hill	1968	-	Center	-
The Huffington Post	5586	Liberal	Left	Skews Left
The Huffington Post UK	5855	-	-	-
The Independent	19799	-	Lean Left	-
The Inquisitr	2467	-	-	-
The Intercept	1268	Liberal	Left	-
The Irish Times	3826	-	Right	-
The Manchester Evening News	8447	-	-	-
The Michelle Malkin Blog	53	-	Right	-
The Moscow Times	1137	-	-	-
The New York Times	5471	Center	Lean Left	Neutral
The Onion	1094	-	Lean Left	-
The Poke	1313	-	-	-
The Political Insider	2680	-	-	-
The Right Scoop	2697	-	-	-
The Shovel	223	-	-	-
The Spoof	696	-	-	-
The Sun	43492	-	-	-
The Telegraph	33763	-	Lean Right	-
The Verge	5951	-	Lean Left	-
The Washington Examiner	469	Conservative	Lean Right	Skews Right
The Week UK	2207	-	Center	-
TheAntiMedia	666	-	-	-
TheBlaze	5287	-	Right	-
theRussophileorg	31422	-	-	-
ThinkProgress	4819	-	Left	-
True Activist	370	-	-	-
True Pundit	13646	-	-	-
Trump Times	86	-	-	-
Unian	3312	-	-	-
USA Today	5814	Center	Center	-
Veterans Today	2621	Conservative	-	-

Vorhersage der politischen Ausrichtung von Nachrichtenartikeln mittels Machine Learning

Vox	4288	-	Left	Skews Left
Waking Times	447	-	-	-
Washington Monthly	551	-	Lean Left	Skews Left
Washington Post	1252	Liberal	Lean Left	Neutral
Western Journal	4729	-	Right	-
Window on Eurasia Blog	840	-	-	-
Wings Over Scotland	147	-	-	-
Wizbang	58	-	-	-
WSJ Washington Wire	79	-	-	-
Yahoo News	1665	-	Left	-

Bemerkung: „-“ bedeutet, dass ein Mediennavigator einem Nachrichtenanbieter keine politische Richtung zuweist

Der Kaggle Datensatz wurde von Andrew Thompson auf Kaggle¹⁵ veröffentlicht, jedoch verweist die Seite auch auf eine erweiterte Version des Datensatzes hin („204,000 News Articles and Essays“, o. J.). Der erweiterte Kaggle Datensatz enthält mehr als 200000 Nachrichten von 18 verschiedenen Nachrichtenanbieter, unter anderem von CNN¹⁶, Fox News¹⁷, The Atlantic¹⁸, der New York Times¹⁹ oder Vox²⁰. Ein Großteil der Nachrichten stammt aus den Jahren 2016 und 2017. Aus diesem Datensatz werden 15407 Einträge mit *NULL*-Werten entfernt. In Tabelle 3 ist sowohl die Anzahl der Nachrichten je Nachrichtenanbieter als auch die politische Richtung der Nachrichtenanbieter aufgeführt.

Tabelle 3: Anzahl der Nachrichtenanbieter im Kaggle Datensatz

Quelle	# Artikeln	SPR	All Sides	Ad Fontes Media
Atlantic	7386	-	Lean Left	Skews Left
Breitbart	23791	Conservative	Right	Hyper-Partisan Right
Business Insider	8344	-	Center	Neutral
Buzzfeed News	10699	Liberal	Lean Left	Skews Left

¹⁵ <https://www.kaggle.com/snapcrack/all-the-news>

¹⁶ <https://us.cnn.com/>

¹⁷ <https://www.foxnews.com/>

¹⁸ <https://www.theatlantic.com/world/>

¹⁹ <https://www.nytimes.com/>

²⁰ <https://www.vox.com/>

Vorhersage der politischen Ausrichtung von Nachrichtenartikeln mittels Machine Learning

CNN	14288	Liberal	Lean Left	Neutral
Fox News	4384	Conservative	Lean Right	Skews Right
Guardian	8685	Liberal	Lean Left	-
Los Angeles Times	9034	-	-	-
National Review	6258	Conservative	Right	Skews Right
New Inquiry	3264	-	-	-
New York Post	18144	Conservative	Right	Neutral
New York Times	30141	Center	Lean Left	Neutral
NPR	12001	Center	Center	Neutral
Reuters	10719	-	Center	Neutral
Talking Points Memo	5323	-	-	Hyper-Partisan Left
Verge	1225	-	Lean Left	-
Vox	4953	-	Left	Skews Left
Washington Post	11152	Liberal	Lean Left	Neutral

Bemerkung: „-“ bedeutet, dass ein Mediennavigator einem Nachrichtenanbieter keine politische Richtung zuweist

3.2.3 Analyse des Datensatzes und der Mediennavigatoren

In den Mediennavigatoren von SPR, AllSides und Ad Fontes Media befinden sich 72, 52 und respektive 90 Nachrichtenanbieter (siehe Tabelle 4). Von den 194 Nachrichtenanbietern im NELA-GT-2018 Datensatz deckt der Mediennavigator von SPR mit 31 Nachrichtenanbieter die wenigsten und der Mediennavigator von AllSides mit 75 Nachrichtenanbieter die meisten Nachrichtenanbieter ab (siehe Tabelle 5). Im Kaggle Datensatz zeichnet sich dasselbe Bild ab. Der Mediennavigator von SPR deckt mit 10 Nachrichtenanbieter die geringste Anzahl von Nachrichtenanbieter ab, während der Mediennavigator von AllSides die höchste Anzahl von Nachrichtenanbieter aufführt (siehe Tabelle 6).

Tabelle 4: Anzahl der Nachrichtenanbieter je Mediennavigator

Mediennavigator	Anzahl Nachrichtenanbieter
SPR	72
AllSides	52
Ad Fontes Media	90

Tabelle 5: Anzahl der Nachrichtenanbieter im NELA-GT-2018 Datensatz je Mediennavigator

Mediennavigator	Anzahl Nachrichtenanbieter
SPR	31
AllSides	75
Ad Fontes Media	38

Tabelle 6: Anzahl der Nachrichtenanbieter im Kaggle Datensatz je Mediennavigator

Mediennavigator	Anzahl Nachrichtenanbieter
SPR	10
AllSides	15
Ad Fontes Media	14

Betrachtet man die Verteilung der politischen Klassen des NELA-GT-2018 Datensatzes anhand des Mediennavigators von SPR, so stellt man fest, dass etwas weniger als 540000 Nachrichten keiner der drei politischen Klassen zugeordnet werden konnte (siehe Tabelle 7). Ungefähr jede zweite Nachricht, dessen politische Klasse bekannt ist, stammt von einem konservativem Nachrichtenanbieter. Der Anteil der liberalen Nachrichten ist mit 24% genauso hoch wie der Anteil der Nachrichten, die von der politischen Mitte stammen (siehe Abbildung 6). Bei der Einteilung der Nachrichten in fünf politische Ausprägungen mithilfe des Mediennavigators von AllSides, beträgt die Anzahl der Nachrichten, die keiner politischen Klassen zugewiesen konnte, etwas mehr als 392000 (siehe Tabelle 8). Die Verteilung der Nachrichten im NELA-GT-2018 Datensatzes ist fast ausgeglichen, wenn man vernachlässigt, dass ungefähr ein Viertel der Nachrichten von rechten Nachrichtenanbietern stammen (siehe Abbildung 7). Die Anzahl der Nachrichten im NELA-GT-2018 Datensatz, die keiner der sieben politischen Ausprägungen des Mediennavigators von Ad Fontes Media zugeordnet werden konnte, liegt ungefähr bei 550000 (siehe Tabelle 9). Obwohl es sieben politische Klassen gibt, findet sich im NELA-GT-2018 Datensatz kein einzelner Nachrichtenanbieter wieder, die sich der politische Ausrichtung *Most Extreme Left* zuordnen ließ. Außerdem sind nur 2518 Nachrichten der gegenüberliegenden Klasse *Most Extreme Right* zugewiesen. Weniger als die Hälfte der Nachrichten stammen von neutralen Nachrichtenanbietern

Vorhersage der politischen Ausrichtung von Nachrichtenartikeln mittels Machine Learning

und es befinden sich mehr Nachrichten aus dem linken politischen Spektrum als aus dem rechten politischen Spektrum (siehe Abbildung 8).

Tabelle 7: Verteilung der Klassen im NELA-GT-2018 Datensatz anhand des Mediennavigators von SPR

Klasse	Anzahl Nachrichten
Liberal	41889
Center	41413
Conservative	89971
Keine Klassen	539264

Tabelle 8: Verteilung der Klassen im NELA-GT-2018 Datensatz anhand des Mediennavigators von AllSides

Klasse	Anzahl Nachrichten
Left	52378
Lean Left	68449
Center	53347
Lean Right	64325
Right	82021
Keine Klassen	392017

Tabelle 9: Verteilung der Klassen im NELA-GT-2018 Datensatz anhand des Mediennavigators von Ad Fontes Media

Klasse	Anzahl Nachrichten
Most Extreme Left	0
Hyper-Partisan Left	28505
Skews Left	21870
Neutral	79753
Skews Right	21870
Hyper-Partisan Right	6993
Most Extreme Right	2518
Keine Klassen	549491

Vorhersage der politischen Ausrichtung von Nachrichtenartikeln mittels Machine Learning

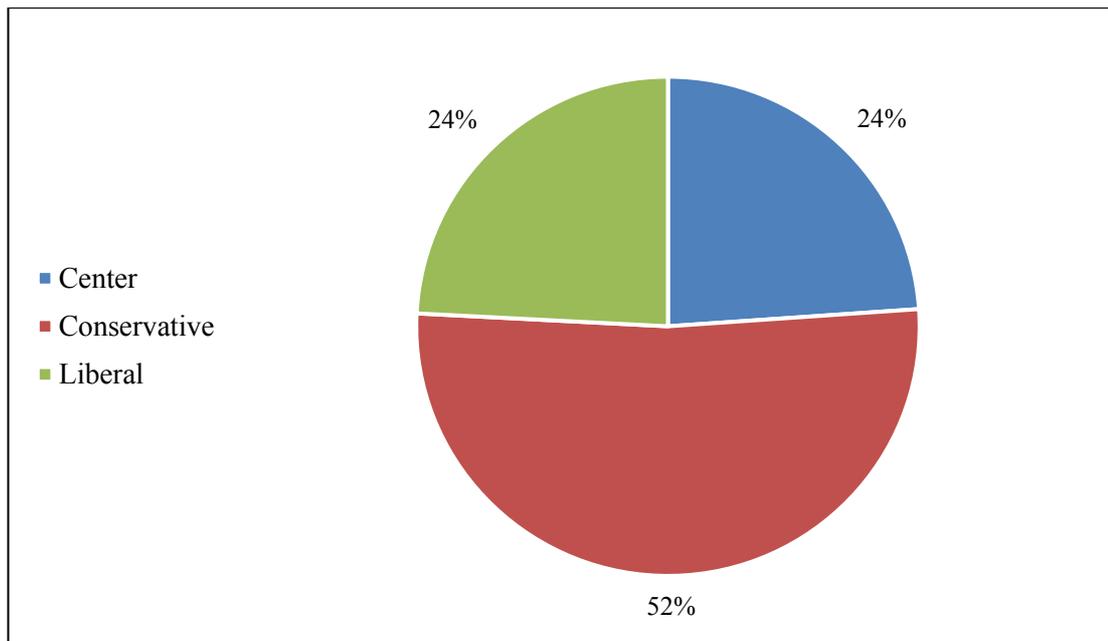


Abbildung 6: Verteilung der Klassen im NELA-GT-2018 Datensatz anhand des Mediennavigators von SPR ohne Berücksichtigung der Nachrichten, die keiner politischen Klasse zugewiesen sind

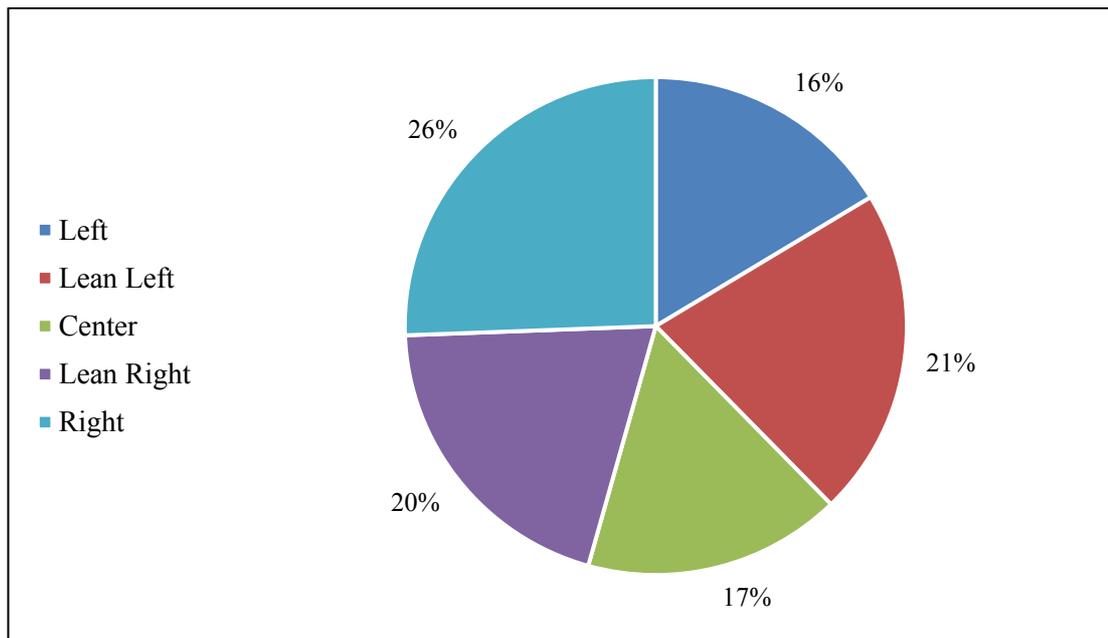


Abbildung 7: Verteilung der Klassen im NELA-GT-2018 Datensatz anhand des Mediennavigators von AllSides ohne Berücksichtigung der Nachrichten, die keiner politischen Klasse zugewiesen sind

Vorhersage der politischen Ausrichtung von Nachrichtenartikeln mittels Machine Learning

