

Der Neutralitätsanspruch von Qualitätsjournalismus am Beispiel des Politpodcasts „Lage der Nation“ – Eine Sentiment-Analyse mit maschinellem Lernen

Inhaltsverzeichnis

1	Sozialwissenschaftliche Analysen der digitalisierten Medienwelt – eine Herausforderung.....	1
2	Podcasts zwischen politischem Journalismus und maschineller Sprachverarbeitung	2
3	„Lage der Nation“, „Spiegel Online“ und „FAZ.NET“ – Beispiele für Qualitätsjournalismus?.....	4
4	Methodik und Technische Umsetzung	8
5	Modelltraining und Sentiment-Analyse: Auswertung	11
5.1	Modellauswahl, Finetuning und Hyperparametertraining	11
5.2	Evaluierung Sentiment-Analyse: „Lage der Nation“	14
5.3	Evaluierung Sentiment-Analyse: „Spiegel Online“.....	16
5.4	Evaluierung Sentiment-Analyse: „FAZ.NET“	18
5.5	„Lage der Nation“, „Spiegel Online“ und „FAZ.NET“ – Berichterstattung im Quervergleich.....	20
6	Machine Learning und geisteswissenschaftliche Textanalyse - Fazit, Einordnung, Ausblick	22
7	Literaturverzeichnis	26
8	Datenquellenverzeichnis.....	29
9	Abbildungsverzeichnis	30
10	Datenschutzbestimmungen und Urheberrechtliche Erklärung	30
	Plagiatserklärung.....	31

1 Sozialwissenschaftliche Analysen der digitalisierten Medienwelt – eine Herausforderung

Das 21. Jahrhundert ist geprägt von Fake News und der Verdrängung von klassischen Printmedien durch eine Flut an Digitalen Medien, die, um zu bestehen, immer schneller und leichter zugänglich publiziert werden müssen (vgl. Sadrozinski, 2013: 84ff; 87). Dennoch besteht eine ungebrochene Relevanz von neutralem Qualitätsjournalismus, dem als Stütze von Demokratie und Rechtsstaatlichkeit eine zentrale Rolle in jedem politischen System zugeordnet wird und der in den letzten Jahren verstärkt an Bedeutung gewinnt (vgl. Patzelt, 2013: 141, 227ff, 422ff).

Doch wie kann die Neutralität von Berichterstattung und insbesondere von politischem Journalismus überprüft und analysiert werden? Wie kann die bei etablierten Medienhäusern in Deutschland altbekannte leichte politische Färbung im politischen Spektrum auf die neuen Arten von Journalismus umgelegt und eingeordnet werden? Eine besondere Form von gesprochenem Journalismus soll die technische Untersuchungsbasis dieser Abhandlung bilden: Podcasts, also Audio-Aufnahmen von Diskussionen und Gesprächen. Dazu soll eine, durch maschinelles Lernen gestützte, Methode entwickelt werden, um journalistische Produkte schnell und in großem Umfang hinsichtlich ihrer politischen Schlagrichtung überprüfen zu können und so ein Werkzeug entwickelt werden, mit dem automatisiert politisch einseitige, der Stimmungsmache dienende, Texte und Medien erkannt werden können und deren politische Denkart und Parteilichkeit identifiziert werden kann. Bei bekannten und etablierten Zeitungen und Zeitschriften ist deren politische Ausrichtung im Allgemeinen bekannt. Bei der sich immer schneller entwickelnden Welt der digitalen Medien dient ein maschinell gestütztes Werkzeug als wichtige Methode, um auch neue und unbekannte mediale Produkte in großem Umfang klassifizieren und sie so in eine politische Richtung einordnen zu können. Dies ist die Motivation dieser Projektanalyse und soll am Beispiel des Politpodcasts „Lage der Nation“ exemplarisch durchgeführt werden. Die erwähnte Einstufung von althergebrachten Medien soll als Vergleichsparameter dienen und beispielhaft anhand der Onlineportale der Medienhäusern des Spiegelverlags und der „Frankfurter Allgemeinen Zeitung“ untersucht werden. Damit soll die Analysemethode bewertbar gemacht und zugleich ein Vergleich der Ergebnisse ermöglicht werden. Dabei soll dasselbe Klassifizierungsmodell angewandt werden, dass nachfolgend auch zur politischen Einordnung und Klassifizierung der „Lage der Nation“

entwickelt wird. Vorhergehend wird der bisherige Forschungsstand im Dreieck der maschinellen Sprachverarbeitung, der neuen Medienform Podcasts und der politischen Einordnung von Medienprodukten aufgearbeitet.

2 Podcasts zwischen politischem Journalismus und maschineller Sprachverarbeitung

Die quantitative und qualitative Klassifizierung von Stimmung und Meinung in politischen Texten ist in den Politik- und Sozialwissenschaften eine etablierte Methode, um wissenschaftliche Erkenntnisse aus den vielen Textquellen, die Politik und Medien liefern, zu generieren. Bereits vor der Entwicklung von IT- und EDV-Lösungen wurden quantitative Untersuchungen in diesem Bereich vollzogen. Als Beispiel dient hier die Untersuchung von Block, die mit ihrer Analyse „Freedom and Equality: Indicators of Political Change in Sweden, 1945-1975“ die politische Werteentwicklung in Schweden anhand einer diagnostischen Inhaltsanalyse der Leitartikel der fünf führenden schwedischen Tageszeitungen über 30 Jahre hinweg untersuchte (vgl. Block, 1995). Die fortschreitende Digitalisierung aller Bereiche des menschlichen Lebens hat durch neue digitale Kommunikationsformen und -foren die Datenlage im Bereich digitaler Texte mit sozialwissenschaftlicher Relevanz geändert. Politische Texte und Meinungsäußerungen finden sich auf allen Kanälen der Gesellschaft, unter unterschiedlich professionalisierten Urheber:innen. Zusätzlich hat sich durch die Digitalisierung aber auch die Entwicklung von informationswissenschaftlichen Werkzeugen, mit denen Texte inhaltlich zu untersucht und klassifiziert werden können, beschleunigt und damit auch die Schlagzahl an politikwissenschaftlichen Inhalts- und Sentiment-Analysen deutlich erhöht. Ein Softwareprodukt nutzen hierbei die Autoren der Studie „Election Forecasts With Twitter: How 140 Characters Reflect the Political Landscape“. Hierbei wird mit „Linguistic Inquiry and Word Count“ (LIWC2007), eine Textanalysesoftware, um Sentiments aus Tweets zu extrahieren und so zu untersuchen, untersucht, inwieweit politische Äußerungen in diesem sozialen Netzwerk die tatsächliche politische Landschaft wieder spiegeln (vgl. Tumasjan et al., 2011). Die Verwendung von Softwareprogrammen mit wenig Gestaltungsspielraum und einfachen Benutzeroberflächen wie LIWC2007 wurde im vergangenen Jahrzehnt schrittweise abgelöst von offenen Lösungen in Programmiersprachen. Im Fokus dieser Untersuchung steht die Analyse von politischen Texten mit der Hilfe von maschinellem Lernen, genauer mit Methoden im Bereich des Natural Language Processing (NLP). Als wichtigstes Beispiel für die Rahmenbedingungen dieser Analyse dient Python

in Kombination mit dem SpaCy- oder dem NLTK-Package, da erstere Kombination im Rahmen dieser Untersuchung verwendet wurde, um maschinelle Sprachverarbeitung durchführen zu können. Im Bereich der Verarbeitung von schriftlicher Sprache im Rahmen von politischer Kommunikation, die mit Methoden des maschinellen Lernens aus der Python-Umgebung realisiert wurden, gibt es eine Vielzahl Beispiele, die auch diese Untersuchung geprägt haben. Als Beispiel kann hier die Untersuchung „NLP Based Text Analytics and Visualization of Political Speeches“ (Katre, 2019) dienen, bei der mittels dem Selenium Webdriver und dem Natural Language Tool Kit (NLTK) politische Reden einer Topic Analysis unterzogen wurden, um so relevanten Themenbereich von zentralem Interesse auszumachen und zu visualisieren. Ähnlich zu der hier geplanten Analyse im Bereich politischer Podcasts war es auch Ziel bei der Studie von Katre aufzuzeigen, welches Potenzial Verfahren im Bereich NLP bieten, um große Mengen an politischen Sprachprodukte zu analysieren. Von besonderem Interesse ist auch der Fokus auf gesprochenen politischen Inhalt, der aber selbstverständlich bereits im Vorfeld der eigentlichen Veräußerung verschriftlicht wurde und deswegen dennoch eine andere Untersuchungsebene trifft, wie die Analyse von reinem gesprochen Inhalt. Die Analyse von Textdokumenten hinsichtlich ihrer (partei-)politischen Ausrichtung ist seit jeher von zentralem Interesse in den Politikwissenschaften, hat aber durch die Implementierung von Prozessen von Maschinellern Lernen einen deutlichen Aufschwung erhalten. Auf die dahingehende Konzeption und das Untersuchungsdesign dieser Studie hat das Projekt „Capstone Project 2: Twitter Political Compass Machine“ (Wu, 2020) Einfluss, bei dem ein Tool entwickelt wurde, um Tweets hinsichtlich ihrer Zugehörigkeit oder Zustimmung zu den beiden großen Parteien der US-Politik zuordnen zu können. Hierbei wurden zuerst verschiedene Modelle des Maschinellen Lernens hinsichtlich ihrer Leistungsfähigkeit beim Klassifizieren politischer Sprache verglichen. Das beste Modell wurde dann verwendet, um den tatsächlichen Klassifizierungstask durchzuführen. NLP-Methodiken, um Podcasts zu analysieren, sind bisher weniger oft im Forschungsdiskurs zu finden, was auch auf die relative Neuartigkeit dieser Medienform zurückgeführt werden kann. In „What are Ben Shapiro & Pod Save America Talking About? An NLP Analysis“ (Overfelt, 2020) wird mithilfe einer NLP-Pipeline eine große Zahl an Podcasts einer Textanalyse unterzogen, ohne Werkzeuge des Maschinellen Lernens zu benutzen.

Die Aufbereitung des Forschungsstands in den vorhergehenden Zeilen zeigt den Anspruch dieser Untersuchung: Viele Untersuchungsebenen und Methoden, die diese Analyse nutzen will, wurden und werden bereits in Forschungsprojekten verwendet und

untersucht. Dieses Projekt soll alle angesprochenen Dimensionen vereinen: Es soll das Medium Podcasts als Form gesprochener politischer Meinungsäußerung hinsichtlich darin enthaltener (partei-)politischer Orientierung und Tendenz mittels NLP-Methodiken untersucht werden.

3 „Lage der Nation“, „Spiegel Online“ und „FAZ.NET“ – Beispiele für Qualitätsjournalismus?

Die einleitend erörterte Fragestellung zur Neutralität von journalistischen Medienprodukten und das dazu benötigte Werkzeug zur politischen Einordnung derselbigen gilt als Grundstein des theoretischen Hintergrunds dieser Untersuchung.

Den beschriebenen, immer schneller produzierten, Medien- und Journalismus-Produkten steht ein ungebrochener Bedarf an neutralem Qualitätsjournalismus gegenüber (vgl. Sadrozinski, 2013: 84ff, 87; Patzelt, 2013: 141, 227ff, 422ff). Grundsätzlich definiert sich Qualitätsjournalismus durch mehrere Vorraussetzungen, neben der hier untersuchten Unparteilichkeit, sind unter anderem Sachlichkeit, Wahrhaftigkeit und Vertraulichkeit zu nennen (vgl. Schulzki-Haddouti et al., 2009: 7). Diese Faktoren greifen ineinander und bedingen sich oft gegenseitig. Zusätzlich zu diesen Voraussetzungen definiert sich Qualitätsjournalismus aber auch durch die gesellschaftlich-strukturellen Aufgaben, die im politikwissenschaftlichen Kontext an ihn gestellt wird. Medien und Journalismus müssen sich ihrer Wichtigkeit und Aufgabe als Kommunikationsakteur zwischen der Gesellschaft und dem zentralen politischen Entscheidungssystem bewusst sein (vgl. Patzelt, 2013: 141, 227ff, 422ff).

