

„Auch heute war die Stimmung im Allgemeinen fest.“ Zero-Shot Klassifikation zur Bestimmung des Media Sentiment an der Berliner Börse zwischen 1872 und 1930

Wehrheim, Lino

lino.wehrheim@ur.de
Universität Regensburg

Borst, Janos

borst@informatik.uni-leipzig.de
Universität Leipzig

Burghardt, Manuel

burghardt@informatik.uni-leipzig.de
Universität Leipzig

Niekler, Andreas

aniekler@informatik.uni-leipzig.de
Universität Leipzig

Motivation und Korpus

In einer Szene des Films „The Wolf of Wall Street“ hält der Börsenmakler Jordan Belfort, verkörpert durch Leonardo Di Caprio, eine Ansprache vor seinen Mitarbeitern, woraufhin eine Parade nackter Musiker durch das Großraumbüro zieht, gefolgt von einer Gruppe leicht bekleideter Frauen und einer orgiastischen Feier in den Büroräumen. Auch wenn sich der Film auf eine besonders schillernde Persönlichkeit konzentriert, so veranschaulicht er doch, wie verrückt es an der Börse bisweilen zugehen kann. Tatsächlich spielen dort wie auch in anderen Lebensbereichen nicht nur „harte“ Informationen, sondern eben auch vermeintliche „weiche“ Aspekte wie Stimmungen und Gefühle eine wichtige Rolle, die sich in kollektiven Irrationalitäten, Hypes und Ängsten niederschlagen (Akerlof and Shiller 2009; Shiller 2015). Der Ökonom John Maynard Keynes prägte 1936 dafür den Ausdruck der „animal spirits“ (Keynes 1936). Ein Ansatz, die Bedeutung der Stimmung für die Börse greifbar zu machen, besteht darin, das in der Finanzpresse zum Ausdruck kommende Media Sentiment mittels Sentiment Analyse (Liu 2015) zu messen und dessen Einfluss auf

die Kursentwicklung zu bestimmen (Tetlock 2007; García 2013; Hanna, Turner und Walker 2020).¹

Das so gemessene Media Sentiment wird dabei als Spiegel der Stimmung der Börsianer, des Investor Sentiment, interpretiert. Das vorliegende Projekt verfolgt das Ziel, Sentiment Daten für die Berliner Börse für den Zeitraum zwischen 1872 und 1930 zu generieren, die dann für verschiedene inhaltliche Fragestellungen genutzt werden können.² Das Ziel des Beitrags besteht darin, die mit der Erhebung dieser Daten verbundenen Herausforderungen aufzuzeigen und unseren Ansatz der Zero-Shot-Klassifikation vorzustellen.

Datengrundlage ist ein Textkorpus, das aus täglichen Marktberichten der Berliner Börsen-Zeitung (BBZ), der wichtigsten Finanzzeitung dieser Epoche, besteht.³ Diese Berichte, die im Median etwa 540 Wörter umfassen, enthalten eine kompakte verbale Beschreibung des täglichen Geschehens an der Berliner Börse: In welcher Stimmung befand sich die Börse? Was beeinflusste die allgemeine Stimmung? Welche Aspekte wiesen Besonderheiten auf? Wie entwickelten sich bestimmte Teilmärkte, etwa Eisenbahnaktien?

Die Berichte wurden von uns in einem mehrstufigen Prozess aus den ganzseitigen Scans der BBZ extrahiert (Liebl und Burghardt 2020), die die Staatsbibliothek Berlin zur Verfügung stellt.⁴ Dazu wurde mittels einer eigenen OCR-Pipeline und Layout-Detection zunächst die relevante Ausgabe (Morgen vs. Abend), dann die relevante Seite und zuletzt der jeweils relevante Seitenabschnitt identifiziert, was sich aufgrund des wandelnden Layouts der Zeitung und insbesondere der Berichte als durchaus komplexes Problem erwies. Insgesamt wurden so knapp 18 000 Berichte extrahiert, was einem Textvolumen von etwa 9,87 Millionen Wörtern entspricht. Für den Großteil des Untersuchungszeitraums zählt das Korpus knapp 300 Berichte pro Jahr (der Börsenhandel fand von Montag bis Samstag statt). Für die Zeit des Ersten Weltkriegs und die Jahre zwischen 1922 bis 1924 ergeben sich aufgrund eines wenig standardisierten und häufig wechselnden Berichtsformats sowie häufiger Börsenfeiertage bislang nur 150 bis 200 Texte pro Jahr. Einige weitere Datenlücken in den 1870er Jahren, die aus unvollständigen Überlieferungen der BBZ resultieren, wurden mit dem Pendant aus der *Vossischen Zeitung* gefüllt, die ebenfalls einen täglichen Börsenbericht veröffentlichte. Dabei wurde durch verschiedene Tests sichergestellt, dass beide Zeitungen eine weitestgehend übereinstimmende Bewertung der Gesamtstimmung an der Berliner Börse angeben.⁵