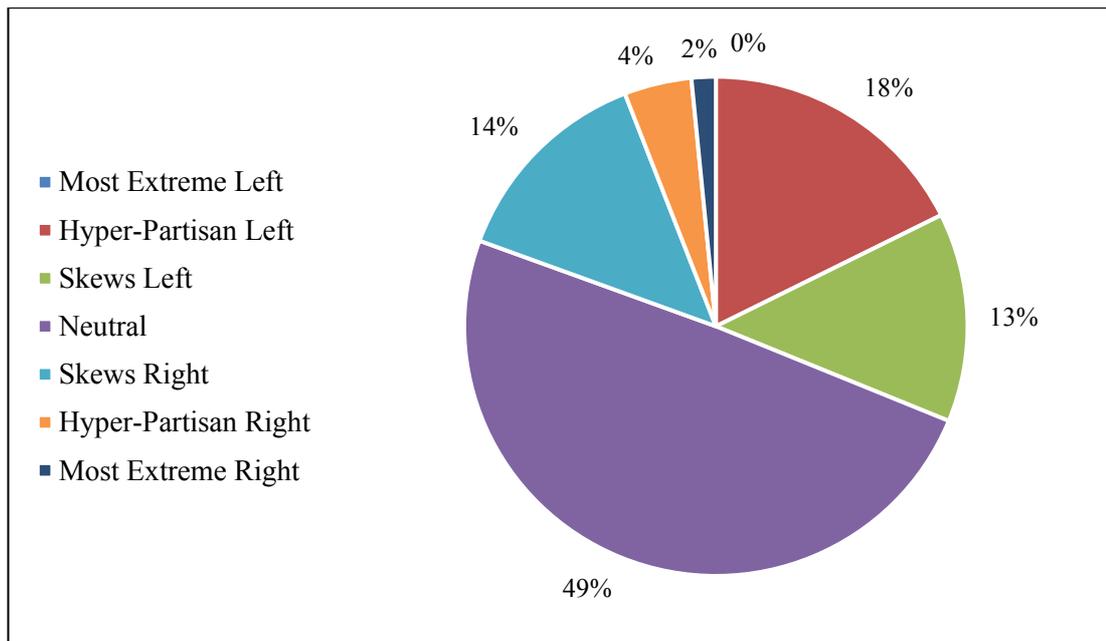


Abbildung 8: Verteilung der Klassen im NELA-GT-2018 Datensatz anhand des Mediennavigators von Ad Fontes Media ohne Berücksichtigung der Nachrichten, die keiner politischen Klasse zugewiesen sind

50248 Nachrichten im Kaggle Datensatz konnten keiner politische Klasse des Mediennavigators von SPR zugewiesen werden (siehe Tabelle 10). Unter den Nachrichten, denen eine politische Richtung zugewiesen werden konnte, ist die politische Richtung fast ausgeglichen verteilt (siehe Abbildung 9). Nur 17621 Nachrichten konnten keiner politischen Ausrichtung des Mediennavigators von AllSides zugewiesen werden (siehe Tabelle 11). Fast die Hälfte der Nachrichten stammt aus dem halb linken politischen Spektrum, während die Klassen *Left* und *Lean Right* gemeinsam nur 6% aller Nachrichten ausmachen, denen eine politische Richtung zugewiesen worden ist (siehe Abbildung 10). Die Anzahl der Nachrichten, die keiner der sieben politischen Richtungen des Mediennavigators von Ad Fontes Media zugeordnet werden konnten, liegt bei 22208 Nachrichten (siehe Tabelle 12). Jedoch sind die beiden Klassen *Most Extreme Left* und *Most Extreme Right* überhaupt nicht im Datensatz vertreten. Weniger als zwei Drittel der Nachrichten stammen aus dem neutralen politischen Spektrum, während etwas mehr als ein Drittel der Nachrichten aus dem halb linken, halb rechten und dem rechten Spektrum kommen (siehe Abbildung 11).

Tabelle 10: Verteilung der Klassen im Kaggle Datensatz anhand des Mediennavigators von SPR

Klasse	Anzahl Nachrichten
Liberal	44824

Vorhersage der politischen Ausrichtung von Nachrichtenartikeln mittels Machine Learning

Center	42142
Conservative	51514
Keine Klassen	50248

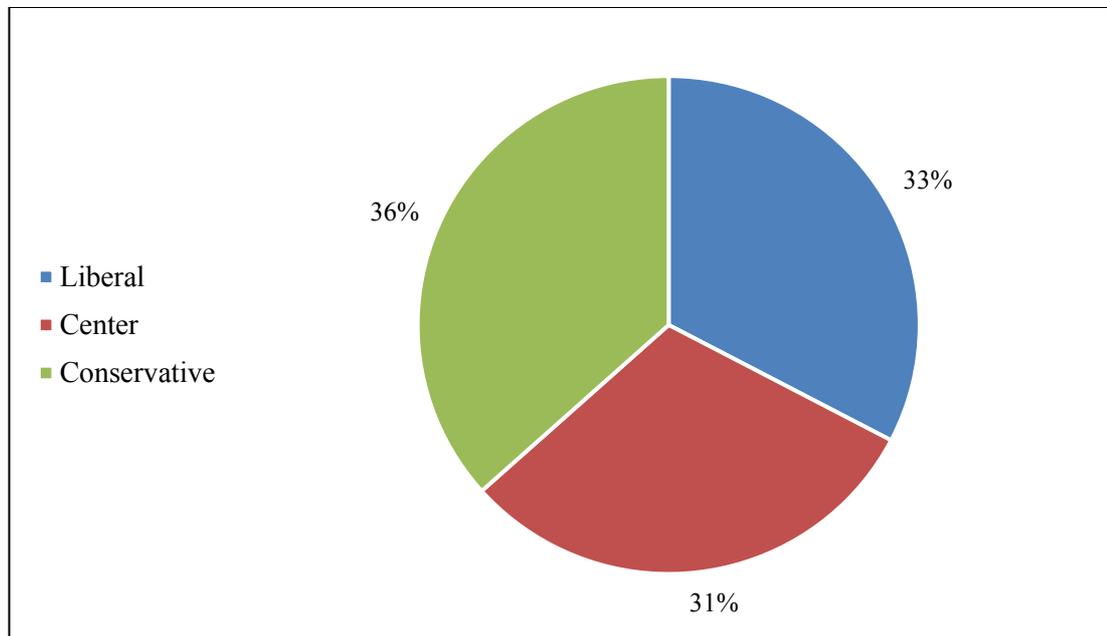


Abbildung 9: Verteilung der Klassen im Kaggle Datensatz anhand des Mediennavigators von SPR ohne Berücksichtigung der Nachrichten, die keiner politischen Klasse zugewiesen sind

Tabelle 11: Verteilung der Klassen im Kaggle Datensatz anhand des Mediennavigators von AllSides

Klasse	Anzahl Nachrichten
Left	4953
Lean Left	83576
Center	31064
Lean Right	4384
Right	47130
Keine Klassen	17621

Tabelle 12: Verteilung der Klassen im Kaggle Datensatz anhand des Mediennavigators von Ad Fontes Media

Klasse	Anzahl Nachrichten
Most Extreme Left	0
Hyper-Partisan Left	5323
Skews Left	23038
Neutral	103726
Skews Right	10642
Hyper-Partisan Right	23791

Vorhersage der politischen Ausrichtung von Nachrichtenartikeln mittels Machine Learning

Most Extreme Right	0
Keine Klassen	22208

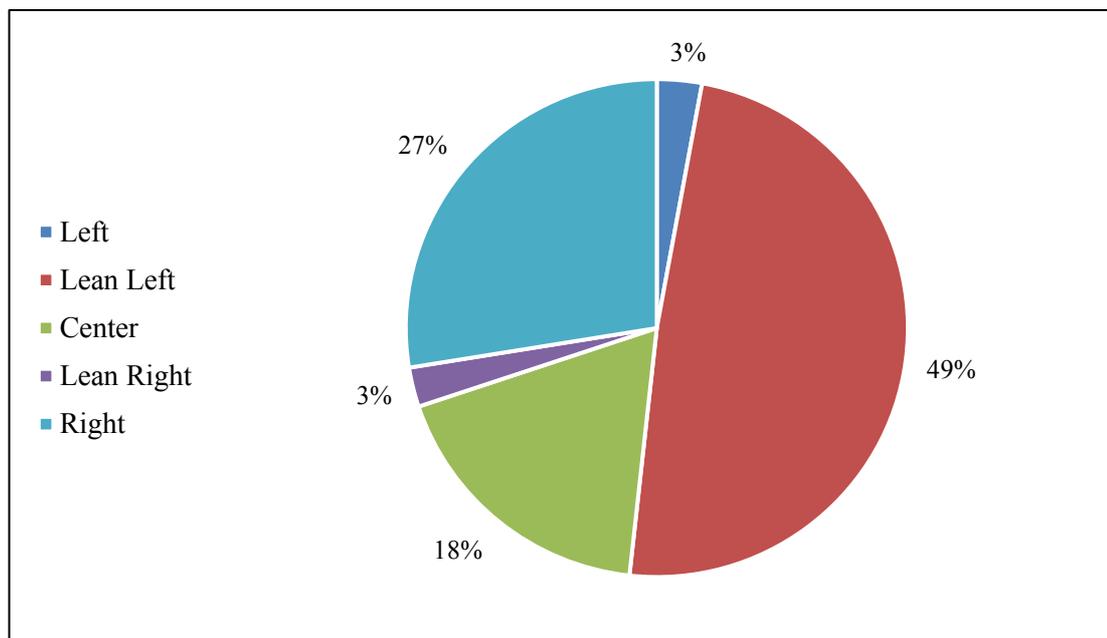


Abbildung 10: Verteilung der Klassen im Kaggle Datensatz anhand des Mediennavigators von AllSides ohne Berücksichtigung der Nachrichten, die keiner politischen Klasse zugewiesen sind

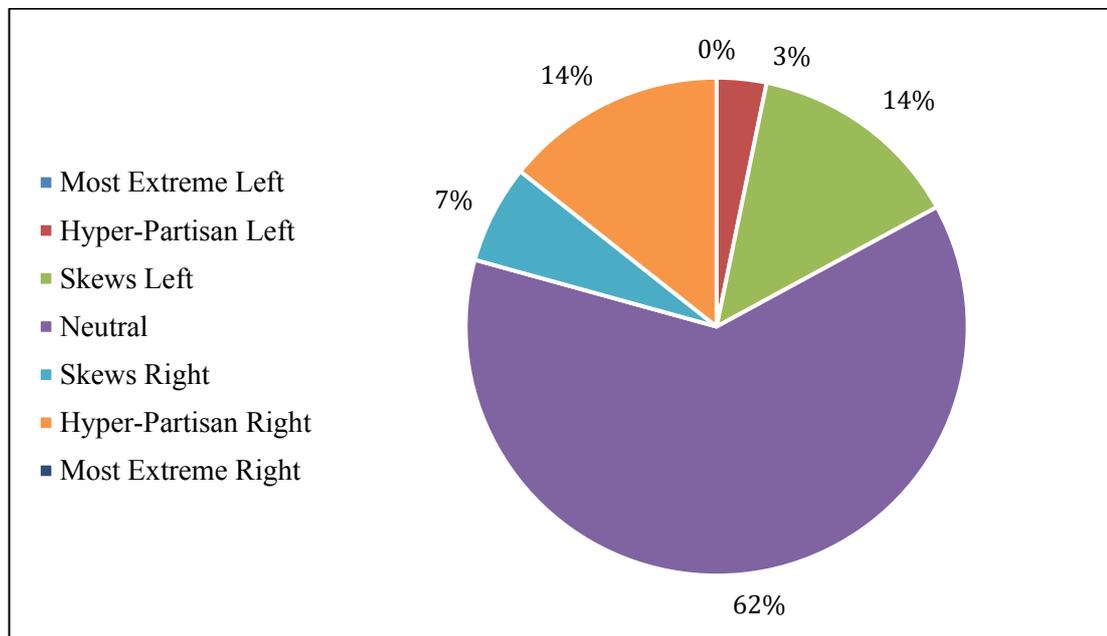


Abbildung 11: Verteilung der Klassen im Kaggle Datensatz anhand des Mediennavigators von Ad Fontes Media ohne Berücksichtigung der Nachrichten, die keiner politischen Klasse zugewiesen sind

Unabhängig vom Datensatz lässt sich feststellen, dass die Anzahl der Nachrichten, die keiner politischen Richtung zugewiesen worden sind, am geringsten ist, wenn der

Mediennavigator von AllSides zugrunde gelegt wird (siehe Abbildung 12). Außerdem deckt der Mediennavigator von AllSides sowohl absolut als auch relativ betrachtet die meisten Nachrichtenanbieter im NELA-GT-2018 und Kaggle Datensatz ab (siehe Abbildung 13, Tabelle 5 und Tabelle 6). Im NELA-GT-2018 Datensatz konnten Nachrichtenanbietern eine politische Ausrichtung zugewiesen werden, auch wenn diese nicht in Abbildung 4 mitaufgeführt werden. Daher deckt der Mediennavigator von AllSides im NELA-GT-2018 Datensatz 144% der Nachrichtenanbieter ab. Ein negativer Aspekt des Mediennavigators von Ad Fontes Media ist, dass in keinem der beiden Datensätze alle sieben politische Ausrichtungen enthalten sind (siehe Abbildung 8 und Abbildung 11). Man kann anmerken, dass bei einer anderen Datengrundlage eine ausgewogenere Klassenverteilung zu Stande kommen kann, aber ebenso gut muss man die Tatsache hervorheben, dass nur insgesamt zwei der 90 Nachrichtenanbieter zu den politischen Richtungen *Most Extreme Left* und *Most Extreme Right* zugewiesen worden sind. Obwohl der Mediennavigator von SPR mit den drei politischen Ausrichtungen im NELA-GT-2018 die ausgeglichene Klassenverteilung begünstigt, so deckt er auch über beide Datensätze hinweg absolut betrachtet die wenigsten Nachrichten ab (siehe Tabelle 5 und Tabelle 6).