Diese gesellschaftswissenschaftlichen Voraussetzungen erfordern eine maschinell gestützte Untersuchung von politischen Texten und Medien, der Umfang und die damit verbundene Frequenz der Veröffentlichungen erfordern dies ebenfalls. Hierbei nutzt diese Analyse, die im Forschungsstand bereits erarbeiteten und in den letzten Jahren deutlich weiterentwickelten, Möglichkeiten des Natural Language Processing mithilfe von maschinellem Lernen. Durch diese ist es möglich, große Mengen an Text statistisch und semantisch auszuwerten, was Ziel dieser Untersuchung ist. Verstärkt wird die Notwendigkeit von maschinell gestützter Klassifizierung von Medienprodukten durch die hier untersuchte und relativ neue Form der medialen Berichterstattung als Podcasts. In der kommunikationswissenschaftlichen Debatte wird dabei vor allem Polit-Podcasts, wie das hier untersuchte Produkt, eine immer wichtigere Rolle im öffentlichen Diskurs attestiert, wohingegen die geisteswissenschaftliche Analyse dieses Mediums im Vergleich

ausbaufähig erscheint, was die bereits dargelegte Motivation dieser Analyse unterstreicht (vgl. Lührmann, 2019: 9f, 13).

Nachdem nun die kommunikations- und politikwissenschaftlichen Grundlagen dieser Untersuchung sowie die damit verbundene Relevanz von NLP-Methoden erläutert wurden, sollen nun noch kurz die drei Untersuchungsobjekte und Vergleichsparameter vorgestellt werden. Besonders im Fokus steht dabei die politische Einordnung der drei journalistischen Produkte.

Begonnen wird mit dem Hauptaugenmerk dieser Untersuchung, dem Podcast „Lage der Nation“ des Juristen Ulf Buermeyer und des Journalisten Philip Banse. In wöchentlicher Ausgabe diskutieren Buermeyer und Banse in ihrem Podcast-Format „Lage der Nation“ politische und gesellschaftliche Prozesse in Deutschland und der Welt. Mit mehr als 200.000 Downloads pro Folge gehört der seit 2016 produzierte Podcast zu den meistgehörten seiner Art in Deutschland und ist mit mehreren Preisen ausgezeichnet worden, unter anderem als „Bester Polit-Podcast 2016“ durch iTunes und als „Bester Podcast des Jahres 2017“ beim Wettbewerb „Die goldenen Blogger“ (vgl. Barczok, 2019; Ostermaier 2016). Neben diesen Belegen für die Qualität betonen die beiden Podcaster auch in den Folgen selbst immer wieder ihre journalistischen Prinzipien, auf die sich Banse auch in einem Interview beruft und dabei seinen journalistischen Ursprung im „traditionellen, öffentlich-rechtlichen Rundfunk“ (Barczok, 2019) betont (vgl. Barczok, 2019). Dabei benennen die beiden Podcaster vor allem die Berichtigung von Fehlern, die Verifizierung von Inhalten und das Vermeiden von Gerüchten als Prinzipien ihrer Berichterstattung (vgl. Barczok, 2019). Der Podcast finanziert sich durch vier Standbeine. Zunächst gibt es während den Folgen bis zu zwei Werbeblöcke und eine werbefreie Premiumversion des Podcast. Weiter werden Einnahmen durch Spenden und gelegentliche Live-Auftritte mit Live-Aufnahmen einer Folge generiert. Zusammenfassend kann „Die Lage der Nation“ also als qualitätsjournalistisches Produkt bezeichnet werden und ist deswegen primärer Untersuchungsgegenstand. Mit Blick auf die nachfolgende Analyse wird erwartet, dass die „Lage der Nation“ einen neutral-kritischen Blick auf die politische Landschaft in Deutschland hat und alle politischen Strömungen in etwa gleich oft erörtert und entsprechend deren politischen Leistungen und Verfehlungen bewertet. Grundsätzlich ist die „Lage der Nation“ mit Blick auf die politische Landschaft der Bundesrepublik noch nicht abschließend eingeordnet, dies stellt eines der Teilziele dieser Analyse dar. Es wird aber von einer politisch neutralen und fakten-basierten Ausrichtung ausgegangen und mit Blick auf den juristischen und öffentlich-rechtlichen Hintergrund des beiden Podcaster

eine politische mittig-zentrierte Ausrichtung erwartet. Eine politische Einordnung des Podcast wird erschwert durch die institutionelle Unabhängigkeit des Formats, dass im Gegensatz zu vielen ähnlichen Produkten nicht an einen etablierten Verlag oder ein bestehendes Medium mit entsprechender bundespolitischer Orientierung und Vergangenheit gebunden ist. Anders ist die bei den beiden Vergleichsparametern, die als Verifizierungsbenchmarks dieser Untersuchung dienen und nachfolgend ebenfalls eingeführt werden.

Als erstes Vergleichsparameter in der Analyse der politischen Einordnung von Medienprodukten dienen Online-Artikel des deutschlandweit bekannten und etablierten Spiegel-Verlags. Die „Spiegel-Gruppe“ ist ein deutscher Medienkonzern, gegründet 1969 durch die Rudolf Augstein GmbH, die seit 1947 das Hauptprodukt des Konzerns, das Spiegelmagazin, vermarktet (vgl. Spiegel Gruppe, 2016). Heute produziert der Medienkonzern neben dem erwähnten Ursprungsprodukt eine Vielzahl an Medienprodukten, darunter verschiedenste Fachmagazine, einen TV-Sender und nicht zuletzt ein breites Angebot an Online-Nachrichten. Letzteres dient als Untersuchungs- und Datenbasis für diese Analyse (vgl. Spiegel Gruppe, 2016). 2021 erwirtschaftete der Gesamtkonzern einen Umsatz von 274,9 Millionen Euro Umsatz und gehört dazu zu den finanzstärksten deutschen Medienhäusern (vgl. Spiegel Gruppe, o.D.). Aber auch mit Blick auf das journalistische Profil gehört der Spiegelverlag zu den etabliertesten deutschen Nachrichtenverlagen, das zeigt schon die Auflagenzahl des wöchentlichen Hauptmagazins „Der Spiegel“ mit rund 680.000 verkauften Exemplaren im Jahr 2021 (vgl. PZ-online, 2022). Noch relevanter ist hierbei auch die vom Spiegel-Konzern veröffentlichte kanalübergreifende Reichweite von wöchentlich 15 Millionen Menschen im Jahr 2021 (vgl. Spiegel Gruppe, o.D.). Zudem zeugen auch die vielen internationalen redaktionellen Standorte der Spiegelgruppe und die sogenannten Spiegelstandards als konzerneigene Leitlinien für qualitätsjournalistische Arbeit vom qualitativen Anspruch des Verlags (vgl. Spiegel Gruppe, 2021). Auch das Online-Nachrichtenportal der Spiegel-Gruppe „Spiegel Online“ genannt, hat in kommunikationswissenschaftlichen Kreisen schon früh eine hochwertige Reputation inne (vgl. Neuberger, 2006: 1). Dafür sprechen auch verschiedene journalistische Preise, die das Portal gewinnen konnte, unter anderem den „Grimme Online Award“ 2005 (vgl. Neuberger, 2006: 2). Zudem wird dem Nachrichtenportal vor allem eine große Stäke im Bereich der schnellen und aktuellen Berichterstattung konstatiert, was die Relevanz des Portals für die Forschungsfrage dieser Untersuchung unterstreicht (vgl. Bönisch, 2006: 5ff). Bezeichnend für die journalistische Qualität des Spiegels ist auch dessen Einstufung als deutsches Leitmedium in der Kommunikationsforschung, auch wenn sich die Leitrollen

der deutschen Medien immer stärker diversifizieren (vgl.: Bönisch, 2006, 89ff). Politisch wird sowohl der Spiegel-Verlag als auch dessen Wochenmagazin „Der Spiegel“ im politik- und kommunikationswissenschaftlichen Kontext links der Mitte im progressive-gemäßigtem Raum eingeordnet (Pew Research Center, 2018; Medienkompass, 2023). Diese politische Färbung des Verlags reicht zurück bis in die Zeit der Gründung durch Rudolf Augstein, der politisch ähnlich bewertet wird (Norddeutscher Rundfunk, 2022). Besonders deutlich wird dies durch die bekannte Spiegel-Affäre 1962, bei der der Verlag Konfrontationen aus dem konservativen Spektrum der Bundespolitik ausgesetzt war (vgl. Spiegel Gruppe, 2016; Pöttker, 2012).

Als zweites Vergleichsparameter wurden Online-Artikel der überregionalen Tageszeitung „Frankfurter Allgemeine Zeitung“ (FAZ) ausgewählt. Die in der Eigenschreibweise „Frankfurter Allgemeine“ genannte Zeitung wird von der „Frankfurter Allgemeine Zeitung GmbH“ produziert. 1949 gegründet, publiziert der Verlag heute für eine Vielzahl von journalistischen Produkten, von Webnachrichten, über Fachzeitungen bis hin zu Podcasts (vgl. Frankfurter Allgemeine, o.D.). Seit 2001 veröffentlicht der Konzern außerdem Nachrichten auf dem hauseigenen Online-Nachrichtenportal „FAZ.NET“ (vgl. Frankfurter Allgemeine, o.D.). Diese Artikel stellen die textuelle Untersuchungsbasis dieser Analyse dar. Mit Umsatzerlösen im Geschäftsjahr 2021 von 233 Millionen Euro, einer täglichen Auflage von 189.948 Auflagen (Stand 4. Quartal 2022) und 940.000 Lesern pro Ausgabe gehört die „FAZ GmbH“ zu den einflussreichsten Medienhäusern in Deutschland (vgl. Springer, 2022; IVW, 2023; Bundesanzeiger, 2022). Diesen Ruf hat der Konzern mit 12,04 Millionen monatlichen Besucher:innen (Stand 01/2023) des hauseigenen Onlineportals FAZ.NET auch im digitalen Raum verfestigt (vgl. ag.ma, 2023). Ähnlich wie der Spiegel-Verlag verfügt die FAZ, um dem Anspruch von internationalem Journalismus und der Vor-Ort-Reportage gerecht zu werden, über eine Vielzahl an internationalen Standorten und Korrespondenzen (Frankfurter Allgemeine, o.D.). Mit Blick auf die journalistische Reputation wird die FAZ als eines der Leitmedien in Deutschland eingestuft und gilt als eine der qualitativ hochwertigsten Tageszeitungen (vgl. Hoeres, 2014). Im politischen Spektrum wird die FAZ im bürgerlich-konservativen Milieu verortet, allerdings wird ihr auch zugesprochen eine gewisse Entwicklung hin zu einem politisch heterogeneren Medium gemacht zu haben (vgl. Hoeres, 2019: 376f, 423-430). Insbesondere hat sich auch die politische Verortung der im Verlag tätigen Journalist:innen diversifiziert (vgl. Hoeres, 2019: 376f, 423-430).. Im Vergleich zum Spiegelverlag ist die FAZ

dennoch konservativer und weniger progressiv einzuordnen (Pew Research Center, 2018; Medienkompass, 2023).

Zusammenfassend besteht die theoretische Motivation dieser Analyse aus einer Trias aus neuen journalistischen Produkten mit neuen Eigenschaften, neuen technischen Analyse-möglichkeiten durch maschinelles Lernen und der politischen Einordnung von Medien. Daraus ergeben sich zwei Fragestellungen, die im nachfolgenden Analyseabschnitt untersucht werden sollen:

1. Ist es möglich mithilfe von Maschinellern eine Methode zu entwickeln, die etablierte journalistische Produkte hinsichtlich ihrer politischen Einordnung und Neutralität klassifizieren kann?
2. Kann diese Methode den noch nicht voll umfänglich politisch eingeordneten Podcast „Lage der Nation“ als neue Form der politischen Berichterstattung auf der politischen Karte der Bundesrepublik verorten?