Methode

Die Börsenberichte der BBZ weisen im Hinblick auf die Bestimmung des in ihnen enthaltenen Sentiments eine vierfache Herausforderung auf. Zunächst wird die Analyse durch die spezielle Domäne verkompliziert. Die Tonalität eines ökonomischen Texts bzw. eines Finanztexts hängt in hohem Maße vom Kontext bzw. der „Richtung“ der Aussage ab (Loughran und McDonald 2011; Malo et al. 2014; Xing et al. 2020). So induziert bspw. das Wort „Verlust“ nur dann eine negative Tonalität, wenn dieser „steigt“, „sich einstellt“, „verharrt“, etc. Geht er jedoch zu-

rück, resultiert dagegen eine Aussage mit positiver Tonalität. Im Börsen-Kontext besonders problematisch: Für die eine Marktseite mögen fallende Kurse etwas Negatives sein. Für die andere, die darauf gesetzt hat, dass die Kurse zurückgehen, ist der gleiche Vorgang hingegen etwas Positives (Hausse- vs. Baissepartei). Hinzu kommt, dass das Börsenjargon durch ein ganz eigenes Vokabular gekennzeichnet ist („Contremine“, „debarrassieren“, „Reprise“, usw.). Neben solch hochspezifischen Ausdrücken finden sich weiterhin viele Beispiele für die Verwendung eines Standardlexikons um eine spezifische Marktstimmung – teilweise geradezu metaphorisch – zu beschreiben: In der Fachsprache der Börse sind die Geschäfte etwa „flau“, „matt“, „fest“ oder „lustlos“ (Krupke 1904; Kautsch 1912). Einerseits standardisiert derlei Jargon zwar den Sprachgebrauch erheblich; andererseits erschwert es den Einsatz von off-the-shelf Lösungen was beispielsweise die automatische Sentiment Analyse solcher Sprachbelege angeht, da in Standardressourcen die Börsen-spezifische Bedeutung diese Begriffe bzw. die Begriffe selbst nicht berücksichtigt werden. Dies wird zuletzt noch dadurch verstärkt, dass wir es teils mit einem veralteten Sprachgebrauch zu tun haben, der sich neben terminologischen Besonderheiten durch Schachtelsätze, Verklammerungen, vielfache Verneinungen, Konjunktive, Querverweise und andere Besonderheiten auszeichnet. Historizität und Sprachwandel erschweren jedoch nicht nur die Sentiment Analyse, sondern auch andere NLP-Ansätze wie etwa Named Entity Recognition (siehe etwa Hellrich et al. 2019, Ehrmann et al. 2021), da im Falle historischer Sprache viele Standard-NLP-Ressourcen nicht ohne Weiteres nutzbar sind.

Vor dem Hintergrund dieser vielfältigen Herausforderungen, die sich durch die spezifische Domäne und die historische Sprachstufe ergeben, waren erste Experimente mit den in Tabelle 1 aufgeführten Sentiment-Lexika nicht erfolgreich. Dies untermauert bestehende Erkenntnisse aus der einschlägigen Fachliteratur bzgl. einer geringeren Performance Wörterbuch-basierter Ansätze gegenüber Machine-Learning-Verfahren (siehe etwa Mishev et al. 2020 und van Atteveldt et al. 2021).⁶

Als Alternative wurden deshalb aktuelle Ansätze aus dem Bereich neuronaler Sprachmodelle erprobt. Durch das Vortrainieren großer Transformer-basierter Sprachmodelle, wie etwa BERT, wurden in den letzten Jahren immer wieder Durchbrüche in verschiedenen Anwendungsbereichen des Natural Language Processing erzielt (Devlin et al. 2019). Das übliche Vorgehen ist es, ein solch vortrainiertes Sprachmodell auf Task-spezifischen Daten nachzutrainieren und es damit an die eigene Anwendungsdomäne anzupassen (Finetuning). Man spricht hier von *transfer learning* (Zhuang et al. 2020). Unter dem Namen Zero-Shot-Klassifikation (früher: Data-less Classification, Chang et al. 2008) existieren seit Kurzem Verfahren, bei denen für die Zielaufgabe gar keine aufwendig erstellten Trainingsdaten mehr für das Finetuning zur Verfügung stehen müssen, aber dennoch ein domänenspezifischer Klassifikator generiert werden kann (vgl. Yin et al. 2019, Veeranna et al. 2016, Brown et al. 2020). Der große Vorteil des Zero-Shot-Verfahrens ist es, dass hierfür keinerlei (also zero) Trainingsdaten vorliegen müssen – deren Generierung je nach Projekt mit hohen Kosten verbunden sein kann –, sondern

Klassifikationsergebnisse im Sinne eines Transfer-Lernens erzielt werden. Einer der erfolgreichsten Ansätze für die Zero-Shot-basierte Textklassifikation stützt sich dabei auf sog. *Entailment*-Modelle, die darauf trainiert sind, einen logischen Widerspruch oder Implikationen zweier Sätze zu erkennen (Yin et al. 2019). Dazu werden die Zielkategorien für den Klassifikator in natürlichsprachliche Sätze umformuliert und mit dem zu klassifizierenden Text verglichen.