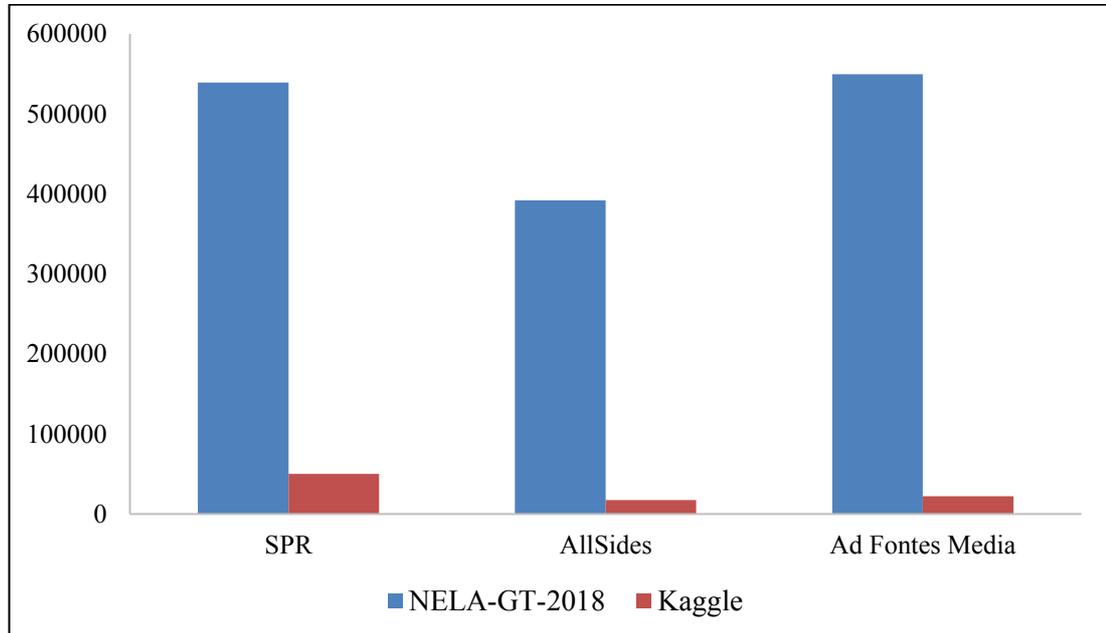


Abbildung 12: Anzahl der Nachrichten je Datensatz, die keiner politischen Klassen zugeordnet wurde

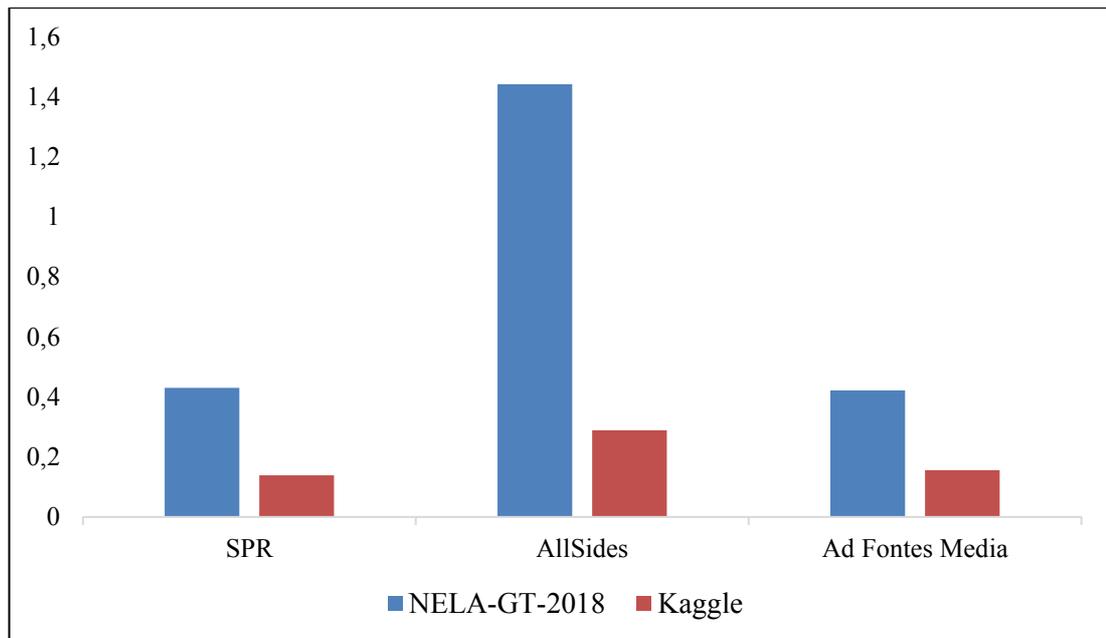


Abbildung 13: Anteil der in den Mediennavigatoren abgedeckten Nachrichtenquellen durch die Datensätze

In der Implementierung wird daher der Mediennavigator von AllSides verwendet, um den Nachrichten eine politische Klasse zuzuweisen. Mit diesem Mediennavigator, welcher auch schon von Kulkarni et al. (2018) verwendet worden ist, werden die meisten Nachrichtenanbieter in den beiden Datensätzen abgedeckt. In einem weiteren Schritt werden der NELA-GT-2018 und der Kaggle Datensatz zusammengeführt und die Klassenverteilung anhand des Mediennavigators von AllSides analysiert. In dem zusammengeführten Datensatz werden nur Nachrichtenanbieter berücksichtigt, denen eine politische Richtung zugewiesen worden ist. Falls für Nachrichtenanbieter keine politische Ausrichtung anhand des Mediennavigators von AllSides zugeordnet werden kann, jedoch von den Mediennavigatoren von SPR und Ad Fontes Media, so wird deren Zuweisung übernommen und die politische Ausrichtung wird in das Namensschema von AllSides angepasst. Die Konvertierung der politischen Richtungen im Mediennavigator von SPR wird wie folgt vorgenommen:

- *Liberal* → *Left*
- *Center* → *Center*
- *Conservative* → *Right*

Die Anpassung der politischen Richtungen im Mediennavigator von Ad Fontes Media an das Namensschemata des von AllSides veröffentlichten Mediennavigator sieht wie folgt aus:

- *Most Extreme Left* → *Left*
- *Hyper-Partisan Left* → *Left*
- *Skews Left* → *Lean Left*
- *Neutral* → *Center*
- *Skews Right* → *Lean Right*
- *Hyper-Partisan Right* → *Right*
- *Most Extreme Right* → *Right*

In dem zusammengeführten Datensatz befinden sich insgesamt 535576 Nachrichten aus über 85 verschiedenen Nachrichten Anbietern (siehe Tabelle 13). Mit etwas mehr als 152000 Nachrichten stellen Nachrichten aus dem halb linken politischen Spektrum die am meisten verbreitete politische Richtung dar, während ein Viertel der Nachrichten von rechten Nachrichtenanbieter die am zweitmeisten verbreitete politische Ausrichtung ist (siehe Tabelle 14 und Abbildung 14). Die übrigen politischen Richtungen sind mit 15% bzw. 16% gleich häufig im Datensatz vertreten.

Dieser zusammengeführte Datensatz bildet eine gute Grundlage für die Implementierung der Modelle, sodass es keine Notwendigkeit besteht weitere Nachrichten aus dem Internet zu *scrapen*.

Tabelle 13: Anzahl der Nachrichten je Nachrichtenanbieter und deren politische Richtung im zusammengeführten Datensatz

Quelle	NELA-GT-2018	Kaggle	Gesamt	Label
21stCenturyWire	322	0	322	Left
ABC News	2808	0	2808	Lean Left
Activist Post	1797	0	1797	Right
Al Jazeera	4517	0	4517	Center
Alternet	4816	0	4816	Left
BBC	16409	0	16409	Center
Bipartisan Report	4060	0	4060	Left
Breitbart	1877	23791	25668	Right
Business Insider	445	8344	8789	Center
Buzzfeed	1661	10699	12360	Lean Left
CBS News	5396	0	5396	Lean Left
Chicago Sun-Times	2113	0	2113	Lean Left
CNBC	2426	0	2426	Center
CNN	8202	14288	22490	Lean Left

Vorhersage der politischen Ausrichtung von Nachrichtenartikeln mittels Machine Learning

CNS News	5263	0	5263	Right
Counter Current News	23	0	23	Lean Left
Daily Beast	6634	0	6634	Left
Daily Kos	994	0	994	Left
Daily Mail	3596	0	3596	Right
Daily Signal	308	0	308	Right
Dick Morris Blog	400	0	400	Lean Right
Drudge Report	18884	0	18884	Lean Right
FiveThirtyEight	556	0	556	Center
Foreign Policy	701	0	701	Center
Fox News	3106	4384	7490	Lean Right
FrontPage Magazine	892	0	892	Right
Hot Air	4642	0	4642	Lean Right
Infowars	2518	0	2518	Right
Intellectual Conservative	377	0	377	Lean Right
Investors Business Daily	730	0	730	Lean Right
LewRockwell	1275	0	1275	Right
Live Action	1054	0	1054	Lean Right
Media Matters for America	2316	0	2316	Left
Mint Press News	1707	0	1707	Left
MotherJones	1128	0	1128	Left
MSNBC	6604	0	6604	Left
National Review	5129	6258	11387	Right
New York Daily News	2042	0	2042	Left
New York Post	25401	18144	43545	Right
New Yorker	265	0	265	Left
Newsweek	9411	0	9411	Lean Left
NPR	5515	12001	17516	Center
oann	14267	0	14267	Lean Right
Observer	541	0	541	Center
Palmer Report	3539	0	3539	Left
PBS	1113	0	1113	Center
Politico	628	0	628	Lean Left
Politics UK	137	0	137	Lean Left
Politicus USA	4018	0	4018	Left
Raw Story	3719	0	3719	Left
Real Clear Politics	7206	0	7206	Center
RedState	4808	0	4808	Right
Reuters	3929	10719	14648	Center
RightWingWatch	1118	0	1118	Right
Salon	1702	0	1702	Left
Shareblue	2134	0	2134	Left
Slate	514	0	514	Left
Talking Points Memo	5841	5323	11164	Left

Vorhersage der politischen Ausrichtung von Nachrichtenartikeln mittels Machine Learning

The American Conservative	439	0	439	Lean Right
The Atlantic	1757	7386	9143	Lean Left
The Daily Caller	11549	0	11549	Right
The Fiscal Times	461	0	461	Lean Right
The Gateway Pundit	5667	0	5667	Right
The Guardian	2195	8685	10880	Lean Left
The Hill	1968	0	1968	Center
The Huffington Post	5586	0	5586	Left
The Independent	19799	0	19799	Lean Left
The Intercept	1268	0	1268	Left
The Irish Times	3826	0	3826	Right
The Michelle Malkin Blog	53	0	53	Right
The New York Times	5471	30141	35612	Lean Left
The Onion	1094	0	1094	Lean Left
The Telegraph	33763	0	33763	Lean Right
The Verge	5951	1225	7176	Lean Left
The Washington Examiner	469	0	469	Lean Right
The Week UK	2207	0	2207	Center
TheBlaze	5287	0	5287	Right
ThinkProgress	4819	0	4819	Left
USA Today	5814	0	5814	Center
Veterans Today	2621	0	2621	Right
Vox	4288	4953	9241	Left
Washington Monthly	551	0	551	Lean Left
Washington Post	1252	11152	12404	Lean Left
Western Journal	4729	0	4729	Right
Yahoo News	1665	0	1665	Left

Tabelle 14: Verteilung der Klassen im zusammengeführten Datensatz

Klasse	Anzahl Nachrichten
Left	80257
Lean Left	152025
Center	84411
Lean Right	82976
Right	135907

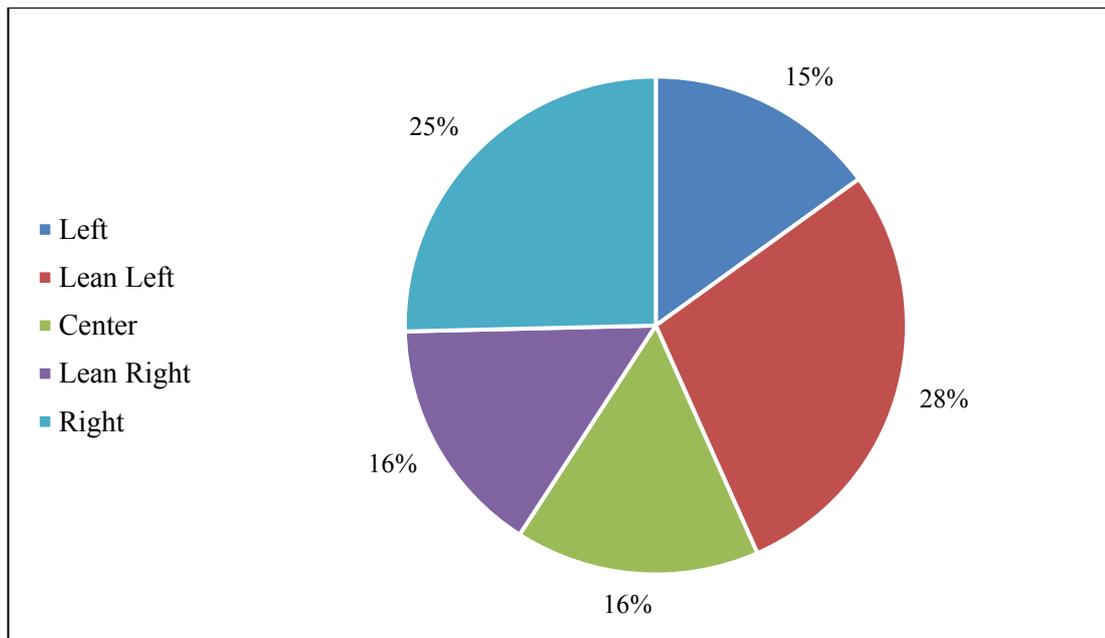


Abbildung 14: Verteilung der Klassen im zusammengeführten Datensatz

3.3 Implementierung

Die Implementierung erfolgt in der Programmiersprache *Python* und hat auf einem vom Lehrstuhl gehosteten Server stattgefunden. Dieser Server ist mit einer NVIDIA Titan RTX mit 24 Gigabyte VRAM ausgestattet. Für das Trainieren der Modelle werden die Bibliotheken *scikit-learn*²¹, *Keras*²² (mit *Tensorflow* als Backend) und *ktrain*²³ (Maiya, 2020) benutzt. In den Unterkapiteln werden jeweils auf Pythonskripte verwiesen, welches unter folgendem *Github Repository*²⁴ aufrufbar sind.