4 Methodik und Technische Umsetzung

Die Neutralität, beziehungsweise die politische Ausrichtung, des hier beispielhaft verwendeten Podcasts „Lage der Nation“ soll durch Named-Entity-Recognition (NER) und einer damit verbundenen Sentiment-Analyse untersucht werden. Als Datenbasis dienen die Episoden 260 – 326 des Podcasts, die von Podcaster Philip Banse in Transkriptform veröffentlicht wurden (vgl. Banse 2023, Datenquellenverzeichnis¹). Als Untersuchungszeitraum wurde dabei die Zeitspanne zwischen der Bundestagswahl im September 2021 und März 2023 gewählt, daraus ergibt sich eine Summe von 66 Folgen, die zur Untersuchung verwendet wurden. Durch NER sollen Passagen im Verlauf der einzelnen Podcastfolgen gefunden werden, in denen entweder eine der derzeit im deutschen Bundestag vertretenen Parteien erwähnt wird oder für die Parteien übliche, offizielle oder umgangssprachlich verwendete Bezeichnungen genannt werden. Zudem werden die Namen von bekannten Politiker:innen der jeweiligen Parteien lokalisiert. Als bekannte Politiker:innen wurden dabei die jeweiligen Parteivorstände, Generalsekretär:innen, Mitglieder der Bundesregierung, Ministerpräsident:innen, Regierende Bürgermeister:innen von Berlin, Bremen und Hamburg, Fraktionsvorsitzende, Bundestagspräsident:innen sowie der Bundespräsident gewählt. Die so definierte Entitätsliste findet sich im Anhang (vgl. Anhang 4).

¹ Die für diese Analyse verwendeten Datenquellen werden gesondert in einem eigenen Datenquellenverzeichnis gelistet.

Aufgrund der Notwendigkeit dieser Entitätsliste wurde auch der oben beschriebene Zeitraum der Analyse gewählt: Da sich naturgemäß nach der Bundestagswahl die Konstellation der Bundespolitik stark verändert hat, was Regierung, Opposition und maßgebliche Politiker:innen betrifft, wäre für die Untersuchung vor der Bundestagswahl eine andere Pattern-Architektur vonnöten gewesen. Durch die Dauer des Untersuchungszeitraums bis April 2023 soll die Aktualität der Untersuchungsbasis mit Stand des Abschlusses der Analyse im Vergleich zu den nachfolgend beschriebenen Vergleichsparametern gewährleistet werden. Um eine übersichtliche Auswertung zu gewährleisten, wurden die sehr kleinteilig definierten Entitätslisten nach der Sentiment-Analyse wieder aufgelöst und die gefunden einzelnen Entitäten den sechs im Bundestag vertretenen Parteien zugeordnet.

Die Sätze, die die so gefundenen Stellen umfassen, sollen dann anhand einer Sentiment-Analyse mit einem zuvor trainierten spezifischem BERT-Modell untersucht werden, um bestmögliche Ergebnisse bei der Klassifizierung der einzelnen Textpassagen hinsichtlich ihres Sentiments zu erzielen. So soll mittels einer Frequenzanalyse ermittelt werden, wie oft die einzelnen Parteien im Podcast negativ oder positiv bewertet werden und ein Vergleich zwischen den einzelnen politischen Gruppierungen gezogen werden.

Analog dazu sollen auch die beiden Vergleichsparameter, der Spiegelverlag und der Weltverlag untersucht werden. Dies dient dazu, die Ergebnisse der oben beschriebenen Analysemethode signifikant zu verifizieren. Wenn das entworfene Einstufungstool richtig arbeitet, sollte vorausgesetzt werden, dass es auch die journalistischen Produkte von Spiegel und FAZ entsprechend ihres oben beschriebenen politischen Profils einstufen kann. Es wird also überprüft, ob die Parteien und Politiker:innen der deutschen Parteienlandschaft im Spiegel und in der FAZ so bewertet werden, wie zu erwarten ist: kritisch-hinterfragend, weitgehend neutral mit leichter konservativer beziehungsweise progressiver Färbung.

Als Untersuchungsbasis dienen hier eine Vielzahl von Online-Artikeln der beiden Medienhäusern zu den verschiedensten Themen. Diese Textdatenbasis wurde durch Webscraping mit den Python-Libraries BeautifulSoup und Request generiert. Anschließend wurden dann ebenfalls zunächst mit NER Textpassagen lokalisiert, die entweder eine der sechs im Bundestag vertretenen Parteien oder prominente Politiker:innen der Parteien enthalten. Nachfolgend wurden die so gefundenen Sätze dann mittels der Methodik der oben beschriebenen Sentiment-Analyse hinsichtlich ihrer Stimmung zu den einzelnen politischen Strömungen untersucht und einer Häufigkeitsanalyse unterzogen.

Sowohl beim Hauptuntersuchungsgegenstand als auch bei beiden Vergleichsparametern wurde für die Sentiment-Analyse ein Klassifikationsmodell aus dem Bereich Deep Learning verwendet. Zunächst wurden drei verschiedene Klassifikationsmodelle anhand eines bereits annotierten Datensatzes trainiert und getestet, das Klassifikationsmodell mit den besten Ergebnissen wurde dann für die eigentliche Analyse des Datensatzes des Untersuchungsgegenstands verwendet. Konkret wurden die Modelle „bert-base-multilingual-cased“ (Modell 1)², „oliverguhr/german-sentiment-bert“ (Modell 2)³ und „citizenlab/twitter-xlm-roberta-base-sentiment-finetuned“ (Modell 3)⁴, gewählt. Alle drei Modelle stammen aus der Bidirectional-Encoder-Representations-from-Transformers-Familie (BERT), einer Gruppe von Sprachmodellen auf Basis von Neuronalen Netzen, die von einem Google-Forschungsteam entwickelt wurden. Modell 1 kann als Basis-Modell der BERT-Familie verstanden werden und ist die multilinguale Erweiterung des Stammmodells BERT, dass die Entwickler 2018 veröffentlichten und kann so für die hier untersuchte deutsche Datenbasis verwendet werden (vgl. Devlin et al., 2018). Modell 1 wurde gewählt, um grundsätzlich evaluieren zu können, wie auf die BERT-Architektur basierende Deep-Learning-Modelle politische Sprache verarbeiten können. Modell 2 ist ein solches Modell, das entwickelt wurde, um Stimmung und Emotion in deutscher Sprache verstehen zu können und wurde aufgrund der guten grundsätzlichen Performance und der hohen Überschneidung mit der Aufgabenstruktur dieser Analyse gewählt. Zudem wurde Modell 2 unter anderem analog zum hier verwendeten Trainingsdatensatz anhand von Twitter-Daten trainiert, woraus ein besonderer Outputvorteil für die hier durchgeführte Untersuchung erwartet wurde (vgl. Guhr et al., 2020). Aus denselben Gründen wurde auch das dritte Modell gewählt, auch hier war eine hohe Taskübereinstimmung zu erkennen, das Modell wurde ebenfalls trainiert, um Text in drei Sentiment-Kategorien klassifizieren zu können. Allerdings basiert dieses Modell auf der RoBERTa-Architektur, einer Weiterentwicklung von BERT. RoBERTa erzielt aufgrund einer umfangreicheren und länger trainierten Datenmenge bessere Ergebnisse. Zudem wurden andere Hyperparameter beim Training verwendet (vgl. Liu et al., 2019: 6ff).

Im Detail wurden die Modelle aus dem Huggingface-Hub einem Finetuning-Prozess verbunden mit Hyperparametertraining unterzogen. Dafür würde ein bereits annotierter Datensatz verwendet, der über 1700 Tweets von deutschen Politiker:innen enthält. Die

² <https://huggingface.co/citizenlab/twitter-xlm-roberta-base-sentiment-finetuned>

³ <https://huggingface.co/oliverguhr/german-sentiment-bert>

⁴ <https://huggingface.co/citizenlab/twitter-xlm-roberta-base-sentiment-finetuned>

Tweets wurden dabei in die Kategorien positiv, neutral und negativ klassifiziert – in diese Kategorien wurden dann auch die Aussagen zur Parteipolitik in der nachfolgend ausgewerteten Sentiment-Analyse kategorisiert (vgl. Weissenbacher, 2021). Der Datensatz wurde aufgrund der darin klassifizierten politischen Sprache gewählt, um die Modelle zielgenau für die hier angestrebte Aufgabe der Verarbeitung politischer Sprache trainieren zu könne. Nachdem die drei Modelle anhand des Datensatzes trainiert wurden, wurde die Modellgüte anhand eines Testdatensatzes bestimmt. Dazu wurde bereits vor dem Modelltraining ein Teil der Daten mit der klassischen Train-Test-Split-Methode abgespalten und nicht für den Deep-Learning-Prozess verwendet. Konkret wurden 20% des Stammdatensatzes verwendet, um den Testdatensatz zu erstellen, die restlichen 80% bilden den Trainingsdatensatz. Nach Abschluss des Trainings wurde das Modell und seine Vorhersagekraft dann anhand des Test-Datensatzes evaluiert, verwendet wurden dazu die klassischen Evaluierungsmetriken Precision, Recall, Accuracy und der F1-Score. Anhand dieser Metriken wurden dann entschieden, welches zuvor für die hier angestrebte Aufgabe nun speziell trainierte Modell im Anschluss für die Sentiment-Analyse verwendet werden soll. Die genauen Maßnahmen des Finetunings und Hyperparametertrainings werden in Verbund mit den daraus resultierenden Ergebnissen im Analyseteil dargestellt.

5 Modelltraining und Sentiment-Analyse: Auswertung

5.1 Modellauswahl, Finetuning und Hyperparametertraining

Das in der Methodik beschriebene Finetuning und Hyperparametertraining liefert ein ambivalentes Ergebnis. Um einerseits die Vergleichbarkeit der Evaluierungsergebnisse der Trainingsprozesse der drei Modelle gewährleisten zu können, wurde durch den Parameter „random_state“ gewährleistet, dass beim Split-Vorgang dieselben Trainings- und Testdaten sowie dieselben Initialisierungsparameter ausgewählt und bei jedem Modell wurden. Um den Trainingsprozesse andererseits repräsentativ zu gestalten, wurde zudem die Methode der Kreuzvalidierung angewandt, bei der die Modelle mehrere Male mit verschiedenen Trainings- und Testdaten trainiert und bewertet wurden. So konnte sichergestellt werden, dass die Ergebnisse im Finetuning-Prozess nicht zufällig auf die Datenaufteilung zurückzuführen sind, sondern das trainierte Modelle stabile und datenunabhängige Klassifikationsergebnisse liefert. Diese beiden Maßnahmen mit dem Ziel der Vergleichbarkeit und der Repräsentativität zeigten Wirkung, über den gesamten Trainingsprozess hinweg lagen die Evaluierungsergebnisse der Modelle auf einem ähnlichen Niveau und konnten

so datenunabhängig als stabil eingeschätzt werden. Bei Modell 3 muss jedoch angemerkt werden, dass die Durchführung der Kreuzvalidierung gezeigt hat, dass die Trainingsergebnisse hier stärkeren Schwankungen unterworfen sind als bei den beiden anderen Modellen, diese Erkenntnis hat sich auf den nachfolgend beschriebenen Entscheidungsprozess für ein Modell ausgewirkt.

Grundsätzlich lieferten die drei in den Trainingsprozess integrierten Modelle ähnliche Ergebnisse (vgl. Abb. 2). In allen Evaluierungsmetriken erreichten die trainierten Modelle Werte zwischen 0,64 und 0,74 (vgl. Abb. 3, Anhang 3). Die Ergebnisse wurden durch wiederholte und kleinstufige Anpassungen der Hyperparameter Batch-Size, Epochenanzahl und Lernrate feinjustiert, blieben jedoch über den gesamten Trainingsprozess auf einem ähnlichen Level. Allerdings konnten die Ergebnisse dadurch dennoch in kleinem Umfang verbessert werden, so konnten unter anderem der F1-Score für alle Modelle um bis zu 0,06 Einheiten verbessert werden. Das ähnliche Niveau der Modelle schlägt sich auch in den Vorhersagekapazitäten der Modelle nieder, wie in den nachfolgenden Confusion-Matrizen erkennbar ist (vgl. Abb. 1).