Tabelle 1: Übersicht zu genutzten Sentiment-Lexika für die deutsche Sprache.

Name	Sprache	URL
BPW dictionary	Deutsch (optimiert für Finanzsprache)	https://www.uni-giessen.de/fbz/fb02/forschung/research-networks/bsfa/textual-analysis/index.html
SentiWS	Deutsch (allgemein)	http://www.uilwaltinger.de/sentiment/
German Emotion Analysis	Deutsch (allgemein)	https://www.romanklinger.de/emotion/
Affective Norms for German Sentiment Terms (ANGST)	Deutsch (allgemein)	https://link.springer.com/article/10.3758/s13428-013-0426-y
Lexikon von Chen/Skiena	Multilingual	https://aclanthology.org/at-tachments/P14-2063.Data-sets.zip

In unserem konkreten Anwendungsfall erstellen wir einen Klassifikator für die Sentiment-Kategorien „positiv“, „neutral“ und „negativ“, welche entsprechend in den folgenden Sätzen verbalisiert werden: „Die Stimmung an der Börse ist {positiv|neutral|negativ}“. Der Klassifikator entscheidet durch automatische Analyse aller sprachlichen *Entailments* sodann, welche dieser drei Hypothesen sich am wahrscheinlichsten aus dem Eingabetext schlussfolgern lässt. Verwendet wurde hierzu ein BERT-Modell⁷, das auf dem deutschsprachigen Teil des XNLI (*cross natural language inference*) Datensatzes (Conneau et al. 2018) getestet wurde.

Vor der Klassifizierung wurden zunächst alle Berichte in einzelne Sätze segmentiert. Nach ersten Experimenten fiel auf, dass einige Sätze in den Marktberichten stark deskriptiven Charakter haben, angezeigt bspw. durch eine längere Auflistung von Unternehmen in direkter Abfolge. Solche Sätze enthalten aber keine relevanten Informationen aus Perspektive des Media Sentiment und wurden deshalb schon im Vorfeld heuristisch herausgefiltert. Dabei wurde Domänenwissen in generalisierbare Heuristiken übertragen, die dann automatisiert auf das gesamte Korpus angewendet wurden. Beispielsweise prüft die Heuristik, ob längere Aufzählungen oder gehäuftes Auftreten von numerischen Werten in einem Satz vorkommen. Letzteren kommt eine besondere Bedeutung zu, da sie meist in Aussagen wie „Der Kurs ist um 2 % gefallen“ oder „Der Kurs stieg bis 218“ auftreten. Zwar können diese Sätze, wie das erste Beispiel zeigt, durchaus eine Form von Tonalität enthalten. Allerdings drücken sie aus unserer Sicht keine Stimmungen oder Gefühle, sondern eine finanzwirtschaftliche Information aus, wie sie auch an anderen Stellen der Zeitung, etwa der Kurstabelle, zu finden ist (siehe mehr dazu im Abschnitt „Herausforderungen“). Das Herausfiltern der Sätze mit numerischen Werten soll sicherstellen, dass unser Sentiment Index nicht durch solche Informationsaussagen verzerrt wird. Weiterhin wurden sehr kurze (kürzer als 30 Zeichen) und sehr lange Sätze (länger als 800 Zeichen) sowie Überschriften entfernt. Alle verbleibenden Sätze

wurden dann nach dem vorgestellten Zero-Shot-Ansatz klassifiziert und einem Sentiment Score zugewiesen (positiv: 1, negativ: -1, neutral: 0), der letztlich über alle Sätze eines Berichts zu einem Gesamt-Sentiment-Score gemittelt wurde.

Um die Qualität der Zero-Shot-Klassifizierung anekdotisch zu evaluieren, wurden 150 Sätze als Gold Standard durch einen Domänenexperten annotiert. Auf dieser Basis ergab sich eine korrekte Klassifikation des automatischen Ansatzes in 74% der Fälle. Von den 26% falsch erkannten Sätzen entfällt der Großteil auf solche, die fälschlicherweise als positiv bewertet wurden, obwohl sie eigentlich als neutral einzustufen sind. Tabelle 2 fasst verschiedene Metriken zur Evaluation des Zero-Shot-Ansatzes sowie die Verteilung der Klassen zusammen.