Insgesamt werden 27 verschiedene Modelle implementiert, um die fünf möglichen politischen Ausrichtungen von Nachrichten vorherzusagen (*Full Experiment*). Angelehnt an die Arbeit von Baly et al. (2019) werden neben dem *Full Experiment* zusätzlich zwei weitere *Ablation Studies* (*Ablation Study 01* und *Ablation Study 02*) durchgeführt. In den *Ablation Studies* wird überprüft, ob die Modelle eine bessere *Accuracy* erzielen, wenn halb linke und halb rechte Nachrichten eher dem politischen Spektrum in der Mitte oder an den äußeren Rändern zugeordnet werden. In beiden *Ablation Studies* werden daher drei statt fünf politische Richtungen klassifiziert. In der *Ablation Study 01* werden die politischen Richtungen wie folgt zusammengefasst:

²¹ <https://scikit-learn.org/stable/index.html>

²² <https://keras.io/>

²³ <https://github.com/amaiya/ktrain>

²⁴ <https://github.com/oakyuez/master-thesis>

- $\{Left, Lean Left\} \rightarrow Left$
- $Center \rightarrow Center$
- $\{Lean Right, Right\} \rightarrow Right$

In der zweiten *Ablation Study 02* werden Nachrichten aus dem halb linken und halb rechten Spektrum zu der politischen Mitte zugewiesen. Die Zuordnung sieht wie folgt aus:

- $Left \rightarrow Left$
- $\{Lean Left, Center, Lean Right\} \rightarrow Center$
- $Right \rightarrow Right$

In den *Ablation Studies* wird je Bibliothek ein Modell angewandt, und zwar das Modell, welches im *Full Experiment* am besten abschneidet.

Um die Trainingszeiten der Modelle zu verkürzen, werden nur 75000 der 535576 Nachrichten berücksichtigt. Außerdem sind in den 75000 Nachrichten alle politischen Richtungen gleich häufig vertreten. Im *Full Experiment* stellt jede politische Richtung daher 20% (entspricht 15000) der Nachrichten dar, während in den *Ablation Studies* aus jeder politischen Richtung 25000 (entspricht ein Drittel) Nachrichten enthalten sind. Der Anteil der Testdaten beträgt 20%. Da *Transformer* Modelle Nachrichten mit mehr 512 Wörtern nicht bearbeiten können, enthält jede Nachricht im Trainings- und Testdatensatz weniger als 512 Wörtern.

Wie in der Abbildung 2 dargestellt worden ist, besteht die Implementierung aus drei einzelnen Komponenten, die in den folgenden Unterkapiteln näher erläutert werden.

3.3.1 Target Labeling

Während der Machbarkeitsstudie ist die csv-Datei *labels.csv* entstanden, in der die zugewiesene politische Richtung der Nachrichtenanbieter enthalten ist. Sowohl der NELA-GT-2018 als auch der Kaggle Datensatz kann ausschließlich als Datenbankdatei heruntergeladen werden. Beide Datensätze werden als *pandas*²⁵ *Dataframe* eingelesen und zu einem *Dataframe* zusammengeführt. Abschließend wird dieser *Dataframe* mit der csv-Datei *labels.csv* anhand der Spalte *Source* über ein *Join*

²⁵ <https://pandas.pydata.org/>

zu einem neuem *Dataframe* verschmolzen. Dieser *Dataframe* wird als csv-Datei *nela_kaggle.csv* abgespeichert und für das *Preprocessing* und die *Vectorization* wieder benutzt. Das Skript zum Target Labeling heißt *03_join.py*.

3.3.2 Preprocessing & Vectorization

Für das Anwenden von Modellen, die auf dem *Bag-of-Words* Ansatz basieren, ist es notwendig, dass die Nachrichten, die in Textform vorliegen, nach gewissen Kriterien bereinigt (engl. *preprocessed*) und zu *Tokens* umgewandelt werden. Bei der Bereinigung des reinen Textes werden auf die Methoden der *nlTK*²⁶ Bibliothek zurückgegriffen. Aus dem reinen Nachrichtentext werden zuerst die Interpunktionen entfernt. Anschließend werden die Wörter in der Nachricht als einzelne *Tokens* aufgefasst und alle Großbuchstaben zu Kleinbuchstaben umgewandelt. Als letzter Schritt in der Bereinigung erfolgt das *Stemming* nach Porter (1980). Alle hier aufgelisteten Bereinigungs-schritte werden gemeinsam in der Methode *preprocess_text_with_porter* durchgeführt. Als Argumente werden der Methode das *Dataframe*, die Spalte, in der sich der reine Nachrichtentext befindet und die neue Spalte, die die *Tokens* (als Resultats des *Preprocessing*) beinhalten soll, mitgegeben. Das Skript zum *Preprocessing* heißt *04_text_relabeling_preprocessing.py*. In diesem Skript werden auch die neuen politischen Richtungen für die beiden *Ablation Studies* erstellt. Das *Dataframe* mit den bereinigten Nachrichten wird als csv-Datei *news_preprocessed_with_additionalLabel.csv* abgespeichert.

Die *Vectorization* der *Tokens* erfolgt mit dem *TfidfVectorizer* der *scikit-learn* Bibliothek. Bei der *Vectorization* werden aus den *Tokens* je zwei Dokumentenmatrizen erzeugt, die beide jeweils die 30000 häufigsten *Tokens* der Nachrichtenartikeln beinhalten. Die Dokumentenmatrizen unterscheiden sich in der Berücksichtigung von *Bi-* und *Trigrams*. Der *TfidfVectorizer* übernimmt auch das Entfernen von *stopwords*, sodass dieser Schritt nicht in der Methode *preprocess_text_with_porter* ausgeführt wird. Die *Vectorization* erfolgt ausschließlich bei der Anwendung der Modelle, die auf dem *Bag-of-Words* Ansatz basieren, und wird daher nur in den Skripten *06_FE_scikitlearn_tfidf.py*, *06_AbSt01_svm_scikit-learn.py* und *06_AbSt02_svm_scikit-learn.py* berücksichtigt.

²⁶ <https://www.nltk.org/>

3.3.3 Modellanwendung & -bewertung

3.3.3.1 *Bag-of-Words* Modelle mit *Scikit-learn*

Aus der *scikit-learn* Bibliothek werden je zwei Entscheidungsbäume, *SVM*, logistische Regressionen und Naive Bayes Klassifikatoren angewandt. Jedes Modell wird je einmal mit den in 3.3.2 erzeugten *Bi-* und *Trigrams* trainiert. Um die Ergebnisse replizieren zu können, wird der *random_state* Parameter beim Initialisieren der Entscheidungsbäume auf 42 gesetzt. Bei der Initialisierung der logistischen Regression werden die Parameter *solver* und *multi_class* auf *saga* bzw. *multinomial* gesetzt.

Alle Modelle werden beim *Full Experiment* in dem Skript *06_FE_scikit_learn_tfidf.py* angewandt. In den *Ablation Studies* wird die *SVM* mit *Trigrams* als Trainingsattributen in den Skripten *06_AbSt01_svm_scikit-learn.py* und *06_AbSt02_svm_scikit-learn.py* angewandt.

3.3.3.2 Neuronale Netze mit *Keras*

Mithilfe der *Keras* Bibliothek werden acht verschiedene künstliche neuronale Netze implementiert. Das Beispiel, wie sie in der *Keras* Dokumentation zu finden ist (Omernick & Chollet, 2019), ist für den Anwendungsfall dieser Masterarbeit angepasst worden. Die Parameter *max_features* und *sequence_length* werden von 20000 auf 30000 bzw. von 500 auf 512 erhöht, während die Dimension der *Embedding* Schicht unverändert bleibt. Die letzte Schicht des neuronalen Netzes enthält fünf statt nur einen Ausgabeknoten. Die weiteren implementierten neuronale Netze sind

- ein *Feed-Forward Network* mit drei Schichten,
- ein *CNN*,
- ein *GRU*,
- ein bidirektionaler *GRU*,
- ein *LSTM*,
- ein bidirektionaler *LSTM*,
- sowie ein bidirektionaler *LSTM* mit zwei *Convolutional* Schichten nach der *Embedding* Schicht.

Alle implementierten neuronalen Netze weisen die gleiche Spezifikation in den Parametern (*max_features*, *sequence_length*, *embedding_dim*) und der Architektur der letzten Schicht mit fünf Ausgabeknoten für die einzelnen politischen Richtungen auf. Jedes künstliche neuronale Netz wird maximal 10 *Epochs* lang trainiert. Das Training wird vorzeitig gestoppt, wenn der Wert des zu minierenden *val_loss* nach zwei *Epochs* wieder ansteigt. Im *Full Experiment* werden die neuronalen Netze im Skript *07_FE_kerasTF.py* angewandt.

In den *Ablation Studies* wird das künstliche neuronale Netz aus der *Keras* Dokumentation angewandt und enthält in der letzten Schicht drei Ausgabeknoten (Omernick & Chollet, 2019). Die Implementierung der beiden *Ablation Studies* findet in den Skripten *07_AbSt01_keras_NeuralNetwork.py* und *07_AbSt02_keras_NeuralNetwork.py* statt.

3.3.3.3 *Transformer* und vordefinierte Modelle mit *ktrain*

Die meisten Modelle (insgesamt 11) sind mit der *ktrain* Bibliothek implementiert, da sie neben vordefinierten Modelle für Textklassifikationen auf die Modelle der *transformers*²⁷ Bibliothek zurückgreifen kann (Maiya, 2020). Neben den vordefinierten Modellen *fasttext* (Joulin et al., 2016), *logreg*, *nbsvm* (Wang & Manning, 2012), *bigru* und *standard_gru* werden die folgenden Modellen aus der *transformers* Bibliothek angewandt:

- *BERT* (Devlin et al., 2019)
- *distilBERT* (Sanh et al., 2020)
- *RoBERTa* (Liu et al., 2019)
- *ALBERT* (Lan et al., 2020)
- *XLNET* (Yang et al., 2020)
- *XLM* (Lample & Conneau, 2019)

Der Klassifikator *logreg* ist eine logistische Regression, welcher auf einer trainierbaren *Embedding* Schicht basiert (Maiya, o. J.). Die Modelle *bigru* und *standard_gru* unterscheiden sich von den in 3.3.3.2 vorgestelltem *GRU* hinsichtlich der zugrunde gelegten *Word Embeddings*. Vor dem Trainieren aller elf Modelle ist das

²⁷ <https://huggingface.co/transformers/#>

Ermitteln der optimalen Lernrate mit der Methode *lr_find* essenziell, da sie im Anschluss an die Methode *lr_fit* übergeben wird. Die Höhe der übergebenen *Epochs* in den Methoden *lr_find* und *lr_fit* ist modellabhängig und ist aufgrund der langen Ausführungszeit erst während der Implementierung der jeweiligen Modelle festgelegt worden.

Die Modelle *fasttext*, *logreg*, *nbsvm*, *bigru* und *stgru* werden beim *Full Experiment* im Skript *08_FE_ktrain_classifiers.py* implementiert. Die übrigen Modelle werden in den Skripten *09_FE_huggingface_distilBERT.py*, *10_FE_huggingface_BERT.py*, *11_FE_huggingface_ALBERT.py*, *12_FE_huggingface_RoBERTa.py*, *13_FE_huggingface_XLM.py* und *14_FE_huggingface_XLNet.py* angewandt.

Für die *Ablation Studies* werden die Modelle *bigru* und *RoBERTa* implementiert. Die dazugehörigen Pythonskripte lauten *08_AbSt01_biGRU.py*, *08_AbSt02_biGRU.py*, *09_AbSt01_RoBERTa.py* und *09_AbSt02_RoBERTa.py*.

Eine Übersicht aller angewandten Modelle ist in der untenstehenden Tabelle 15 entnehmbar.

Tabelle 15: Übersicht der angewandten Modelle

Modell Nr.	Klassifikator	Spezifikation/Erläuterung	Berücksichtigt in der Ablation Study
Mod_01	Entscheidungsbaum	Mit Bigrams trainiert	Nein
Mod_02	Entscheidungsbaum	Mit Trigrams trainiert	Nein
Mod_03	Naive Bayes	Mit Bigrams trainiert	Nein
Mod_04	Naive Bayes	Mit Trigrams trainiert	Nein
Mod_05	SVM	Mit Bigrams trainiert	Nein
Mod_06	SVM	Mit Trigrams trainiert	Ja
Mod_07	Logistische Regression	Mit Bigrams trainiert	Nein
Mod_08	Logistische Regression	Mit Trigrams trainiert	Nein
Mod_09	CNN + Dense layer	Beispiel aus der Keras Dokumentation (Omernick & Chollet, 2019)	Ja

Vorhersage der politischen Ausrichtung von Nachrichtenartikeln mittels Machine Learning

Mod_10	Feed-Forward Network	3 <i>hidden Dense</i> Schichten	Nein
Mod_11	CNN	3 <i>hidden Convolutional</i> Schichten	Nein
Mod_12	GRU	1 <i>hidden</i> Schicht mit GRU und 2 <i>hidden Dense</i> Schichten	Nein
Mod_13	Bidirektionaler GRU	3 <i>hidden</i> Schicht mit GRU	Nein
Mod_14	LSTM	1 <i>hidden</i> Schicht mit LSTM Zellen und 2 <i>hidden Dense</i> Schichten	Nein
Mod_15	Bidirektionaler LSTM	3 <i>hidden</i> Schicht mit LSTM Zellen	Nein
Mod_16	CNN + Bidirektionaler LSTM	2 <i>hidden Convolutional</i> Schichten + 2 <i>hidden</i> Schicht mit LSTM Zellen + 1 <i>hidden Dense</i> Schicht	Nein
Mod_17	fasttext	Wird 20 Epochs trainiert	Nein
Mod_18	logreg	Wird 10 Epochs trainiert	Nein
Mod_19	nbsvm	Wird 10 Epochs trainiert	Nein
Mod_20	bigru	Wird 7 Epochs trainiert	Ja
Mod_21	standard_gru	Wird 7 Epochs trainiert	Nein
Mod_22	BERT	Wird ein Epoch trainiert	Nein
Mod_23	distilBERT	Wird 5 Epoch trainiert	Nein
Mod_24	ALBERT	Wird ein Epoch trainiert	Nein
Mod_25	RoBERTa	Wird ein Epoch trainiert	Ja
Mod_26	XLM	Wird ein Epoch trainiert	Nein
Mod_27	XLNet	Wird ein Epoch trainiert	Nein

Bemerkung: Die fett gedruckten Modelle werden sowohl im *Full Experiment* als auch in den *Ablation Studies* angewandt.