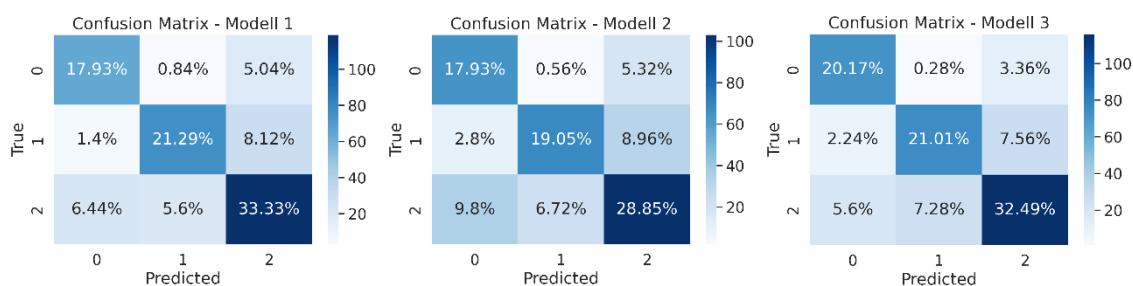


Abbildung 1: Confusion Matrix Modelltraining, $n = 357$ (0 = positiv, 1 = negativ, 2 = neutral)

Die bisher beschriebene Qualität muss durch ein bekanntes Phänomen im Machine-Learning-Prozess eingeschränkt werden, dass auch im hier dokumentierten Finetuning-Prozess zu Tage tritt. Bei allen Modellevaluierung wurde deutlich, dass der Training Loss des Modells mitunter schon nach wenigen Epochen deutlich abfiel, während der Validation Loss oft deutlich anstieg. Als exemplarisches Beispiel hierfür dient die Entwicklung der beiden Werte über die Trainingsepochen beim Training mit Modell 1, wie in Abbildung 1 zu sehen ist (vgl. Abb. 2). Dies deutet auf Overfitting hin, was bedeutet, dass Machine-Learning-Modell so stark auf die Trainingsdaten angepasst wird, dass es für neue, unbekannte Daten schlechte Vorhersagen trifft. Das Modell ist im Wesentlichen zu komplex, zu gut an die Trainingsdaten angepasst worden und enthält zu vielen unabhängige Variablen, was dazu führt, dass es Muster und Zusammenhänge lernt, die spezifisch

für die Trainingsdaten sind, aber nicht für andere Daten gelten (vgl. Backhaus et al., 2015: 94; IBM 2023).

Als Effekt davon leidet die Genauigkeit der Vorhersagefähigkeit der Modelle für noch unbekannte Daten, was sich auf die Signifikanz der in dieser Analyse auswirkt und bei der nachfolgenden Einordnung der Ergebnisse beachtet werden muss. Da das Phänomen Overfitting bei allen Modellen auftrat, zog dies keine daraus basierende Auswirkung auf den Modellentscheidungsprozess mit sich, welcher sich aufgrund der sehr ähnlichen Evaluierungsergebnisse besonders detailliert geführt wurde. Vor allem Modell 1 und 3 liegen hier sehr dicht aneinander, Modell 2 fällt etwas zurück (vgl. Abb. 3).

Schlussendlich wurde die Entscheidung für Modell 1 getroffen. Entscheidender Faktor hierfür war bei bereits identifizierten annähernd gleichen Evaluierungsmetriken die bei der Kreuzvalidierung von Modell 3 festgestellte Schwankung der Ergebnisse bei unterschiedlich randomisierter Auswahl von Test- und Trainingsdate. Allerdings zeigt sich mit Blick auf Training und Validation Loss der beiden Modelle auch, dass letzterer bei Modell weniger stark steigt, was ein Indiz für ein schwächer ausgeprägtes Overfitting sein kann, so dass auch Argumente für Modell 3 nicht verschwiegen werden können. Allerdings wurde im Entscheidungsprozess die Gewährleistung eines stabilen und ausreißerlosen Modelloutputs höher gewichtet als ein marginal schwächeres Overfitting.

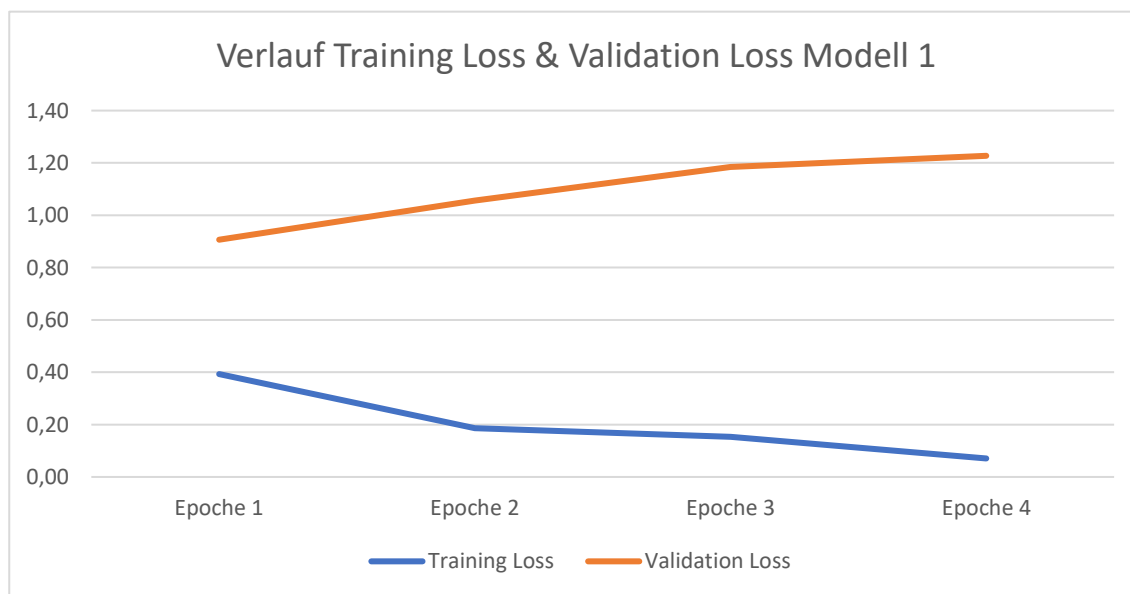


Abbildung 2: Verlauf Training Loss & Validation Loss Modell 1 ("bert-base-multilingual-cased")

Modell 1		Epoch	Training Loss	Validation Loss	Precision	Recall	Acc	F1
Batch Size	32	1	0.393100	0.906605	0.730946	0.734308	0.731092	0.731842
Epochenanzahl	4	2	0.185900	1.058602	0.741454	0.736427	0.739496	0.738048
Lernrate	3,00E-05	3	0.153200	1.185098	0.722336	0.712964	0.714286	0.713776
Testgröße	0,20	4	0.070700	1.227207	0.726732	0.726139	0.725490	0.725349
Modell 2		Epoch	Training Loss	Validation Loss	Precision	Recall	Acc	F1
Batch Size	32	1	0.308400	1.070258	0.653009	0.661310	0.647059	0.645865
Epochenanzahl	4	2	0.216700	1.117317	0.705587	0.694945	0.697479	0.697355
Lernrate	3,00E-05	3	0.130200	1.518438	0.672668	0.681127	0.672269	0.673332
Testgröße	0,20	4	0.119000	1.539757	0.659797	0.668975	0.658263	0.659453
Modell 3		Epoch	Training Loss	Validation Loss	Precision	Recall	Acc	F1
Batch Size	32	1	0.392400	0.735043	0.739515	0.734176	0.733894	0.736480
Epochenanzahl	4	2	0.221900	0.896989	0.733185	0.755087	0.733894	0.740638
Lernrate	2,00E-05	3	0.133200	1.059207	0.719472	0.742721	0.722689	0.724188
Testgröße	0,20	4	0.111300	1.056536	0.722805	0.736350	0.725490	0.727500

Abbildung 3: Evaluierungsmetriken Trainingsmodelle ("bert-base-multilingual-cased", „oliverguhr/german-sentiment-bert“, „citizenlab/twitter-xlm-roberta-base-sentiment-finetuned“)

Möglichkeiten und Maßnahmen, um den Trainings- und Finetuning-Prozess noch zu verbessern und Overfitting zu vermeiden werden im, der Analyse folgenden, Ausblick erörtert.

5.2 Evaluierung Sentiment-Analyse: „Lage der Nation“

Wie in der Evaluierung des Trainingsprozesses der Modelle bereits aufgezeigt, sind die Ergebnisse der mit Modell 1 durchgeführten Sentiment-Analysen einer gewissen Unsicherheit unterworfen. Abbildung 1 zeigt, dass falsche Klassifizierungen nicht ausgeschlossen werden können und sind so sehr wahrscheinlich auch bei den Ergebnissen der Sentiment-Analyse zu finden.

Mit Blick auf die entitätsunabhängige Verteilung der Sentiment-Label bestätigt sich die theoretisch formulierte Annahme, dass die „Lage der Nation“ die Akteure der deutschen Bundespolitik neutral-kritisch bewertet und ich in ihrer Rolle des kritisch-konstruktiven Journalismus vor allem darauf fokussiert Missstände und Verbesserungspotenziale zu verbessern. Das belegt die hohe Anzahl an neutral und negativ bewerteten Aussagen und Textstellen (vgl. Abb. 4).

Sentiment-Label	Häufigkeit absolut	Häufigkeit relativ
positiv	54	6,12%
negativ	277	31,37%
neutral	552	62,51%
Gesamt	883	100,00%

Abbildung 4: Entitätsunabhängige Sentiment-Label-Verteilung LAGE DER NATION, n = 883

Die aus der allgemeinen Sentiment-Verteilung (vgl. Abb. 4) gewonnene Erkenntnis einer kritische-neutralen Berichterstattung der „Lage der Nation“ bestätigt sich auch mit Blick auf die Verteilung der Sentiment-Label nach Parteien (vgl. Abb. 5 & 6). Keine der Parteien wird in auffälligem Maß besonders oft positiv oder negativ erwähnt. Grundsätzlich werden alle politischen Strömungen am häufigsten neutral in die Berichterstattung und Diskussion des Podcasts aufgenommen. Etwas aus der Reihe fallen hier die Alternative für Deutschland, die besonders oft negativ konnotiert wird (41,27% aller AFD-Nennungen im Sample) und die Linke, die grundsätzlich verschwindend gering im Podcast erwähnt wird und in der „Lage der Nation“ eine untergeordnete Rolle zu spielen scheint. Am häufigsten wird im Podcast dagegen über die FDP und die SPD gesprochen, dies könnte auf die Regierungsbeteiligung der beiden Parteien zurückzuführen sein. Diese Beobachtungen sind für sich genommen bemerkenswert, können der „Lage der Nation“ aber keine parteipolitische Färbung assistieren. Der Polit-Podcast spricht zwar über manche Parteien weniger oft und scheint sich vor allem mit der Regierungspolitik und der größten Oppositionspartei zu beschäftigen, eine Ausrichtung im politischen Spektrum der Bundesrepublik lässt sich daraus nicht ableiten. So bestätigt sich die zuvor theoretisch begründete Annahme einer mittig-zentrierten und vor allem neutralen Ausrichtung bestätigen, die „Lage der Nation“ beschäftigt sich vor allem mit den Parteien der politischen Mitte und zeigt weniger Interesse für die an den gegenüberliegenden Rändern positionierte AFD und Linke.