Die Modellgüte aller Modelle wird mit der *Accuracy* bewertet. Die *Accuracy* gibt in Klassifikationsproblemen den Anteil aller richtig klassifizierten Instanzen unter allen klassifizierten Instanzen wieder (Kowsari et al., 2019). Für alle Modelle, die

mehr als ein *Epoch* trainiert werden, wird jeweils die *Accuracy* angegeben, wo der *val_loss* am geringsten ist. Bei dem besten und schlechtesten Modell wird zusätzlich die Konfusionsmatrix genauer betrachtet.

Im nächsten Kapitel werden die *Accuracy* Werte für die einzelnen Modelle sowohl grafisch als auch tabellarisch wiedergeben. Darüber hinaus wird in dem Kapitel auf die Dauer der Trainingszeit der einzelnen Modelle eingegangen. Das Kapitel wird abgeschlossen mit der Analyse von zwei Konfusionsmatrizen.

4 Ergebnisse

Die *Accuracy* Werte für alle im *Full Experiment* und in den *Ablation Studies* angewandten Modelle sind in Tabelle 16 aufgelistet. In Abbildung 15 sind die *Accuracy* Werte für alle im *Full Experiment* angewandten Modelle dargestellt. Ferner zeigt die Abbildung 15 als horizontale Linie die Güte eines Klassifikators auf, der die politische Ausrichtung zufällig bestimmen würde. Die *Accuracy* jenen Klassifikators beträgt 20%, da der Klassifikator zufällig eine politische Klasse von fünf möglichen Ausprägungen bestimmt.

In dem *Full Experiment* schneidet das *BERT* Modell (Mod_22) mit einer *Accuracy* von ca. 17% am schlechtesten ab, während das *RoBERTa* Modell (Mod_25) mit einer *Accuracy* von 86,77% am besten abschneidet. Die Modelle innerhalb der *scikit-learn* Bibliothek (Mod_01 bis Mod_08) erzielen *Accuracy* Werte zwischen ungefähr 53% und 70%. Die *SVM* (Mod_05 und Mod_06) erzielen unter allen implementierten Klassifikatoren aus der *scikit-learn* Bibliothek die besten Resultate unabhängig davon, ob die Modelle mit *Bi-* oder *Trigrams* trainiert werden. Die Naive Bayes (Mod_03 und Mod_04) Klassifikatoren hingegen weisen die schlechtesten *Accuracy* Werte auf. Anhand der *Accuracy* Werte jener acht Modelle, lässt sich nicht ableiten, dass mit *Bigrams* trainierte Modelle bessere Resultate erzielen als Modelle, die mit *Trigrams* trainiert worden sind. Generell scheint es indifferent zu sein, ob Modelle mit *Bi-* oder *Trigrams* trainiert werden, da der Unterschied von maximal 0,26 Prozentpunkten (bei Mod_01 und Mod_02) nicht stark ins Gewicht fällt.

Unter allen künstlichen neuronalen Netzen, die mit der *Keras* Bibliothek implementiert sind, schneidet der *GRU* (Mod_12) am besten ab. Das zweitbeste Modell (Mod_09) ist das neuronale Netz basierend auf der Architektur von Omernick & Chollet (2019). Beide künstliche neuronalen Netze sind die einzigen Modelle, die

eine *Accuracy* von mindestens 70% aufweisen können. Die Modellgüte ist beim *LSTM* (Mod_14) mit einer *Accuracy* in Höhe von 67,49% am schlechtesten. Jedoch beträgt der Unterschied zwischen dem besten (Mod_09) und dem schlechtesten Modell (Mod_14) ungefähr drei Prozentpunkte. Obwohl die bidirektionale *LSTM* (Mod_15) besser abschneidet als die nicht-bidirektionale *LSTM* (Mod_14), erzielt der bidirektionale *GRU* (Mod_13) eine schlechtere *Accuracy* als der nicht-bidirektionale *GRU* (Mod_12). Es ist daher nicht möglich anzunehmen, dass bidirektionale *RNN* (Mod_13 und Mod_15) generell bessere Resultate erzielen werden als nicht-bidirektionale *RNN* (Mod_12 und Mod_14). Beide *GRU* Modelle (Mod_12 und Mod_13) weisen eine höhere *Accuracy* auf als die beiden *LSTM* Klassifikatoren (Mod_14 und Mod_15). Anhand dieser Resultate lässt sich deduzieren, dass bei der Implementierung von *RNN* die *GRU* gegenüber den *LSTM* favorisiert werden sollen.

Unter den vordefinierten Modellen in der *ktrain* Bibliothek (Mod_17 bis Mod_21) weist nur der *nbsvm* (Mod_19) eine *Accuracy* von weniger als 70% auf. Die besten Modelle innerhalb der vordefinierten Modelle sind die *bigru* (Mod_20) und *standard_gru* (Mod_21) Modelle, wobei sogar ersterer einen *Accuracy* Wert von etwas weniger als 80% erzielt. Beide Modelle (Mod_20 und Mod_21) erzielen höhere *Accuracy* Werte als jedes mit der *Keras* Bibliothek implementierte neuronale Netz (Mod_09 bis Mod_16). Tatsächlich schneidet nur der *nbsvm* als einziges vordefiniertes Modell der *ktrain* Bibliothek schlechter ab als die von Grund auf implementierten künstlichen neuronalen Netzen (Mod_09 bis Mod_16). Daraus lässt sich schließen, dass man bei der Auswahl eines geeigneten Modells für die automatisierte Identifizierung der politischen Richtung in Nachrichten bereits vordefinierte Modelle, wie dem *bigru* oder *standard_gru* Modell, für die Implementierung in Erwägung ziehen sollte.

Die Modelle aus der *transformers* Bibliothek, welche ebenfalls mit der *ktrain* Bibliothek implementiert worden sind, weisen sowohl vielversprechende als auch erwähnenswerte Ergebnisse auf. Das *RoBERTa* Modell (Mod_25) erzielt unter allen Modellen eine *Accuracy* von 86,77%, welche um mindestens sieben Prozentpunkte höher ausfällt als die *Accuracy* des zweitbesten Modells (*bigru*, Mod_20). Die *Accuracy* Werte der *BERT*, *distilBERT*, *ALBERT*, *XLM* und *XLNET* Modelle (Mod_21 bis Mod_24, Mod_26 und Mod_27) sind die schlechtesten *Accuracy* Werte aller Modelle im *Full Experiment*. Es ist seltsam, dass alle Modelle bis auf das *XLM* Modell (Mod_26) eine *Accuracy* von 20% oder weniger aufweisen. Dies bedeutet, dass diese

Vorhersage der politischen Ausrichtung von Nachrichtenartikeln mittels Machine Learning

Modelle (Mod_21 bis Mod_24 und Mod_27) schlechter abschneiden als ein Klassifikator, welches die politische Ausrichtung einer Nachricht zufällig und ohne vorher gelernte Information identifiziert. Es stellt sich die Frage, ob diese Modelle generell schlechte Resultate erzielen, oder, ob die Implementierung in der *ktrain* Bibliothek fehlerhaft ist. Letzteres scheint plausibler zu sein, da das *BERT* Modell in der Arbeit von Chun et al. (2019) gegenüber einem *CNN* oder einer *SVM* weitaus überlegener ist.

Tabelle 16: Accuracy Werte der Modelle im *Full Experiment* und in den *Ablation Studies*

Modell Nr.	Full Experiment	Ablation Study 01	Ablation Study 02
Mod_01	62,66%	-	-
Mod_02	62,40%	-	-
Mod_03	52,91%	-	-
Mod_04	52,89%	-	-
Mod_05	70,19%	-	-
Mod_06	70,22%	74,19%	80,50%
Mod_07	67,71%	-	-
Mod_08	67,97%	-	-
Mod_09	70,19%	36,44%	38,41%
Mod_10	69,29%	-	-
Mod_11	69,42%	-	-
Mod_12	70,55%	-	-
Mod_13	69,21%	-	-
Mod_14	67,49%	-	-
Mod_15	68,83%	-	-
Mod_16	68,12%	-	-
Mod_17	72,90%	-	-
Mod_18	70,84%	-	-
Mod_19	66,52%	-	-
Mod_20	79,65%	81,35%	86,47%
Mod_21	76,24%	-	-
Mod_22	16,97%	-	-
Mod_23	20,00%	-	-

Vorhersage der politischen Ausrichtung von Nachrichtenartikeln mittels Machine Learning

Mod_24	19,45%	-	-
Mod_25	86,77%	88,57	90,76%
Mod_26	40,49%	-	-
Mod_27	20,00%	-	-

Bemerkung: Leere Zellen bedeuten, dass diese Modelle nicht in den *Ablation Studies* berücksichtigt wurden

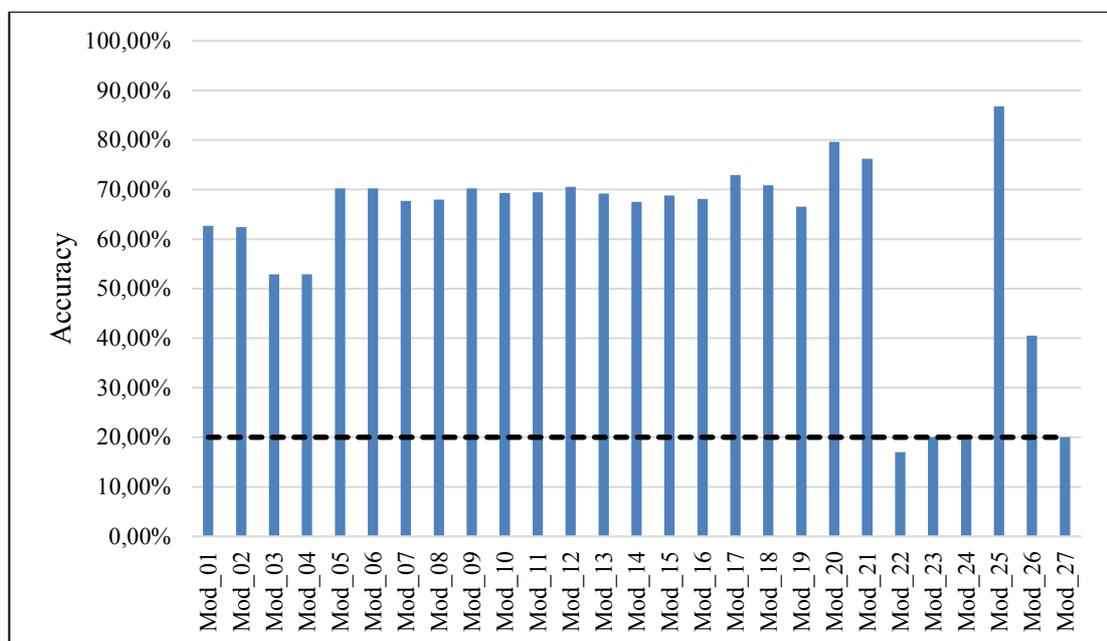


Abbildung 15: Accuracy Werte der Modelle, die im *Full Experiment* angewandt worden sind

Bemerkung: Die horizontale Linie stellt die *Accuracy* eines Klassifikators dar, das die politische Richtung einer Nachricht zufällig bestimmt

Die Abbildung 16 stellt die *Accuracy* Werte der Modelle dar, die sowohl im *Full Experiment* als auch in beiden *Ablation Studies* angewandt worden sind. Diese Modelle sind die mit *Trigrams* trainierte *SVM* (Mod_06), das neuronale Netz basierend auf der Architektur von Omernick & Chollet (2019) (Mod_09), das *bigru* Modell (Mod_20) und das *RoBERTa* Modell (Mod_25). Die *Accuracy* Werte der vier angewandten Modelle ist in der *Ablation Study 02* durchgehend höher als in der *Ablation Study 01*. Aus diesem Ergebnis lässt sich empfehlen, dass bei der Zuordnung von Nachrichten Anbietern aus dem halb linken und rechten politischen Spektrum, deren Zuteilung seitens der Mediennavigatoren nicht eindeutig sind, diese eher zu der Klasse *Center* zugewiesen werden anstatt in das Äußerste des politischen Spektrums. Während bei allen Modellen die *Accuracy* Werte in den *Ablation Studies* höher ausfällt als im *Full Experiment*, reduziert sich die *Accuracy* des neuronalen Netzes basierend

auf der Architektur von Omernick & Chollet (2019) (Mod_09) um fast die Hälfte. Dieses Ergebnis bestärkt die Empfehlung, dass man bei der Klassifikation des politischen *Bias* von Nachrichten mit bereits vorhandenen Modellen wie dem *bigru* Modell der *ktrain* Bibliothek schon erfolgsversprechende Resultate erzielen kann.

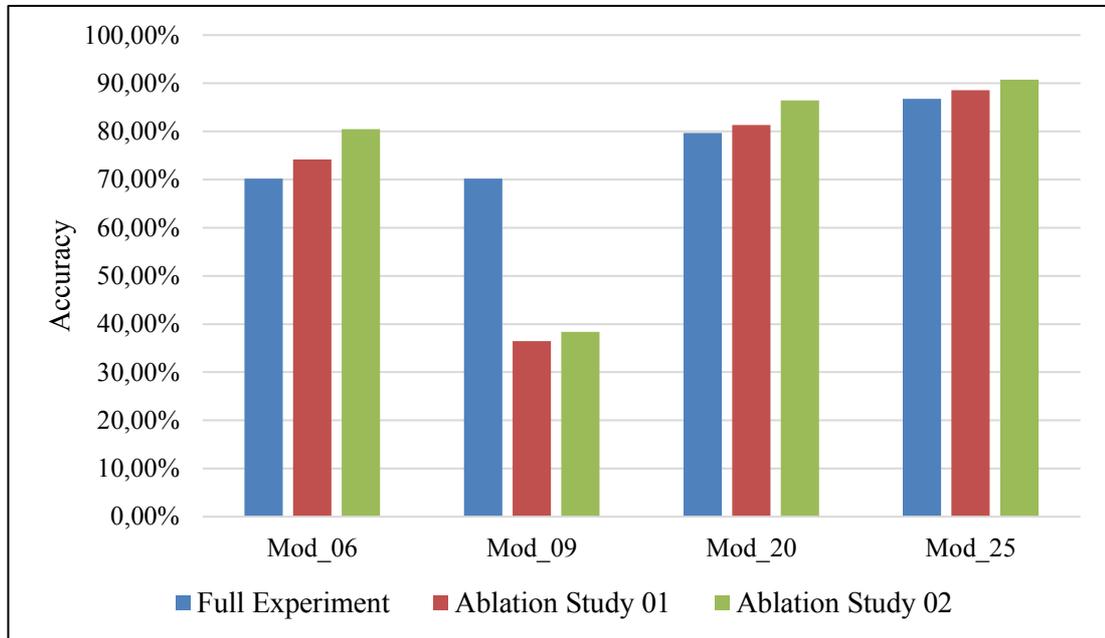


Abbildung 16: Accuracy Werte der Modelle, die im *Full Experiment* und in den *Ablation Studies* angewandt worden sind

Nachdem die *Accuracy* Werte für alle Modelle im *Full Experiment* und in den *Ablation Studies* inspiziert worden sind, liegt der Fokus nun auf die Länge der Trainingsdauer der im *Full Experiment* angewandten Modelle.

Das *bigru* Modell (Mod_20) benötigt zum Trainieren eines *Epochs* mit durchschnittlich 84 Sekunden nicht unwesentlich länger als Modelle, die mit der *Keras* Bibliothek implementiert worden sind (Mod_09 bis Mod_16) (siehe Abbildung 17). Das Trainieren jener Modelle aus der *transformers* Bibliothek (Mod_22 bis Mod_27) nimmt mindestens 25 Minuten für je ein *Epoch* in Anspruch. Im Vergleich dazu brauchen die Modelle aus der *scikit-learn* Bibliothek zum Trainieren nicht mehr als ungefähr 100 Sekunden, wobei die *SVM*, die logistische Regression und der Naive Bayes Klassifikator bereits innerhalb von zehn Sekunden trainiert worden sind. Die Betrachtung der Zeit, die die Modelle zum Trainieren in Anspruch nehmen, ist von wichtigem Ausmaß, da sie je nach Anwendungsfall häufig wieder trainiert werden müssen. Ein weiterer Grund, welches ein häufiges Trainieren der Modelle erfordert, ist, dass sich die politische Orientierung der Nachrichtenanbieter mit der Zeit verschieben kann.

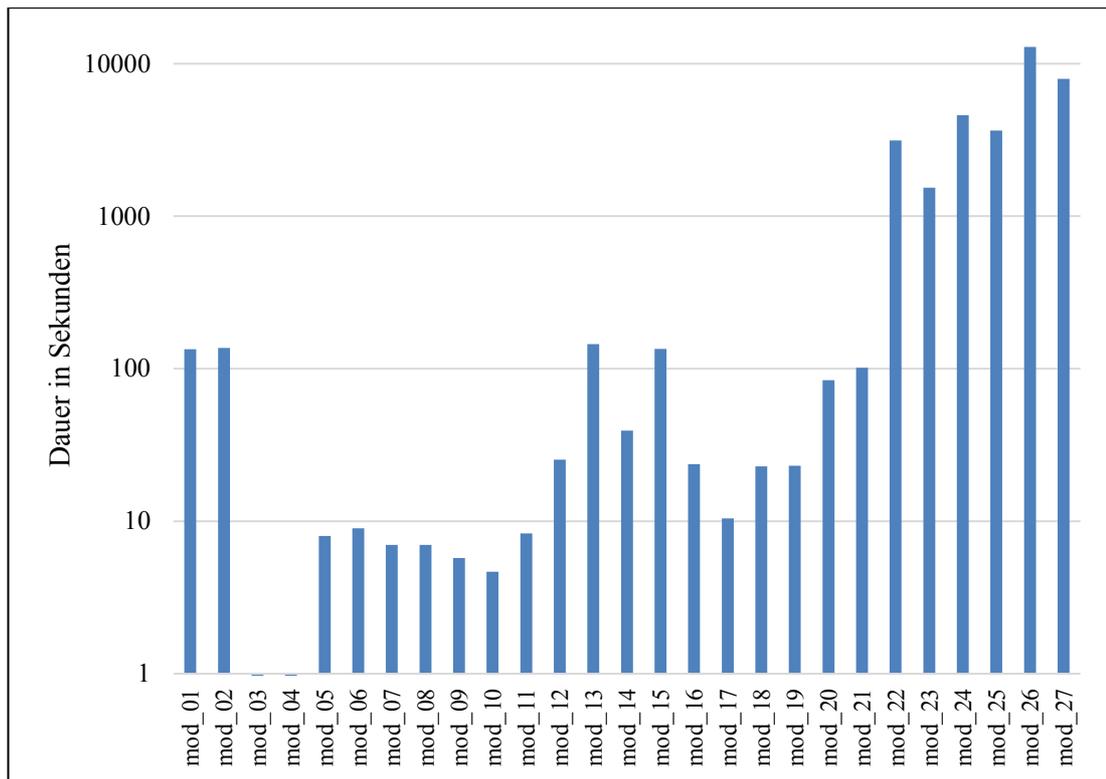


Abbildung 17: Dauer (in Sekunden), die ein Modell im *Full Experiment* zum Trainieren für durchschnittlich ein *Epoch* benötigt

Die Abbildung 18 zeigt die Relation der Modelle zwischen ihrer Trainingsdauer in Sekunden für je ein *Epoch* im Vergleich zu ihrer erzielten *Accuracy* im *Full Experiment*. Man erkennt, dass sich die 27 verschiedenen Modelle zu fünf homogenen Clustern gebildet haben. Das erste Cluster enthält die beiden Naive Bayes Klassifikatoren (Mod_03 und Mod_04) und ist dadurch charakterisiert, dass die beiden Modelle innerhalb von einer Sekunde trainiert worden sind und eine *Accuracy* von knapp 53% erzielen. Das zweite Cluster enthält 19 Modelle, und zwar unter anderem alle mit der *Keras* Bibliothek implementierten neuronalen Netze, die vordefinierten Modelle der *ktrain* Bibliothek sowie die übrigen Klassifikatoren der *scikit-learn* Bibliothek (Mod_01, Mod_02, Mod_05 bis Mod_21). Die Modelle in diesem Cluster sind dadurch gekennzeichnet, dass sie zum Trainieren für je ein *Epoch* maximal zweieinhalb Minuten benötigen und *Accuracy* Werte zwischen 60% und 80% liefern. In dem dritten Cluster befinden sich die *BERT*, *distilBERT*, *ALBERT* und *XLNET* Modelle (Mod_22 bis Mod_24 und Mod_27). Das Trainieren dieser Modelle dauert mindestens 1000 Sekunden für je ein *Epoch* und die Modelle sind dennoch nicht besser als ein Klassifikator, der die politische Richtung von Nachrichten zufällig klassifiziert. Das *XLM* Modell (Mod_26) stellt als einziges Modell das vierte Cluster dar, während

Vorhersage der politischen Ausrichtung von Nachrichtenartikeln mittels Machine Learning

das *RoBERTa* Modell (Mod_25) ebenfalls allein das fünfte Cluster repräsentiert. Beide Modelle benötigen sehr lange zum Trainieren eines *Epochs*, unterscheiden sich aber jedoch in ihrer Modellgüte außerordentlich. Während das *RoBERTa* Modell (Mod_25) die beste *Accuracy* aufweist, ist das *XLM* Modell (Mod_26) unter allen Modellen, die besser abschneiden als zufälliges Raten, das schlechteste Modell und beansprucht zum Trainieren eines *Epochs* mit mindestens zehntausend Sekunden die längste Zeit. Die grafische Darstellung des dritten Clusters bestätigt die oben genannte Empfehlung, dass möglichst auf bereits vorhandene Klassifikatoren wie dem *bigru* Modell zurückgriffen werden sollte. Das *bigru* Modell (Mod_20) benötigt zwar doppelt so lang wie der bidirektionale *GRU* (Mod_12), aber die knapp eine Minute längere Trainingsdauer des Modells verspricht auch eine fast um 10 Prozentpunkte höhere *Accuracy*.

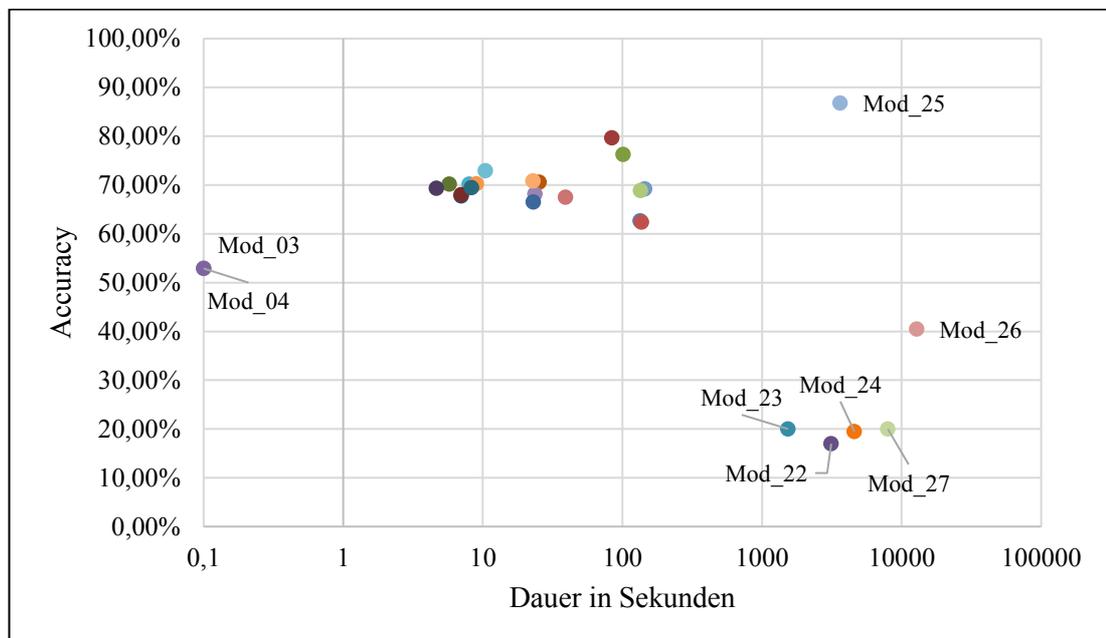


Abbildung 18: Verhältnis der Dauer (in Sekunden) zum Trainieren eines *Epochs* der angewandten Modelle im *Full Experiment* im Vergleich zu ihrer erzielten *Accuracy*

Abschließend werden die Konfusionsmatrizen des *RoBERTa* (Mod_25) und des *BERT* (Mod_22) Modells genauer betrachtet, welche im *Full Experiment* sowohl am besten als auch am schlechtesten abgeschnitten haben (siehe Tabelle 17 und Tabelle 18). Die Konfusionsmatrix des *RoBERTa* Modells (siehe Tabelle 17) zeigt deutlich, dass die meisten falsch klassifizierten Nachrichten sich in der Nähe der tatsächlichen politischen Ausrichtung befinden. Betrachtet man beispielsweise die Klasse *Center*, so sieht man, dass das *RoBERTa* Modell insgesamt 296 Nachrichten falsch klassifiziert. 182 der 296 falsch klassifizierten Nachrichten wurden um die Klasse

Vorhersage der politischen Ausrichtung von Nachrichtenartikeln mittels Machine Learning

Center zu den politischen Klassen *Lean Left* und *Lean Right* zugewiesen. Diese Beobachtung gilt für alle fünf politischen Klassen und bestätigt die Überlegenheit des *RoBERTa* Modells.

Im Kontrast dazu deckt die Konfusionsmatrix des *BERT* Modells (Mod_22) die Ursachen auf, weshalb das *BERT* Modell einen sehr niedrigen *Accuracy* Wert aufweist (siehe Tabelle 18). Das *BERT* Modell klassifiziert nahezu jede Nachricht entweder zu der Klasse *Center* oder *Right*. Die *Recall* Werte für die politische Richtung *Center* und *Right* betragen 47% und respektive 38% Prozent. Dies verdeutlicht, dass das *BERT* Modell die Mehrheit der Nachrichten in diesen politischen Richtungen fehlerhaft klassifiziert. Man erkennt auch, dass das *BERT* Modell 1233 Nachrichten aus dem linken Spektrum als rechte Nachrichten vorhersagt. Gleichzeitig sagt das *BERT* Modell die politische Richtung von nur 1145 rechten Nachrichten zutreffend voraus. Dies illustriert beispielhaft, dass das *BERT* Modell nicht in der Lage ist zwischen linken und rechten Nachrichten zu differenzieren.

Tabelle 17: Konfusionsmatrix des *RoBERTa* Modells im *Full Experiment*

<i>RoBERTa</i>	Left	Lean Left	Center	Lean Right	Right
Left	2713	130	70	38	49
Lean Left	148	2568	101	61	122
Center	138	100	2515	62	185
Lean Right	77	100	81	2449	293
Right	30	47	44	109	2770

Bemerkung: Die Spalten stellen die vorhergesagten Klassen dar, während die Zeilen die tatsächliche Klasse der Nachrichten wiedergeben.

Tabelle 18: Konfusionsmatrix des *BERT* Modells im *Full Experiment*

<i>BERT</i>	Left	Lean Left	Center	Lean Right	Right
Left	0	0	1767	0	1233
Lean Left	0	0	2150	1	849
Center	0	0	1400	0	1600
Lean Right	0	0	2005	0	995
Right	0	0	1854	1	1145

Bemerkung: Die Spalten stellen die vorhergesagten Klassen dar, während die Zeilen die tatsächliche Klasse der Nachrichten wiedergeben.

Zusammenfassend lassen sich folgende Designprinzipien für *Machine Learning* Modelle, die den politischen *Bias* von Nachrichten vorhersagen, deduzieren:

1. Es ist indifferent, ob Klassifikatoren mit *Bi-* oder *Trigrams* trainiert werden.
2. Es ist unerheblich, ob bidirektionale oder nicht-bidirektionale *Recurrent Neural Networks* bei der Klassifikation zum Einsatz kommen.
3. Bei der Implementierung von *Recurrent Neural Networks* sind *Gated Recurrent Units* den *Long Short-Term Memory* vorzuziehen.
4. Der Einsatz von bereits implementierten Klassifikatoren sollte intensiv in Erwägung gezogen und die Modellarchitektur für den eigenen Anwendungsfall angepasst werden.
5. Falls es keine verlässliche Zuweisung der politischen Richtung für Nachrichtenanbieter aus dem halb linken und rechten Spektrum existiert, so sollten Nachrichten jenen Anbieters im Zweifel zur politischen Mitte zugeordnet werden.
6. Darüber hinaus sollte man bei der Auswahl eines Modells die Zeit zum Trainieren miteinkalkulieren. Die Dringlichkeit das Modell im produktiven Betrieb einzusetzen, die Größe des Trainingsdatensatzes sowie die verwendete Hardware bestimmen maßgeblich die Dauer des Trainierens.