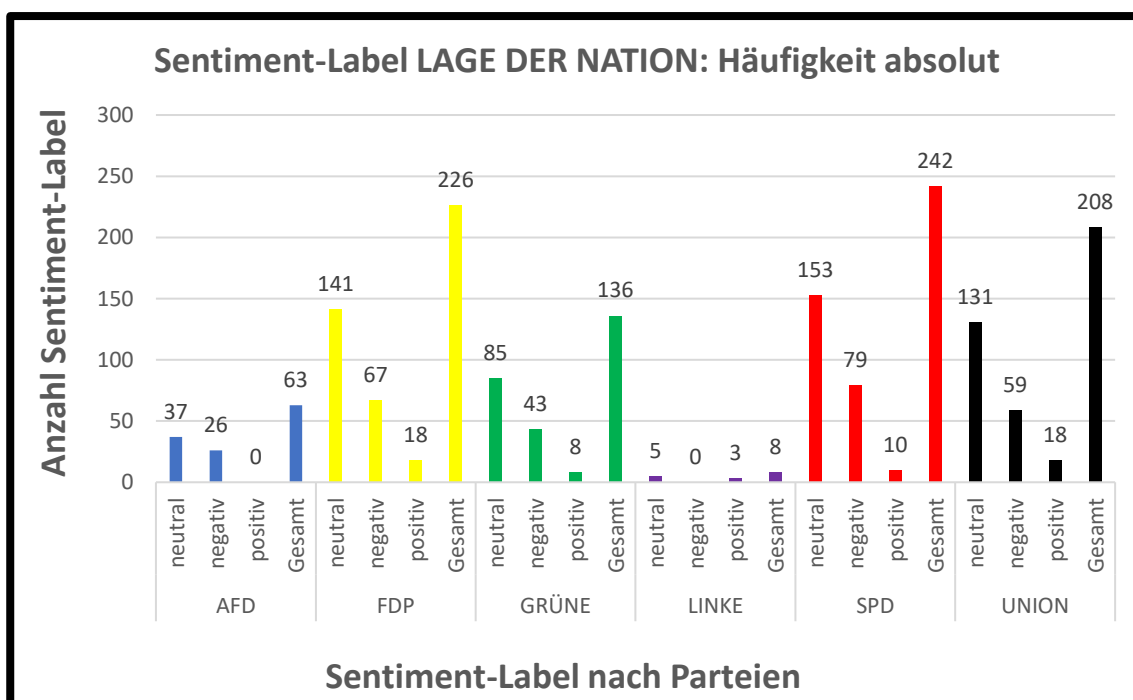


Abbildung 5: Sentiment-Label-Verteilung „Lage der Nation“ Häufigkeit absolut, $n = 833$, Balken Gesamt entsprechen Anteil Parteien am Gesamtsample Lage der Nation

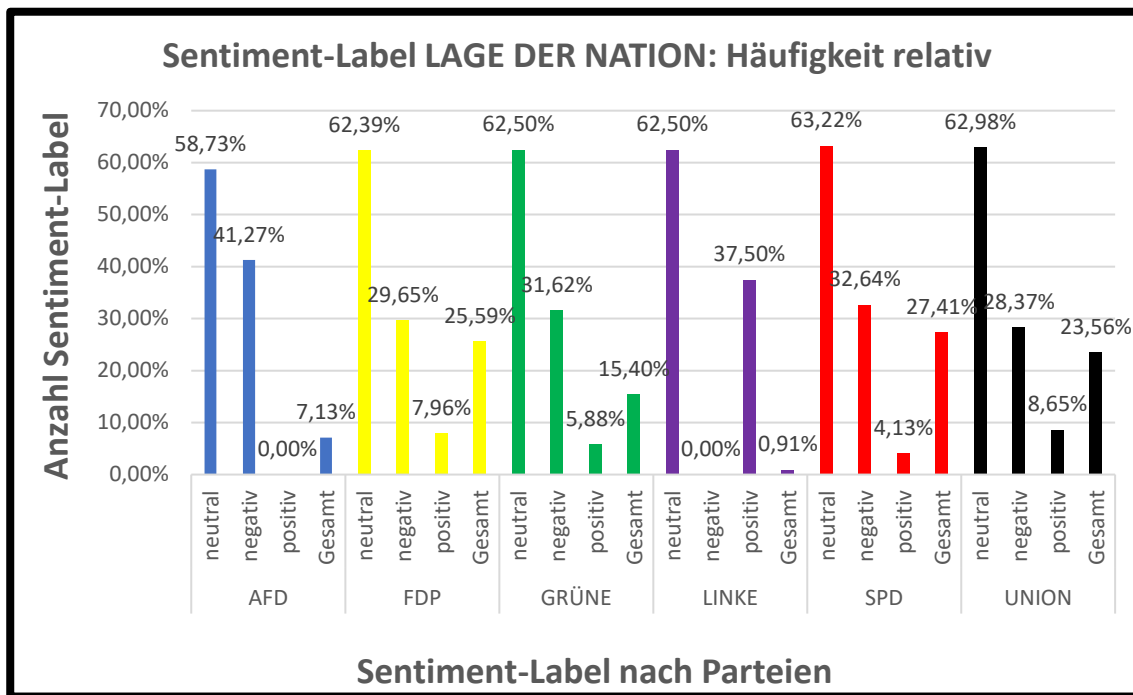


Abbildung 6: Sentiment-Label-Verteilung „Lage der Nation“ Häufigkeit relativ, n = 833, Balken Gesamt entsprechen Anteil Parteien am Gesamtsample Lage der Nation

5.3 Evaluierung Sentiment-Analyse: „Spiegel Online“

Durch die Analyse der entitätsunabhängigen Verteilung der Sentiment-Label verfestigt sich der theoretisch erarbeitete und recherchierte journalistische Qualitäts- und Neutralitätsanspruch des Spiegels. 62% der im Sample enthaltenen Textstellen aus der Berichterstattung des Online-Portals des Medienhauses wurden als neutral klassifiziert (vgl. Abb. 7). Durch die sehr niedrige Häufigkeit positiv klassifizierter Sätze zeigt sich zudem ein kritisch-neutraler Charakter, der auch auf die vorrangig berichtende und weniger wertende und einordnende Arbeitsweise des Onlineportal zurückzuführen ist.

Sentiment-Label	Häufigkeit absolut	Häufigkeit relativ
positiv	31	2,79%
negativ	388	34,95%
neutral	691	62,25%
Gesamt	1110	100,00%

Abbildung 7: Entitätsunabhängige Sentiment-Label-Verteilung SPIEGEL ONLINE, n = 1110

Erneut bestätigt sich die aus der Analyse der parteiungebundenen Verteilung der Sentiment-Label gewonnen Erkenntnis der vorrangig neutralen Berichterstattung mit Blick auf die Klassifizierung nach parteipolitischer Zuordnung (vgl. Abb. 8 & 9). Im Sample ist bei allen Parteien die Gruppe der als neutral eingestuften Textstellen am häufigsten vertreten. Auffällig ist hierbei aber, dass die FDP sehr häufig in negativem Kontext erwähnt wird,

fast jede zweite analysierte Textpassage wurde als negativ klassifiziert. Besonders oft werden erneut die Regierungsparteien der Ampelkoalition in die Berichterstattung mitaufgenommen, die Union fällt als Oppositionsführer hier zurück. Mit zusammen etwas weniger als 8% der Nennungen spielen die AFD und Linke eine untergeordnete Rolle, werden aber in etwa gleichem Verhältnis wie die anderen Parteien positiv, negativ und neutral in die Berichterstattung von „Spiegel Online“ integriert.

Die theoretisch dargelegte und im medienpolitischen Diskurs der Bundesrepublik verfestigte, leichte links-progressive Ausrichtung des Spiegelverlags kann anhand der hier generierten Datenanalyse nicht bestätigt werden. Zwar zeigen sich mit den beschriebenen, auffällig häufig negativ gefärbten Nennungen der FDP und die rein neutralen und negativen Berichterstattung der rechtspopulistischen AFD sowie die etwas öfter gefundene positive Schreibweise über die links der Mitte angesiedelten Parteien SPD, Grüne und Linke vereinzelt Indizien für eine progressiv und leicht linke Orientierung, allerdings werden diese durch die annähernd gleich gearteten Erwähnung von SPD, Grünen und der Union im neutralen und negativen Segment entkräftigt. Zusammenfassend sind Tendenzen einer leicht progressiven Ausrichtung also zu finden, können aber auch mit Blick auf die methodisch-technisch begründete Unschärfe des Klassifizierungsmodells nicht bestätigt werden. Der Spiegel-Verlag berichtet in seinem Online-Portal weitgehend neutral und ohne parteipolitische Färbung.

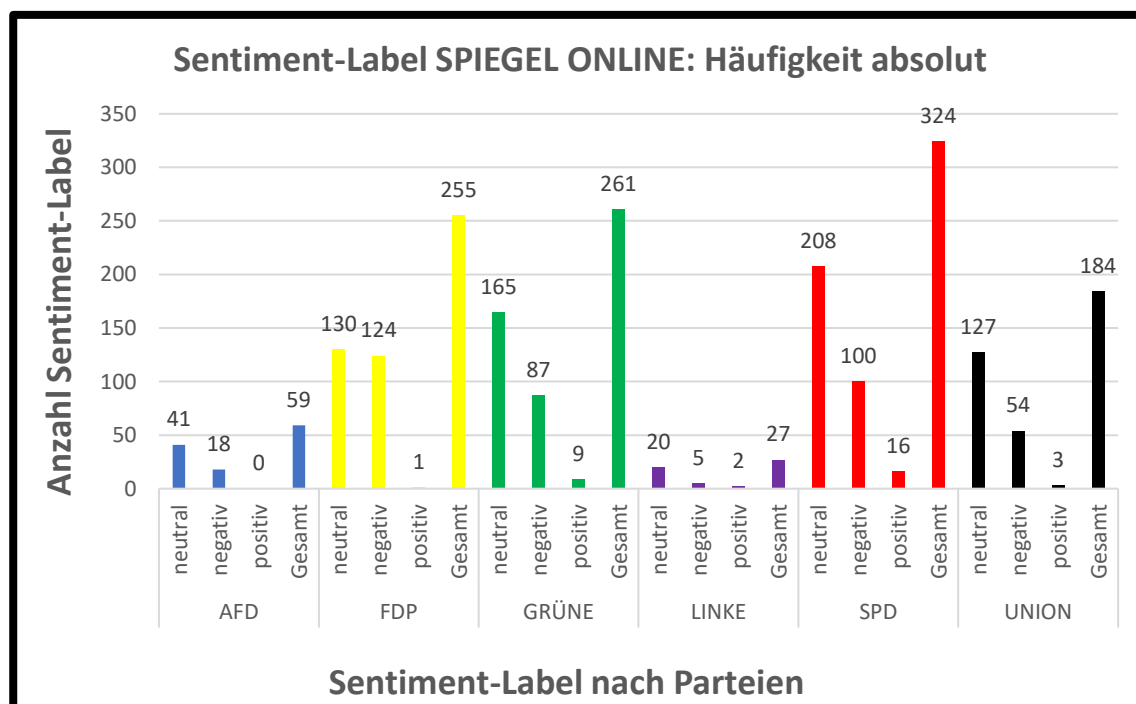


Abbildung 8: Sentiment-Label-Verteilung SPIEGEL ONLINE Häufigkeit absolut, n = 1110, Balken Gesamt entsprechen Anteil Parteien am Gesamtsample Spiegel Online

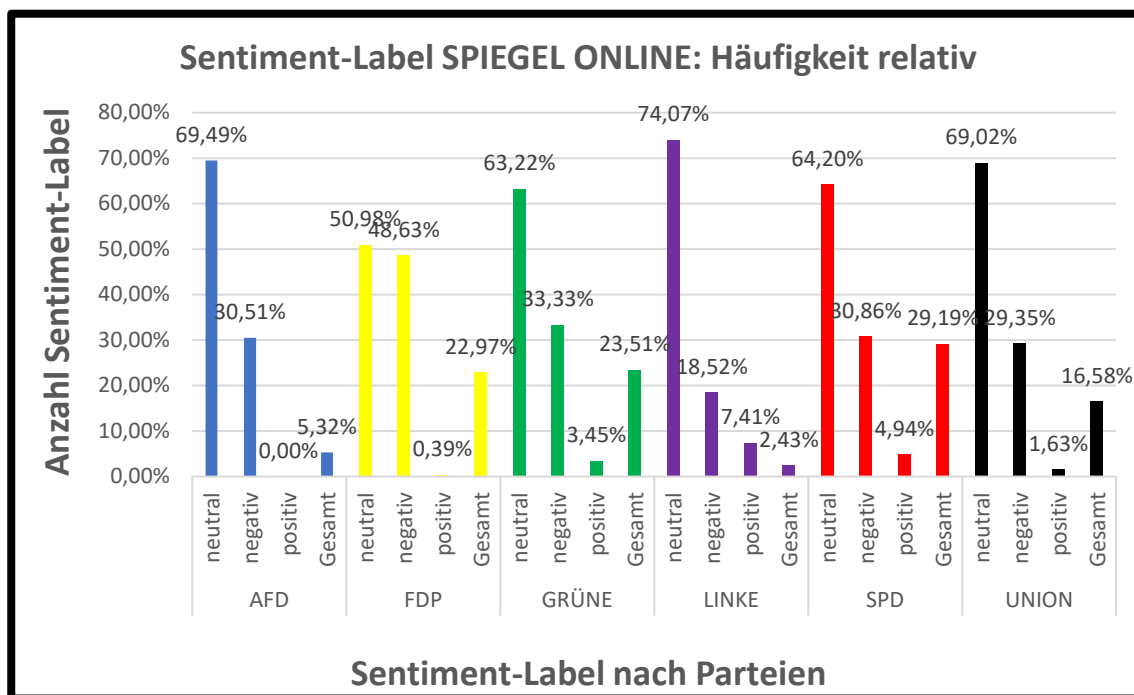


Abbildung 9: Sentiment-Label-Verteilung SPIEGEL ONLINE Häufigkeit relativ, $n = 1110$, Balken Gesamt entsprechen Anteil Parteien am Gesamtsample Spiegel Online

5.4 Evaluierung Sentiment-Analyse: „FAZ.NET“

Mit Blick auf die parteiunabhängige Verteilung der Sentiment-Labels im hier generierten Sample wird deutlich, dass das Online-Portal der „Frankfurter Allgemeinen“ seine grundsätzlichen Ruf eines neutralen und politisch heterogen gefärbten Medienhauses gerecht. Mehr als zwei Drittel der analysierten Textstellen von „FAZ.NET“ wurden neutral bewertet, nur 7% aller Sätze in der Stichprobe wurden positiv bewertet, dadurch zeigt sich ein neutral-kritischer Charakter der Berichterstattung.