5 Diskussion

Während der Implementierung konnten essenzielle Erkenntnisse gewonnen werden wie man mit Machine Learning Modellen den politischen *Bias* von Nachrichten automatisiert ermitteln kann. Die hier abgeleiteten sechs Designprinzipien sind das Resultat einer kontrollierten Testumgebung, in der einzig und allein die Maximierung der *Accuracy* im Vordergrund stand. Darin liegt auch der erste Schwachpunkt dieser Arbeit begründet. Anhand der Ergebnisse wurde beispielsweise abgeleitet, dass bei der Implementierung von *RNN* die Modellarchitektur des *GRU* den gegenüber des *LSTM* präferiert werden soll (3. Designprinzip). Diese Empfehlung basiert ausschließlich darauf, dass die *GRU* eine höhere *Accuracy* als die *LSTM* erzielt haben. Darüber hinaus wird keine mögliche Erklärung aufgeführt, weshalb die *GRU* den *LSTM* überlegen sind. Grundsätzlich werden alle hier implementierten Modelle als Blackbox interpretiert, deren Funktionsweise für den spezifischen Anwendungsfall dieser Arbeit nicht näher

durchleuchtet wird. Dies schränkt die Optimierung der Modelle nachträglich ein, da keine genauen Informationen gewonnen werden, weshalb die Modelle Nachrichten in ein bestimmtes politisches Spektrum zuweisen. Im Speziellen ist dieses Wissen von großer Relevanz, um beispielsweise nachzuvollziehen, warum das *BERT* Modell fast alle Nachrichten als neutrale oder rechte Nachrichten klassifiziert hat.

Der Fokus der Arbeit liegt nicht vorrangig in der Entwicklung eines Modells, welches die politische Ausrichtung nahezu jeder Nachricht korrekt ermittelt. Die Formulierung von Designprinzipien hingegen ist Gegenstand dieser Arbeit gewesen, welche nun in weiterführenden Arbeiten validiert werden müssen. Zum Beispiel können nachfolgende Arbeiten das *Full Experiment* auf dem NELA-GT-2019 Datensatz von Gruppi et al. (2020) replizieren und feststellen, ob die hier ermittelten Designprinzipien weiterhin ihre Gültigkeit besitzen. Ferner könnte in einer weiteren Arbeit validiert werden, ob die hier ermittelten Designprinzipien auch in der Klassifikation von nicht-englischsprachigen Nachrichten gelten. Falls weitere Arbeiten die Gültigkeit der Designprinzipien für deutsche oder französische Nachrichten bestätigen können, dann liegt die Annahme nahe, dass die Designprinzipien sprachunabhängig und somit losgelöst von länderspezifischen Mediennavigatoren und -landschaften sind.

Eine umfangreichere, aber viel realitätsnähere und weiterführende Arbeit ist das Einbetten von performanten Modellen wie beispielsweise dem *bigru* Modell oder das *RoBERTa* Modell in eine mobile App oder Webapplikation. Diese Applikation ist im Idealfall in der Lage Nachrichten von verschiedenen Anbietern in nahezu Echtzeit zu *scrapen* und im Anschluss die politische Richtung vorherzusagen. Diese Lösung bietet den Vorteil, dass medienkritische Bürger zeitnah den Konsum von überaus parteilichen Nachrichten anhand ihres politischen *Bias* vermeiden können. Ein weiterer Nutzen für die Gesellschaft läge darin, dass die Nachrichten der verschiedenen Nachrichtenanbieter bezüglich ihrer enthaltenen politischen Ausrichtung transparent sind. Ein Kritikpunkt dieser Arbeit ist, dass die hier in den Trainings- und Testdatensätzen berücksichtigten Nachrichten nicht aktuell sind und dadurch nicht die heute gegenwärtigen gesellschaftlichen, politischen und wirtschaftlichen Themen und Entwicklungen abdecken. Dieses Defizit könnte durch die Operationalisierung eines Modells in eine Applikation beseitigt werden. Weiterhin könnte die Applikation eine Funktion beinhalten (ähnlich wie das *Community*

*Feedback*²⁸ von *AllSides*), in der die Benutzer dem Modell mitteilen können, ob sie mit der Vorhersage der politischen Ausrichtung einer Nachricht übereinstimmen. Zusätzlich könnten die Benutzer bei fehlerhaft klassifizierten Nachrichten die politische Richtung angeben, die ihrer Meinung nach richtig ist. Die Rückführungen der Vorhersagen als neue Trainingsdatensätze seitens der Applikation befähigt das Modell aus seinen eigenen Fehlern zu lernen. Das mehrmalige Trainieren kann entweder periodisch (täglich, wöchentlich, 14-tägig, monatlich, usw.) oder ereignisgesteuert erfolgen. Im letzteren Szenario wäre es vorstellbar, dass die Applikation das Training des Modells veranlasst, sobald eine bestimmte Anzahl (z.B. 1000) von Klassifikationen als neue Trainingsinstanzen verfügbar ist. In der Praxis ist auch eine hybride Trainingsprozedur der beiden vorangestellten Methoden realisierbar.

Neben der Operationalisierung eines performanten Modells bietet die Applikation die Möglichkeit das 6. Designprinzip im Hinblick auf einen Vergleich der Modellgüte zwischen dem *bigru* Modell und dem *RoBERTa* Modell zu evaluieren. In den Ergebnissen wurde die Empfehlung ausgesprochen das *bigru* Modell gegenüber dem auf der Architektur von Omernick & Chollet (2019) basierenden neuronalen Netz vorzuziehen. Obwohl das Trainieren des *bigru* Modells ungefähr doppelt so viel Zeit in Anspruch nimmt, erzielte es eine knapp um 10 Prozentpunkte bessere *Accuracy*. Die Formulierung des 4. Designprinzips erfolgte, da der zeitliche Mehraufwand das *bigru* Modell zu trainieren nicht mehr als eine Minute beträgt. Allerdings erstreckt sich das Training des *RoBERTa* Modells auf ungefähr eine Stunde, während das *bigru* Modell in dieser Zeit 84-mal trainiert werden kann. Es stellt sich hierbei die Frage, inwiefern der enorme zeitliche Mehraufwand die um sieben Prozentpunkte höhere *Accuracy* des *RoBERTa* Modells rechtfertigt. Hier sollte die Möglichkeit in Betracht gezogen werden, ob das *bigru* Modell durch häufiges Trainieren eine ähnliche Leistung wie das *RoBERTa* Modell erreichen kann. Dennoch muss hier erwähnt werden, dass das *bigru* Modell im *Full Experiment* bereits sieben *Epochs* lang trainiert wurde, während das *RoBERTa* Modell aufgrund der langen Trainingszeit nur ein *Epoch* trainiert wurde. Gleichzeitig muss der Einwand erhoben werden, dass die nachgewiesene *Accuracy* des *RoBERTa* Modells von ca. 86% noch höher hätte

²⁸ <https://www.allsides.com/media-bias/media-bias-rating-methods>

ausfallen können, weil mit Sicherheit nicht ausgeschlossen werden kann, dass der *val_loss* sein Minimum erreicht hat.

Zu guter Letzt muss der Punkt angemerkt werden, dass die Mediennavigatoren und deren Zuweisung der politischen Richtungen mittelbar die *Accuracy* Werte aller Modelle beeinflussen. Alle Modelle lernen nicht unbedingt die wahre politische Ausrichtung einer Nachricht, sondern sie generalisieren Merkmale, die Nachrichten einer zugewiesenen politischen Klasse gemein haben. Im *Target Labeling* ist es theoretisch möglich linke Nachrichten als rechte Nachrichten aufzufassen, und umgekehrt. Die Klassen halb links und halb rechts können ebenfalls analog vertauscht werden. Nun wäre es interessant zu beobachten, ob die Modelle (trotz fehlerhafter Klassenzuweisung) ähnliche *Accuracy* Werte erzielen. Beide *Ablation Studies* zeigen in drei von vier Fällen, dass sich die *Accuracy* Werte verbessern, wenn drei statt fünf Klassen berücksichtigt werden. Dies legt die Vermutung nahe, dass die Modellgüte durch eine zum Teil willkürliche Zuordnung der politischen Richtung künstlich verbessert werden kann. Es ist aber nicht im Interesse, dass vollkommen perfekte Modelle in der Praxis eingesetzt werden, die eine zugewiesene Klasse einer Nachricht ermitteln. Der gesellschaftliche Beitrag eines in der Praxis eingesetzten Modells ist es die politische Ausrichtung von Nachrichten für die Meinungsbildung der Bürger transparent zu machen. Deshalb ist die Validierung der vorhandenen Mediennavigatoren durch Wissenschaftler und unabhängigen Organisationen unerlässlich. Zum einen beruht das *Target Labeling* dadurch auf einem soliden Fundament und zum anderen kann durch die Offenlegung der Methodik Vertrauen in der Bevölkerung für die in der Praxis eingesetzten Applikation geschaffen werden.

6 Fazit

In dieser Arbeit wurde das Ziel gesetzt Designprinzipien für die automatisierte Ermittlung der politischen Ausrichtung von Nachrichten zu formulieren. Die durch die Literaturrecherche gewonnenen Anregungen wurden in der Implementierung, vor allem in der Modellanwendung, erfolgreich umgesetzt. Anhand der Resultate des *Full Experiments* und der beiden *Ablation Studies* konnten insgesamt sechs Designprinzipien abgeleitet werden. Mit der Formulierung der Designprinzipien ergeben sich in der Forschung neue Pfade, die sowohl die Validität der Designprinzipien als auch die Operationalisierung eines Modells betreffen. Zum einen

müssen die Designprinzipien dahingehend überprüft werden, ob sie auf weiteren Datensätzen ihre Gültigkeit besitzen und zum anderen ist es von großem Interesse herauszufinden, ob die Designprinzipien sprachunabhängig sind. Ferner haben in dieser Arbeit angewandten Modelle das Potential einen gesellschaftlichen Beitrag zu leisten. Die Operationalisierung eines performanten Modells kann maßgeblich zur Meinungsbildung in offenen Gesellschaften beitragen. Mit der Formulierung der Designprinzipien ist ein Ausgangspunkt für zukünftige Arbeiten gelegt worden, die nicht nur der wissenschaftlichen Gemeinde von Nutzen ist.

Literaturverzeichnis

- 204,000 news articles and essays. (o. J.). *Components*. Abgerufen 8. Juni 2020, von <https://components.one/datasets/all-the-news-articles-dataset/>
- Ad Fontes Media. (o. J.-a). *About Ad Fontes Media*. Ad Fontes Media. Abgerufen 15. Juni 2020, von <https://www.adfontesmedia.com/about-ad-fontes-media/>
- Ad Fontes Media. (o. J.-b). *How Ad Fontes Ranks News Sources*. Ad Fontes Media. Abgerufen 15. Juni 2020, von <https://www.adfontesmedia.com/how-ad-fontes-ranks-news-sources/>
- Ahmed, A., & Xing, E. (2010). Staying Informed: Supervised and Semi-Supervised Multi-View Topical Analysis of Ideological Perspective. *Proceedings of the 2010 Conference on Empirical Methods in Natural Language Processing*, 1140–1150. <https://www.aclweb.org/anthology/D10-1111>
- Aires, V. P., Nakamura, F. G., & Nakamura, E. F. (2019). A Link-based Approach to Detect Media Bias in News Websites. *Companion Proceedings of The 2019 World Wide Web Conference*, 742–745. <https://doi.org/10.1145/3308560.3316460>
- Bahdanau, D., Cho, K., & Bengio, Y. (2016). Neural Machine Translation by Jointly Learning to Align and Translate. *arXiv:1409.0473 [cs, stat]*. <http://arxiv.org/abs/1409.0473>
- Baly, R., Karadzhov, G., Alexandrov, D., Glass, J., & Nakov, P. (2018). Predicting Factuality of Reporting and Bias of News Media Sources. *arXiv:1810.01765 [cs, stat]*. <http://arxiv.org/abs/1810.01765>
- Baly, R., Karadzhov, G., Saleh, A., Glass, J., & Nakov, P. (2019). Multi-Task Ordinal Regression for Jointly Predicting the Trustworthiness and the Leading Political Ideology of News Media. In J. Burstein, C. Doran, & T. Solorio (Hrsg.),

- Proceedings of the 2019 Conference of the North* (S. 2109–2116). Association for Computational Linguistics. <https://doi.org/10.18653/v1/N19-1216>
- Biessmann, F. (2016). Automating Political Bias Prediction. *arXiv:1608.02195 [cs]*. <http://arxiv.org/abs/1608.02195>
- Cho, K., van Merriënboer, B., Gulcehre, C., Bahdanau, D., Bougares, F., Schwenk, H., & Bengio, Y. (2014). Learning Phrase Representations using RNN Encoder-Decoder for Statistical Machine Translation. *arXiv:1406.1078 [cs, stat]*. <http://arxiv.org/abs/1406.1078>
- Chun, S. A., Holowczak, R. D., Dharan, K., Wang, R., Basu, S., & Geller, J. (2019). Detecting Political Bias Trolls in Twitter Data. *WEBIST*. <https://doi.org/10.5220/0008350303340342>
- Chung, J., Gulcehre, C., Cho, K., & Bengio, Y. (2014). Empirical Evaluation of Gated Recurrent Neural Networks on Sequence Modeling. *arXiv:1412.3555 [cs]*. <http://arxiv.org/abs/1412.3555>
- Cohen, R., & Ruths, D. (2013). Classifying Political Orientation on Twitter: It's Not Easy! *ICWSM*.
- Conover, M. D., Goncalves, B., Ratkiewicz, J., Flammini, A., & Menczer, F. (2011). Predicting the Political Alignment of Twitter Users. *2011 IEEE Third International Conference on Privacy, Security, Risk and Trust and 2011 IEEE Third International Conference on Social Computing*, 192–199. <https://doi.org/10.1109/PASSAT/SocialCom.2011.34>
- Dallmann, A., Lemmerich, F., Zoller, D., & Hotho, A. (2015). Media Bias in German Online Newspapers. *HT '15*. <https://doi.org/10.1145/2700171.2791057>
- Dardis, F. E., Baumgartner, F. R., Boydston, A. E., De Boef, S., & Shen, F. (2008). Media Framing of Capital Punishment and Its Impact on Individuals'