Sentiment-Label	Häufigkeit absolut	Häufigkeit relativ
positiv	37	7,10%
negativ	122	23,42%
neutral	362	69,48%
Gesamt	521	100,00%

Abbildung 10: Entitätsunabhängige Sentiment-Label-Verteilung FAZ.NET, $n = 521$

Die neutral-kritische Tendenz und Ausrichtung der journalistischen Arbeitsweise bei „FAZ.NET“ bestätigt sich auch bei der Analyse der Häufigkeitsverteilung der Klassifizierungslabel nach Partei (vgl. Abb. 11 & 12). Die Linke spielt in der digitalen Berichterstattung der Frankfurter allgemeinen eine marginalisierte Rolle und kommt im Vergleich zu anderen Parteien nicht oft genug im Sample vor, um valide Aussagen über die im Medienhaus getroffene Gewichtung der Partei treffen zu können. Auch die AFD fällt

im bei der Analyse der Anteile am Gesamtsample zurück, erneut konnte das verwendete Klassifizierungsmodell keinen positiv konnotierten Satz finden. Mit einem Anteil von fast 90% der mittiger eingeordneten Parteien am Gesamtsample wird der zentralisierte Fokus von „FAZ.NET“ deutlich. Über Union, SPD, Grüne und FDP wird auf ähnlichem Niveau positiv, negativ und neutral geschrieben. Bei allen Parteien überwiegt die neutrale Berichterstattung mit mindestens 60%. Die in der Bundesrepublik etablierte Einschätzung der „Frankfurter Allgemeinen“ als konservatives Medium kann durch diese Analyse nicht bestätigt werden. Zwar wird die Union im vorliegenden Sample im Vergleich zu den anderen Parteien mit knapp 10% etwas häufiger positiv erwähnt, allerdings kann diese Einschätzung mit Blick auf die grundsätzlich sehr niedrigen Fallzahlen in der positiven Kategorie (7%) nicht als signifikant gewertet werden und bedarf weiterer Analysen.

Inwieweit sich die drei journalistischen Produkte „FAZ.NET“, „Spiegel Online“ und „Lage der Nation“ unterscheiden und ähneln, soll in einem abschließenden Quervergleich untersucht werden. Die bisherigen Erkenntnisse durch die pro Medienhaus für sich stehenden Ergebnisse der Sentiment-Analysen sollen so in Kontext gesetzt werden.

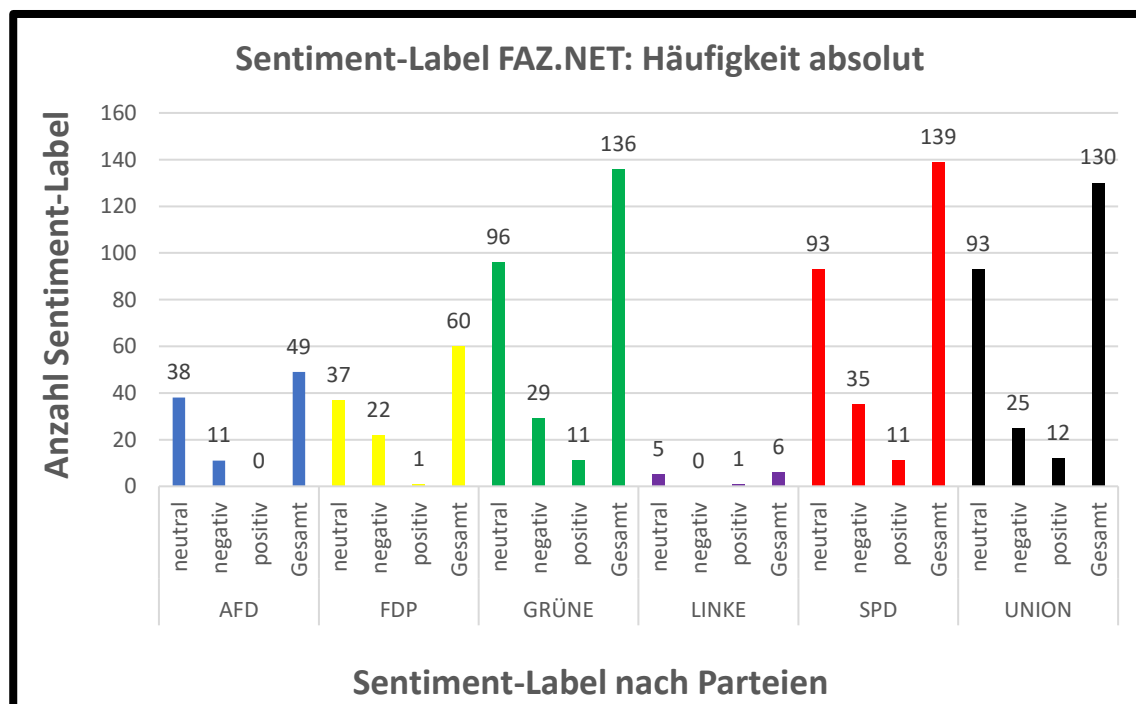


Abbildung 11: Sentiment-Label-Verteilung FAZ.NET Häufigkeit absolut, n = 521, Balken Gesamt entsprechen Anteil Parteien am Gesamtsample FAZ.NET

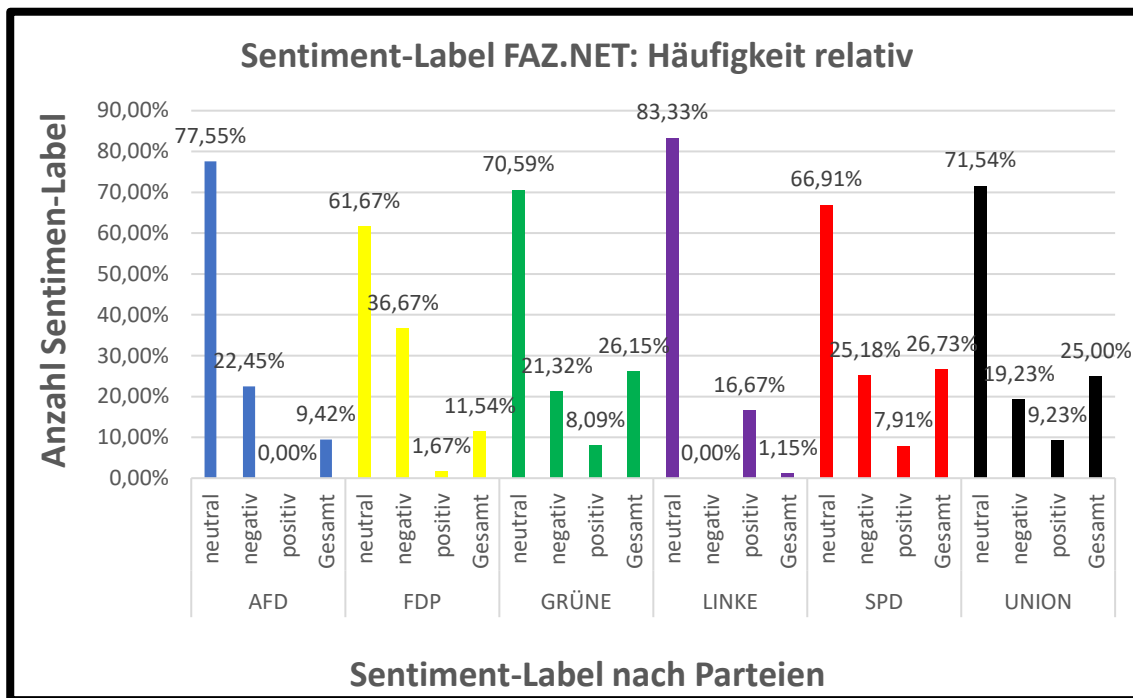


Abbildung 12: Sentiment-Label-Verteilung FAZ.NET Häufigkeit relativ, n = 521, Balken Gesamt entsprechen Anteil Parteien am Gesamtsample FAZ.NET

5.5 „Lage der Nation“, „Spiegel Online“ und „FAZ.NET“ – Berichterstattung im Quervergleich

Die Auswertung der Sentiment-Analysen der „Lage der Nation“ und der Vergleichsparameter „Spiegel Online“ und „FAZ.NET“ hat gezeigt, dass alle drei journalistischen Produkte keine parteipolitische Färbung aufweisen. Diese Erkenntnis bestätigt sich im intermedialen Vergleich der drei Untersuchungsobjekte. Mit einem Anteil von je über 60% an neutral klassifizierten Sätzen weisen sowohl die „Lage der Nation“ als auch „FAZ.NET“ und „Spiegel Online“ eine mehrheitlich neutrale Berichterstattung auf. Bei allen drei journalistischen Produkten überwog auch kategorisiert nach Parteien die neutrale Berichterstattung deutlich (vgl. Abb. 4-12). Nennenswerte Ausreißer waren hier nicht zu finden, Tendenzen bei FAZ und Spiegel hinsichtlich ihrer theoretisch erarbeiteten politischen Färbung wurden nur vereinzelt entdeckt und aufgrund geringer Fallzahl als nicht aussagekräftig gewertet. Neben der neutralen Art der journalistischen Arbeit überwiegt sowohl bei der „Lage der Nation“ als auch bei „Spiegel Online“ und „FAZ.NET“ der kritische Charakter in der Berichterstattung. Bei allen drei Analyseeinheiten liegt der Anteil der positiv bewerteten Textstellen im einstelligen Prozentbereich, vor allem im Sample von „Spiegel Online“ findet sich mit weniger als 3% kaum positives Sentiment (vgl. Abb. 4, 7, 10, 13).

Sentiment-Label	Häufigkeit absolut	Häufigkeit relativ
positiv	122	4,85%
negativ	787	31,30%
neutral	1605	63,84%
Gesamt	2514	100,00%

Abbildung 13: Entitätsunabhängige Sentiment-Label-Verteilung Gesamtsample (Lage der Nation, Spiegel Online, FAZ.NET), n = 2514

Grundsätzlich berichtet die „Lage der Nation“ demzufolge politisch nicht gefärbter über das politische Geschehen, als die beiden Onlineportale, auch wenn das durch den Charakter der eher nüchtern-sachlichen Tatsachenschilderung und Kurzinformation auf Onlineportalen im Vergleich zur meinungsbildenden Diskussion im Podcast vermutet werden kann. Dies spricht für die sachlich-neutrale Qualität der Lage der Nation. Analog zum grundsätzlich ähnlichen Charakter der journalistischen Arbeit gleichen sich die drei untersuchten Medienprodukte auch in der Bewertung der einzelnen Parteien. AFD und Linke finden bei allen drei Medienhäusern weniger oft den Weg in die Berichterstattung, die AFD wurde im ganzen Sample nur neutral oder negativ erwähnt. Mit leichten Abweichungen versehen werden am häufigsten die regierenden Ampelparteien und die Union als größte Oppositionspartei in die Meldungen und Reportagen integriert, die SPD als Kanzlerpartei am häufigsten. Über die „Lage der Nation“ und die Vergleichsparameter „Spiegel Online“ und „FAZ.NET“ gibt es hier keine großen Abweichungen, das Onlineportal des Spiegelverlags berichtet etwas seltener über die Union. Bei der Verteilung der Sentiment-Klassen bei SPD, Grünen, FDP und Union zeigt sich ebenfalls ein Muster über die drei Medienprodukte hinweg. Die neutrale Schreibart liegt bei allen vier Parteien im Gesamtsample bei einem Anteil von rund zwei Drittel, weder „FAZ.NET“ oder „Spiegel Online“ noch die „Lage der Nation“ verweisen hier Ausreißer. Negativ konnotierte Sätze in denen entweder die Regierungsparteien oder die Union genannt wurde, vereinnahmen rund 20% der drei Stichprobengruppen, positive Erwähnungen liegen in dieser Analyse-schiene immer im einstelligen Prozentbereich (vgl. Abb. 14, Anhang 5).

Zusammenfassend kann durch die vollzogene Quervergleichsanalyse zwischen der „Lage der Nation“, „Spiegel Online“ und „FAZ.NET“ allen drei journalistischen Produkten politische Neutralität zu gesprochen werden. Keines der drei Medienhäuser neigt zu tendenziöser Berichterstattung, weder grundsätzlich, noch parteipolitisch. Sowohl der Podcast als auch die beiden Onlineportale liegen auf sehr ähnlichem Niveau, was die allgemeine

und parteibezogene Häufigkeit von neutral, positiv und negativ konnotierten Textstellen betrifft.

Partei	Sentiment-Label	Häufigkeit absolut	Häufigkeit relativ
AFD	neutral	116	77,55%
	negativ	55	22,45%
	positiv	0	0,00%
	Gesamt	171	6,80%
FDP	neutral	308	61,67%
	negativ	213	36,67%
	positiv	20	1,67%
	Gesamt	541	21,53%
GRÜNE	neutral	346	70,59%
	negativ	159	21,32%
	positiv	28	8,09%
	Gesamt	533	26,15%
LINKE	neutral	30	83,33%
	negativ	5	0,00%
	positiv	6	16,67%
	Gesamt	41	1,63%
SPD	neutral	454	66,91%
	negativ	214	25,18%
	positiv	37	7,91%
	Gesamt	705	28,05%
UNION	neutral	351	71,54%
	negativ	138	19,23%
	positiv	33	9,23%
	Gesamt	522	20,77%
GESAMT		2513	100,00%

Abbildung 14: Entitätsabhängige Sentiment-Label-Verteilung Gesamtsample (Lage der Nation, Spiegel Online, FAZ.NET), $n = 2513$

6 Machine Learning und geisteswissenschaftliche Textanalyse - Fazit, Einordnung, Ausblick

Ein Fazit der hier durchgeführten Analyse, eine kritische Einordnung der verwendeten Methodik und ein Ausblick hinsichtlich weiterer Forschungs- und Technologieaspekte soll anhand der beiden theoretisch erarbeiteten Fragestellungen abgehandelt werden.

1. Ist es möglich mithilfe von Maschinellern eine Methode zu entwickeln, die etablierte journalistische Produkte hinsichtlich ihrer politischen Einordnung und Neutralität klassifizieren kann?

Die grundsätzliche technische Möglichkeit, journalistische Produkte zu klassifizieren, kann anhand der durchgeführten Analysemethoden dieser Untersuchung klar bestätigt werden, das wissenschaftliche Potenzial von Maschinellern in der Textanalyse im sozial- und geisteswissenschaftlichen Rahmen hat sich deutlich gezeigt. Die Ergebnisse und vor allem die Signifikanz maschinell gestützter Methoden müssen mit Blick auf die Erfahrungen dieser Untersuchung aber heterogener bewertet und eingeordnet werden und sind abhängig von zahlreichen inhaltlich-methodischen und modelltechnologischen Faktoren. Im inhaltlich-methodischen Ansatz dieser Analyse taten sich zahlreiche Problemstellungen und Erweiterungsdimensionen hervor. Schon bei der Datenakquise wurde deutlich, dass die verwendete Datenbasis großen Einfluss auf die Ergebnisse hat. Die Entscheidung für die Onlineportale des Spiegelverlages und der „Frankfurter Allgemeinen“ ist vor allem aus Gründen der technischen Handhabung und Zugänglichkeit gefallen. Aus inhaltlicher Sicht hätte eine Analyse der Printprodukte der beiden Medienhäuser einen besseren Vergleichswert und eine höhere Aussagekraft hinsichtlich der theoretisch erarbeiteten politischen Färbung gebracht. Dies liegt daran, dass auf den Onlineportalen vor allem Kurznachrichten, aktuelle Entwicklungen und Liveticker veröffentlicht werden, wohingegen tiefergehende politische Analysen, meinungsbehaftete Artikel und Kolumnen sich zumeist in den jeweiligen Printprodukten finden oder aber im digitalen Raum hinter Paywalls nur gebührenpflichtig zugänglich sind, ein Umstand der im Rahmen dieser Analyse nicht umgangen werden konnte. Auch hinsichtlich der hier konstruierten Methode, journalistische Texte nur anhand von Sentiment-Klassifizierung in Textstellen die bundespolitische Entitäten enthalten, ergeben sich dimensionale Erweiterungsmöglichkeiten. Nur durch die Nennung und die Kategorisierung des Sentiments dieser Nennung kann zwar ein Richtungsweiser geschaffen werden, wie politisch ausgerichtet ein Medienprodukt ist, um voll umfänglich und abschließend bestimmen zu können, welcher politischen Strömung das Medium angehört und wie stark, sind weitere, mitunter auch qualitative Methoden vonnöten. Von besonderem Interesse wäre in diesem Zusammenhang die Befragung von Expert:innen. Der Einbezug von Expert:innenwissen im quantitativen Umfang würde die Methodik um einen sinnvollen Mix-Methods-Ansatz erweitern und die analytische Aussagekraft erhöhen. Abschließend hat auch die Wahl der Entitäten, die in die Sentiment-Analyse mit einbezogen werden, Einfluss auf den Untersuchungsausgang. Als Praxisbeispiel soll hier eine Problemstellung der vorangegangenen Analyse dienen. Die Doppeldeutigkeit der Entitäten und der kontextuelle Abhängigkeit stellt die Methodik vor Herausforderungen, die Entität „Union“ ist im bundespolitischen Kontext eine

zentral, hat aber mit Blick auf die häufige sprachliche Verwendung im Zusammenhang mit der Europäischen Union die Genauigkeit der Stichprobe verschlechtert. Zusätzlich kann eine stärkere sprach- und politikwissenschaftliche Vor- und Aufbereitung der Named-Entity-Recognition zum Beispiel mit Blick auf parteiunabhängige politische Inhalte oder Umgangssprache die Aussagekraft der hier entworfenen Modellierung entscheidend erhöhen.

Ebenso breit gesät wie die inhaltlichen Möglichkeiten, um diese Studie zu verbessern, sind die technischen Aspekte und Limitationen. Als primäres Ziel für die Erfahrungen des hier durchgeführten Modell-Trainings gilt klar, das eingehend beschriebene Phänomen des Overfittings zu vermeiden oder weniger stark ausfallen zu lassen. Um dies zu erreichen, gibt es verschiedene Methoden und Stellschrauben, die im beschriebenen Finetuning-Prozess und beim Hyperparametertraining angepasst und verbessert werden können. Grundsätzlich ist die Vorhersagekraft des trainierten Sprachmodells immer abhängig vom für das Training verwendeten Datensatz. Die hier verwendete Datenbasis wurde aus Gründen der darin enthaltenen politischen Sprache gewählt. Diese schon sehr enge Anpassung an die Problemstellung, die das Modell lösen soll, kann zur Folge haben, dass dieses nach dem Training zu spezifisch auf den Trainingsdatensatz war. Als Folge daraus hat das Modell dann Probleme, neue, dem Trainingsdatensatz fremdere, Daten zu klassifizieren, weist dadurch einen höheren Validation Loss auf und leidet unter Overfitting (vgl. Backhaus et al., 2015: 94; IBM 2023). Plakativ formuliert lernt das Modell den Trainingsdatensatz nur auswendig und kann keine allgemein gültigen Regeln und Gewichtungen in die verschiedenen Ebenen des neuronalen Netzes integrieren - verschärft wird dieser Effekt durch den eher geringeren Umfang des Trainingsdatensatz. In der Retrospektive hätte so eine Modellierung mit einem größeren und allgemeineren Korpus, der dann ohnehin auch Fragmente politischer Sprache enthält, zu besseren Evaluierungsergebnissen und damit zu einer höheren Modellgüte geführt. Für die Verarbeitung eines größeren Datensatzes wären aber auch mehr Kapazitäten im Bereich Rechenzeit und -leistung vonnöten gewesen, die im Rahmen dieser Analyse nicht zur Verfügung standen. Auch mit Blick auf das Hyperparametertraining hätte die Modellgüte durch noch feinstufigeres Finetuning und längeres exploratives Anpassen der Parameter verbessert werden können. Besonders die adaptive Justierung der Lernrate könnte in weiterführenden Analysen eine Rolle spielen. Insgesamt gibt es aus der technisch-informatischen Sicht zahlreiche Perspektiven, um journalistische Produkte noch signifikanter hinsichtlich ihrer Qualität und politischen Färbung einordnen zu können.

2. Kann diese Methode den noch nicht voll umfänglich politisch eingeordneten Podcast „Lage der Nation“ als neue Form der politischen Berichterstattung auf der politischen Karte der Bundesrepublik verorten?

Grundsätzlich konnte die „Lage der Nation“ als kritisch-neutrales Medium ohne nennenswerte politische Färbung eingeordnet werden. Die beiden Podcaster Ulf Buermeyer und Philip Banse berichten vor allem über die Regierungsparteien und die Union als größte Oppositionspartei und ordnen die Leistungen und Aktionen, dieser um die politische Mitte zentrierten, Parteien kritisch ein. Die zweite Fragestellung muss dementsprechend analog zur ersten Fragestellung dieser Studie zweischneidig beantwortet werden. Während eine grundsätzliche Einordnung und Klassifizierung möglich war, ist die Aussagekraft durch die zuvor beschriebenen inhaltlichen und technischen Limitationen als eingeschränkt zu bewerten. Um journalistische Produkte abschließend signifikant bewerten und einordnen zu können, bedarf es mehrdimensionaler Untersuchungsebenen, unter denen die hier gewählte Methode der maschinell gestützten Sentiment-Analyse aber als zentral eingeschätzt werden kann.

7 Literaturverzeichnis

ag.ma. "Reichweite der Top-10-Nachrichtenseiten in Deutschland im Januar 2023 (in Millionen Unique User)." Chart. 3. März, 2023. Statista. Zugegriffen am 23. März 2023.

<https://de.statista.com/statistik/daten/studie/165258/umfrage/reichweite-der-meistbesuchten-nachrichtenwebsites/>

Backhaus, K., Erichson, B., Plinke, W., & Weiber, R. (2015). *Multivariate Analysemethoden—Eine anwendungsorientierte Einführung*. Springer.

Barczok, A. (2019). *Die Macher des Podcasts „Lage der Nation“ im c't-Gespräch* (Burmeyer, U., & Banse, P). [Heise Online]. <https://www.heise.de/hintergrund/Die-Macher-des-Podcast-Lage-der-Nation-im-c-t-Gespraech-4578047.html>

Block, E. (1981). Freedom and Equality: Indicators of Political Change in Sweden, 1945-1975. In *Advances In Content Analysis* (S. 241–250). Sage.

Bönisch, J. (2006). *Meinungsführer oder Populärmedium? Das journalistische Profil von Spiegel Online*. (Bd. 3). Lit Verlag.