- Cognitive Responses. *Mass Communication and Society*, 11(2), 115–140.
<https://doi.org/10.1080/15205430701580524>
- Devlin, J., Chang, M.-W., Lee, K., & Toutanova, K. (2019). BERT: Pre-training of Deep Bidirectional Transformers for Language Understanding. *arXiv:1810.04805 [cs]*. <http://arxiv.org/abs/1810.04805>
- Dumitru, V. C., & Rebedea, T. (2019). Fake and Hyper-partisan News Identification. *RoCHI*.
- Elman, J. L. (1990). Finding Structure in Time. *Cognitive Science*, 14(2), 179–211.
https://doi.org/10.1207/s15516709cog1402_1
- Gangula, R. R. R., Duggenpudi, S. R., & Mamidi, R. (2019). Detecting Political Bias in News Articles Using Headline Attention. *Proceedings of the 2019 ACL Workshop BlackboxNLP: Analyzing and Interpreting Neural Networks for NLP*, 77–84. <https://doi.org/10.18653/v1/W19-4809>
- Golbeck, J., & Hansen, D. (2014). A method for computing political preference among Twitter followers. *Social Networks*, 36, 177–184.
<https://doi.org/10.1016/j.socnet.2013.07.004>
- Groseclose, T., & Milyo, J. (2005). A Measure of Media Bias. *The Quarterly Journal of Economics*, 120(4), 1191–1237.
<https://doi.org/10.1162/003355305775097542>
- Gross, J. H., Acree, B., Sim, Y., & Smith, N. A. (2013). Testing the Etch-a-Sketch Hypothesis: A Computational Analysis of Mitt Romney’s Ideological Makeover During the 2012 Primary vs. General Elections. *APSA 2013 Annual Meeting Paper*. American Political Science Association 2013 Annual Meeting.
https://papers.ssrn.com/sol3/papers.cfm?abstract_id=2299991

- Gruppi, M., Horne, B. D., & Adali, S. (2020). NELA-GT-2019: A Large Multi-Labelled News Dataset for The Study of Misinformation in News Articles. *arXiv:2003.08444 [cs]*. <http://arxiv.org/abs/2003.08444>
- Hamborg, F., Donnay, K., & Gipp, B. (2019). Automated identification of media bias in news articles: An interdisciplinary literature review. *International Journal on Digital Libraries*, 20(4), 391–415. <https://doi.org/10.1007/s00799-018-0261-y>
- Hausladen, C. I., Schubert, M. H., & Ash, E. (2020). Text classification of ideological direction in judicial opinions. *International Review of Law and Economics*, 62, 105903. <https://doi.org/10.1016/j.irlle.2020.105903>
- Hirst, G., Riabinin, Y. V., & Graham, J. (2010). *Party status as a confound in the automatic classification of political speech by ideology*.
- Hochreiter, S., & Schmidhuber, J. (1997). Long short-term memory. *Neural Computation*, 9(8), 1735–1780. <https://doi.org/10.1162/neco.1997.9.8.1735>
- Horne, B. D., Dron, W., Khedr, S., & Adali, S. (2018). Sampling the News Producers: A Large News and Feature Data Set for the Study of the Complex Media Landscape. *arXiv:1803.10124 [cs]*. <http://arxiv.org/abs/1803.10124>
- How AllSides Rates Media Bias: Our Methods*. (2016, August 10). AllSides. <https://www.allsides.com/media-bias/media-bias-rating-methods>
- Iyyer, M., Enns, P., Boyd-Graber, J., & Resnik, P. (2014). Political Ideology Detection Using Recursive Neural Networks. *Proceedings of the 52nd Annual Meeting of the Association for Computational Linguistics (Volume 1: Long Papers)*, 1113–1122. <https://doi.org/10.3115/v1/P14-1105>

- Joulin, A., Grave, E., Bojanowski, P., & Mikolov, T. (2016). Bag of Tricks for Efficient Text Classification. *arXiv:1607.01759 [cs]*. <http://arxiv.org/abs/1607.01759>
- Kim, J.-G., & Lee, K.-S. (2014). Predicting Political Orientation of News Articles Based on User Behavior Analysis in Social Network. *IEICE Trans. Inf. Syst.* <https://doi.org/10.1587/transinf.E97.D.685>
- Knight, B., & Chiang, C.-F. (2008). *Media Bias and Influence: Evidence from Newspaper Endorsements*. National Bureau of Economic Research. <https://doi.org/10.3386/w14445>
- Kowsari, K., Jafari Meimandi, K., Heidarysafa, M., Mendu, S., Barnes, L., & Brown, D. (2019). Text Classification Algorithms: A Survey. *Information*, 10(4), 150. <https://doi.org/10.3390/info10040150>
- Kulkarni, V., Ye, J., Skiena, S., & Wang, W. Y. (2018). Multi-view Models for Political Ideology Detection of News Articles. *Proceedings of the 2018 Conference on Empirical Methods in Natural Language Processing*, 3518–3527. <https://doi.org/10.18653/v1/D18-1388>
- Lample, G., & Conneau, A. (2019). Cross-lingual Language Model Pretraining. *arXiv:1901.07291 [cs]*. <http://arxiv.org/abs/1901.07291>
- Lan, Z., Chen, M., Goodman, S., Gimpel, K., Sharma, P., & Soricut, R. (2020). ALBERT: A Lite BERT for Self-supervised Learning of Language Representations. *arXiv:1909.11942 [cs]*. <http://arxiv.org/abs/1909.11942>
- Lin, W.-H., Xing, E., & Hauptmann, A. (2008). A Joint Topic and Perspective Model for Ideological Discourse. In W. Daelemans, B. Goethals, & K. Morik (Hrsg.), *Machine Learning and Knowledge Discovery in Databases* (S. 17–32). Springer. https://doi.org/10.1007/978-3-540-87481-2_2

- Liu, Y., Ott, M., Goyal, N., Du, J., Joshi, M., Chen, D., Levy, O., Lewis, M., Zettlemoyer, L., & Stoyanov, V. (2019). RoBERTa: A Robustly Optimized BERT Pretraining Approach. *arXiv:1907.11692 [cs]*. <http://arxiv.org/abs/1907.11692>
- Maiya, A. S. (o. J.). *Text Classification Example: Sentiment Analysis with IMDb Movie Reviews*. Abgerufen 15. Oktober 2020, von <https://nbviewer.jupyter.org/github/amaiya/ktrain/blob/master/tutorials/tutorial-04-text-classification.ipynb>
- Maiya, A. S. (2020). ktrain: A Low-Code Library for Augmented Machine Learning. *arXiv:2004.10703 [cs]*. <http://arxiv.org/abs/2004.10703>
- Malte, A., & Ratadiya, P. (2019). Evolution of transfer learning in natural language processing. *arXiv:1910.07370 [cs]*. <http://arxiv.org/abs/1910.07370>
- Matthew Gentzkow, & Jesse M. Shapiro. (2010). What Drives Media Slant? Evidence From U.S. Daily Newspapers. *Econometrica*, 78(1), 35–71. <https://doi.org/10.3982/ECTA7195>
- Maynard, D., & Funk, A. (2012). Automatic Detection of Political Opinions in Tweets. In R. García-Castro, D. Fensel, & G. Antoniou (Hrsg.), *The Semantic Web: ESWC 2011 Workshops* (S. 88–99). Springer. https://doi.org/10.1007/978-3-642-25953-1_8
- Media Bias Ratings AllSides*. (2020, März 12). <https://www.allsides.com/media-bias/media-bias-ratings>
- Mikolov, T., Chen, K., Corrado, G., & Dean, J. (2013). Efficient Estimation of Word Representations in Vector Space. *arXiv:1301.3781 [cs]*. <http://arxiv.org/abs/1301.3781>

- Mikolov, T., Sutskever, I., Chen, K., Corrado, G. S., & Dean, J. (2013). Distributed Representations of Words and Phrases and their Compositionality. In C. J. C. Burges, L. Bottou, M. Welling, Z. Ghahramani, & K. Q. Weinberger (Hrsg.), *Advances in Neural Information Processing Systems 26* (S. 3111–3119). Curran Associates, Inc. <http://papers.nips.cc/paper/5021-distributed-representations-of-words-and-phrases-and-their-compositionality.pdf>
- Misra, A., & Basak, S. (2016). *Political Bias Analysis*.
- Norregaard, J., Horne, B. D., & Adali, S. (2019). NELA-GT-2018: A Large Multi-Labelled News Dataset for The Study of Misinformation in News Articles. *arXiv:1904.01546 [cs]*. <http://arxiv.org/abs/1904.01546>
- Omernick, M., & Chollet, F. (2019, November 6). *Keras documentation: Text classification from scratch*. https://keras.io/examples/nlp/text_classification_from_scratch/
- Park, S., Ko, M., Kim, J., Liu, Y., & Song, J. (2011). The politics of comments: Predicting political orientation of news stories with commenters' sentiment patterns. *Proceedings of the ACM 2011 conference on Computer supported cooperative work*, 113–122. <https://doi.org/10.1145/1958824.1958842>
- Patricia Aires, V., G. Nakamura, F., & F. Nakamura, E. (2019). A Link-based Approach to Detect Media Bias in News Websites. In Liu (Hrsg.), *Companion Proceedings of The 2019 World Wide Web Conference* (S. 742–745). ACM / Association for Computing Machinery. <https://doi.org/10.1145/3308560.3316460>
- Pennington, J., Socher, R., & Manning, C. (2014). GloVe: Global Vectors for Word Representation. *Proceedings of the 2014 Conference on Empirical Methods in*

Natural Language Processing (EMNLP), 1532–1543.

<https://doi.org/10.3115/v1/D14-1162>

Porter, M. F. (1980). An algorithm for suffix stripping. *Program*, 14(3), 130–137.

<https://doi.org/10.1108/eb046814>

Potthast, M., Kiesel, J., Reinartz, K., Bevendorff, J., & Stein, B. (2018). A Stylometric

Inquiry into Hyperpartisan and Fake News. *Proceedings of the 56th Annual*

Meeting of the Association for Computational Linguistics (Volume 1: Long

Papers), 231–240. <https://doi.org/10.18653/v1/P18-1022>

Preoțiuc-Pietro, D., Liu, Y., Hopkins, D., & Ungar, L. (2017). Beyond Binary Labels:

Political Ideology Prediction of Twitter Users. In R. Barzilay & M.-Y. Kan

(Hrsg.), *Proceedings of the 55th Annual Meeting of the Association for (S.*

729–740). Association for Computational Linguistics.

<https://doi.org/10.18653/v1%2FP17-1068>

Provost, F., & Fawcett, T. (2013). *Data Science for Business*. O'Reilly Media, Inc.

<https://www.oreilly.com/library/view/data-science-for/9781449374273/>

Recker, J. (2013). *Scientific research in information systems: A beginner's guide*.

Springer. <https://doi.org/10.1007/978-3-642-30048-6>

Sanh, V., Debut, L., Chaumond, J., & Wolf, T. (2020). DistilBERT, a distilled version

of BERT: Smaller, faster, cheaper and lighter. *arXiv:1910.01108 [cs]*.

<http://arxiv.org/abs/1910.01108>

Sapiro-Gheiler, E. (2018). „Read My Lips“: Using Automatic Text Analysis to

Classify Politicians by Party and Ideology. *arXiv:1809.00741 [cs, econ, q-fin]*.

<http://arxiv.org/abs/1809.00741>

Stukal, D., Sanovich, S., Tucker, J. A., & Bonneau, R. (2019). For Whom the Bot

Tolls: A Neural Networks Approach to Measuring Political Orientation of

Twitter Bots in Russia. *SAGE Open*.

<https://doi.org/10.1177/2158244019827715>

Sutskever, I., Vinyals, O., & Le, Q. V. (2014). *Sequence to Sequence Learning with Neural Networks*. <https://arxiv.org/abs/1409.3215v3>

Swiss Propaganda Research. (2019, März 18). *The Media Navigator*. Swiss Policy Research. <https://swprs.org/media-navigator/>

Thomas, M., Pang, B., & Lee, L. (2006). Get out the vote: Determining support or opposition from Congressional floor-debate transcripts. *Proceedings of the 2006 Conference on Empirical Methods in Natural Language Processing*, 327–335. <https://www.aclweb.org/anthology/W06-1639>

Vaswani, A., Shazeer, N., Parmar, N., Uszkoreit, J., Jones, L., Gomez, A. N., Kaiser, L., & Polosukhin, I. (2017). Attention Is All You Need. *arXiv:1706.03762 [cs]*. <http://arxiv.org/abs/1706.03762>

Volkova, S., Coppersmith, G., & Van Durme, B. (2014). Inferring User Political Preferences from Streaming Communications. *Proceedings of the 52nd Annual Meeting of the Association for Computational Linguistics (Volume 1: Long Papers)*, 186–196. <https://doi.org/10.3115/v1/P14-1018>

Wang, S., & Manning, C. (2012). Baselines and Bigrams: Simple, Good Sentiment and Topic Classification. *Proceedings of the 50th Annual Meeting of the Association for Computational Linguistics (Volume 2: Short Papers)*, 90–94. <https://www.aclweb.org/anthology/P12-2018>

Wilke, J. (2012, Mai 31). *Funktionen und Probleme der Medien*. bpb.de. <https://www.bpb.de/politik/grundfragen/deutsche-verhaeltnisse-eine-sozialkunde/139163/funktionen-und-probleme>

- Yan, H., Lavoie, A., & Das, S. (2017). *The Perils of Classifying Political Orientation From Text*. 38–50.
- Yang, Z., Dai, Z., Yang, Y., Carbonell, J., Salakhutdinov, R., & Le, Q. V. (2020). XLNet: Generalized Autoregressive Pretraining for Language Understanding. *arXiv:1906.08237 [cs]*. <http://arxiv.org/abs/1906.08237>
- Yu, B., Kaufmann, S., & Diermeier, D. (2008). Classifying Party Affiliation from Political Speech. *Journal of Information Technology & Politics*, 5(1), 33–48. <https://doi.org/10.1080/19331680802149608>
- Zhao, J., Ryzhik, A., & Lee, N. (2018). *DeepNews. AI: Detecting Political Bias*. *CS 230 Deep Learning*. https://cs230.stanford.edu/projects_spring_2018/reports/8286849.pdf
- Zhou, D. X., Resnick, P., & Mei, Q. (2011). Classifying the Political Leaning of News Articles and Users from User Votes. *ICWSM*.