Bundesanzeiger. Frankfurter Allgemeine Zeitung Gesellschaft mit beschränkter Haftung - Jahresabschluss zum Geschäftsjahr vom 01.01.2021 bis zum 31.12.2021. (2023). Bundesanzeiger. <https://www.bundesanzeiger.de/pub/de/suchergebnis?9>

Das sind die Goldenen Blogger 2017. (2018). Die Goldenen Blogger. https://die-goldenen-blogger.de/neuigkeiten/das-sind-die-goldenen-blogger-2017-2_57084.html

Der Spiegel: Verkaufte Auflage 2021 | Statista. (o. J.). Abgerufen 20. September 2022, von <https://de.statista.com/statistik/daten/studie/13232/umfrage/auflage-der-wochenzeitschrift-der-spiegel-seit1995/>

Devlin, J., Chang, M.-W., Lee, K., & Toutanova, K. (2018). *BERT: Pre-training of Deep Bidirectional Transformers for Language Understanding*.

Frankfurter Allgemeine. (o.D.). *Über uns—Zeitung für Deutschland*. Frankfurter Allgemeine. <https://www.frankfurterallgemeine.de/die-faz>

Guhr, O., Schumann, A.-K., Bahrmann, F., & Böhme, H.-J. (2020). Training a Broad-Coverage German Sentiment Classification Model for Dialog Systems. *Proceedings of the 12th Conference on Language Resources and Evaluation (LREC 2020)*. 12th Conference on Language Resources and Evaluation (LREC 2020), Marseille.

Hoeres, P. (2014, Oktober 21). Frankfurter Allgemeine Langeweile. *Frankfurter Allgemeine Zeitung*, S. 22.

Hoeres, P. (2019). *Zeitung für Deutschland: Die Geschichte der FAZ*. Benevento.

IVW. "Ranking der auflagenstärksten überregionalen Tageszeitungen in Deutschland im 4. Quartal 2022." Chart. 23. Januar, 2023. Statista. Zugriffen am 23. März 2023. <https://de.statista.com/statistik/daten/studie/73448/umfrage/auflage-der-ueberregionalen-tageszeitungen/>

IBM. (2023). *What is overfitting?* IBM Topics. <https://www.ibm.com/topics/overfitting>

Katre, P. D. (2019). NLP Based Text Analytics and Visualization of Political Speeches. *International Journal of Recent Technology and Engineering*, 8(3), 8574–8579.

Liu, Y., Ott, M., Goyal, N., Du, J., Mandar, J., Chen, D., Lewis, M., Zettlemoyer, L., & Stoyanov, V. (2019). *RoBERTa: A Robustly Optimized BERT Pretraining Approach*.

Lührmann, K. (2019). *Podcasts als Raum politisch-medialer Kommunikation* (V. Wehdeking & G. E. Grimm, Hrsg.; Bd. 9). Tectum.

Medienkompass. (2023). *Deutsche Medienlandschaft*. Medienkompass.org. <https://medienkompass.org/deutsche-medienlandschaft/>

Neuberger, C. (2006). Vorwort. In J. Bönisch (Hrsg.), *Meinungsführer oder Populärmedium? Das journalistische Profil von Spiegel Online* (Bd. 3). Lit Verlag.

Norddeutscher Rundfunk. (2022). *Rudolf Augstein: Ein streitbarer Geist*. Norddeutscher Rundfunk. <https://www.ndr.de/geschichte/koepfe/Rudolf-Augstein-Ein-streitbarer-Geist,rudolfaugstein101.html>

Ostermaier, S. (2016). Das Beste aus 2016: Apple veröffentlicht die besten Apps, Filme, Bücher und mehr. *Caschys Blog*. <https://stadt-bremerhaven.de/das-beste-aus-2016-apple-veroeffentlicht-die-besten-apps-filme-buecher-und-mehr/>

Overfelt, F. (2020). *What are Ben Shapiro & Pod Save America Talking About? An NLP Analysis*. floverfelt.org. <https://floverfelt.org/posts/podcast-nlp-analysis.html>

Patzelt, W. J. (2013). *Einführung in die Politikwissenschaft: Grundriß des Faches und studiumbegleitende Orientierung*. Rothe.

Pew Research Center. (2018). *News Media and Political Attitudes in Germany*. <https://www.pewresearch.org/global/fact-sheet/news-media-and-political-attitudes-in-germany/>

Pöttker, H. (2012). Meilenstein der Pressefreiheit – 50 Jahre "Spiegel“-Affäre. *Aus Politik und Zeitgeschichte*, Nr. 29-31/2012.

PZ-online. "Verkaufte Auflage des Nachrichtenmagazins Der Spiegel in den Jahren 1995 bis 2021." Chart. 22. Januar, 2022. Statista. Zugegriffen am 23. März 2023. <https://de.statista.com/statistik/daten/studie/13232/umfrage/auflage-der-wochenzeitschrift-der-spiegel-seit1995/>

Sadrozinski, J. (2013). Zwischen Beruf und Berufung. Wie sich das Bild des Journalisten wandelt. In L. Kramp, D. Ballwieser, L. Novy, & K. Wenzlaff (Hrsg.), *Journalismus in der digitalen Moderne Einsichten – Ansichten – Aussichten*. Springer.

Schulzki-Haddouti, C., Bunjes, M., & Jakob, G. (2009). Begrenzter Journalismus. Was beeinflusst die Entfaltung eines Qualitätsjournalismus. *Schweigen, Lügen und Vertuschen – Wenn die Wahrheit nicht mehr öffentlich wird*. 14. MainzerMedienDisput, Rengsdorf/Hardert.

Spiegel Gruppe. (2016). *70 Jahre DER SPIEGEL – 70 Jahre investigativer Journalismus: Am 4. Januar 1947 erschien die erste Ausgabe des deutschen Nachrichten-Magazins*. <https://gruppe.spiegel.de/news/pressemitteilungen/detail/70-jahre-der-spiegel-70-jahre-investigativer-journalismus-am-4-januar-1947-erschien-die-erste-ausgabe-des-deutschen-nachrichten-magazins>

Spiegel Gruppe. (2021). *Die SPIEGEL-Standards*. <https://gruppe.spiegel.de/journalismus/die-spiegel-standards>

Spiegel Gruppe. (o. D.). *Zahlen und Daten zur SPIEGEL-Gruppe*. <https://gruppe.spiegel.de/unternehmen/zahlen-und-daten>

Springer, Axel. "Reichweite der Frankfurter Allgemeinen Zeitung in den Jahren 2003 bis 2021 (in Millionen Lesern)." Chart. 26. Januar, 2022. Statista. Zugegriffen am 23. März 2023. <https://de.statista.com/statistik/daten/studie/381991/umfrage/reichweite-der-frankfurter-allgemeinen-zeitung/>

Tumasjan, A., Sprenger, T. O., Sandner, P. G., & Welpe, I. M. (2011). Election Forecasts With Twitter: How 140 Characters Reflect the Political Landscape. *Social Science Computer Review*, 29(4), 402–418.

Wu, M. L. (2020). *Capstone Project 2: Twitter Political Compass Machine*. GitHub. <https://github.com/minglwu427/Twitter-Political-Compass-Machine>

8 Datenquellenverzeichnis

Trainingsdatensatz: Huggingface 2021. `mox/german_politicians_twitter_sentiment`. Max Weissenbacher. https://huggingface.co/datasets/mox/german_politicians_twitter_sentiment. Zuletzt abgerufen am 30.03.2023.

Datensatz Transkript „Lage der Nation“: 2016 – 2023. Lage der Nation: Podcast Transkript. Philip Banse. <https://archive.org/details/ldn-transkript>. Zuletzt abgerufen am 30.03.2023.

Datensatz „Spiegel Online“ (Webcrawling). 2023. Spiegel Online. <https://www.spiegel.de/>. Zuletzt abgerufen am 30.03.2023.

Datensatz „FAZ.NET" (Webcrawling). 2023. FAZ.NET. <https://www.faz.net/aktuell/>. Zuletzt abgerufen am 30.03.2023.

9 Anhang

Alle für diese Studie verwendeten Dateien und der Code finden sich auf GitHub. Die folgende Auflistung dient als Übersicht und Linkverzeichnis.

Anhang 1: Code,

https://github.com/BergJakob/project_podcast/tree/main/SRC

Anhang 2: Data Retrieval,

https://github.com/BergJakob/project_podcast/tree/main/Data/data_retrieval

Anhang 3: Finetuning,

https://github.com/BergJakob/project_podcast/tree/main/Data/finetuning

Anhang 4: Named-Entity-Recognition,

https://github.com/BergJakob/project_podcast/tree/main/Data/ner

Anhang 5: Sentiment-Analyse,

https://github.com/BergJakob/project_podcast/tree/main/Data/sa

10 Abbildungsverzeichnis

Abbildung 1: Confusion Matrix Modelltraining	12
Abbildung 2: Verlauf Training & Validation Loss Modell 1 ("bert-base-multilingual-cased")	13
Abbildung 3: Evaluierungsmetriken Trainingsmodelle.....	14
Abbildung 4: Entitätsunabhängige Sentiment-Label-Verteilung LAGE DER NATION	14
Abbildung 5: Sentiment-Label-Verteilung „Lage der Nation“ Häufigkeit absolut.....	15
Abbildung 6: Sentiment-Label-Verteilung „Lage der Nation“ Häufigkeit relativ.....	16
Abbildung 7: Entitätsunabhängige Sentiment-Label-Verteilung SPIEGEL ONLINE.....	16
Abbildung 8: Sentiment-Label-Verteilung SPIEGEL ONLINE Häufigkeit absolut	17
Abbildung 9: Sentiment-Label-Verteilung SPIEGEL ONLINE Häufigkeit relativ	18
Abbildung 10: Entitätsunabhängige Sentiment-Label-Verteilung FAZ.NET	18
Abbildung 11: Sentiment-Label-Verteilung FAZ.NET Häufigkeit absolut	19
Abbildung 12: Sentiment-Label-Verteilung FAZ.NET Häufigkeit relativ	20
Abbildung 13: Entitätsunabhängige Sentiment-Label-Verteilung Gesamtsample.....	21
Abbildung 14: Entitätsabhängige Sentiment-Label-Verteilung Gesamtsample	22

11 Datenschutzbestimmungen und Urheberrechtliche Erklärung

Die gesammelten Daten dieser Analyse wurden im Einklang mit dem Datenschutz- und Urheberrecht erhoben. Dabei wurde darauf geachtet, dass die Daten ausschließlich aus öffentlich zugänglichen Quellen stammen und keine personenbezogenen Daten erhoben wurden. Zudem wurde bei der Verarbeitung der Daten darauf geachtet, dass diese in anonymisierter Form vorliegen und nicht auf einzelne Personen zurückgeführt werden können. Darüber hinaus habe ich darauf geachtet, die Daten korrekt zu zitieren und auf die Urheberrechte der jeweiligen Quellen zu achten. Die gesammelten Daten werden ausschließlich für Forschungszwecke verwendet und wurden nicht aus kommerziellem Interesse gesammelt. Jegliche Weitergabe oder Veröffentlichung der Daten erfolgt unter Wahrung der Persönlichkeitsrechte der Betroffenen und im Einklang mit den geltenden gesetzlichen Bestimmungen.

Plagiatserklärung

Lehrstuhl für Informationswissenschaft

Universität Regensburg

Universitätsstraße 31

D-93053 Regensburg

Eidesstattliche Erklärung*

Hiermit erkläre ich,

Name: Berg

Vorname: Jakob

Geb.-Datum: 16.09.1998

an Eides statt gegenüber dem Lehrstuhl für Informationswissenschaft der Universität Regensburg, dass die vorliegende, an diese Erklärung angefügte Hausarbeit mit dem Thema:

Der Neutralitätsanspruch von Qualitätsjournalismus am Beispiel des Politpodcasts „Lage der Nation“ – Eine Sentiment-Analyse mit maschinellem Lernen

Titel der Lehrveranstaltung: Natural Language Engineering 2

im Semester: Sommersemester 2022

selbständig und nur unter Zuhilfenahme der im Quellen- und Literaturverzeichnis genannten Werke angefertigt wurde.

Regensburg, 31.03.2023



Eigenhändige Unterschrift