Tabelle 2: Evaluationsmetriken und Klassenverteilung des Zero-Shot Klassifikators

	Negativ	Neutral	Positiv
Precision	83.67%	36.84%	82.35%
Recall	70.69%	46.67%	83.17%
F1-score	76.64%	41.18%	82.76%
Distribution	30.69%	15.87%	53.44%

In einer ersten Voranalyse des Korpus (siehe Abb. 1) ergeben sich die nachfolgenden Zeitreihen, die jeweils die diachronen Sentiment-Werte für ein Gleitfenster von 30 und 365 Tagen darstellt. Eine Darstellung der Sentimentausschläge auf Tagesbasis wurde verworfen, da diese starken Fluktuationen unterworfen ist und damit nur schwer interpretierbar ist.

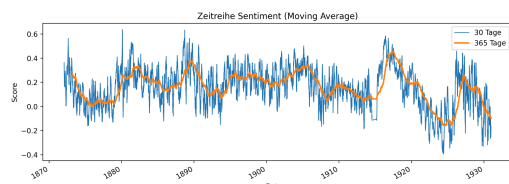


Abbildung 1: Zeitreihe Sentiment über den gesamten Zeitraum des Korpus (Moving Average).

In der geglätteten Darstellung auf Monats- (blau) und Jahresbasis (orange) zeichnet sich dagegen eine Entwicklung mit Phasen unterschiedlicher Stimmungslagen ab, die zumindest auf den ersten Blick plausibel erscheint.⁸

Beispielsweise ist die Spekulationsblase der frühen 1870er Jahre mit einer sehr positiven Stimmung verbunden, die sich nach dem Börsenkrach von 1873 rasch eintrübt. Das gleiche Schema ergibt sich auch für andere Phasen rasch steigender Kurse, wie etwa in den Jahren 1888–90 und 1926/27. Auch der massive Einsturz der Kurse während des Ersten Weltkriegs ist mit einem starken, wenn auch zeitlich verzögerten Rückgang des Sentiment-Index verbunden. Gerade der Anstieg des Sentiment-Index zu Beginn des Ersten Weltkriegs wirft jedoch Fragen auf. Zum einen stellt sich an dieser Stelle die technische Frage, wie belastbar die Ergebnisse angesichts einer Fehlerrate von 26% sind; diesen Punkt führen wir weiter unten aus. Zum anderen zeigt dieser Aspekt inhaltliche Fragestellungen auf: In welchem Verhältnis ste-

hen Stimmung und Kursentwicklung zueinander? Sind beispielsweise steigende (fallende) Kurse immer mit einer positiven (negativen) Stimmung verbunden? Welche Trendwenden lassen sich in der Börsenstimmung identifizieren? Und welche historischen Ereignisse und Entwicklungen lassen sich als Ursachen für veränderte Stimmungslagen festmachen?

Herausforderungen

Unsere aktuellen Experimente zeigen deutlich, dass – eingedenk der eingangs beschriebenen Korpus-Probleme – die Anwendung eines Zero-Shot-Verfahrens auf dem Korpus von Marktberichten bereits vielversprechende Ergebnisse liefert. Gleichzeitig ist die derzeitige Fehlerrate von 26% zu hoch, um belastbare Aussagen treffen zu können. Zudem ist die Klassifizierung der Aussagen in nur eine Kategorie mit drei Ausprägungen angesichts der in den Börsenberichten enthaltenen Aussagen nicht unproblematisch, da sie nur ein sehr grobkörniges und, was schwerer wiegt, potentiell verzerrtes Bild liefert. Konkret scheint uns die Berücksichtigung zweier weiterer Kategorien geboten, die von den uns bekannten Studien zu Finanzmarktsentiment allerdings ausgeblendet werden. Grundsätzlich beziehen sich die einzelnen Aussagen eines BBZ-Berichts auf drei verschiedene Ebenen. Aussagen wie „Das Aussehen der Börse war heute überaus unfreundlich“ beziehen sich auf die gesamte Börse, andere wie „Schantung-Aktien setzten nach niedrigerem Beginn eine Besserung durch“ dagegen auf einzelne Wertpapiere. Dazwischen rangieren Aussagen wie „Elektrische Werte blieben behauptet“, die sich auf einen bestimmten Teilbereich des Börsenhandels beziehen. Damit ein Sentiment Score die Gesamtstimmungslage an der Börse korrekt widerspiegelt, muss eine Gewichtung dieser Aussagentypen erfolgen, da ansonsten die Gefahr einer Verzerrung der ermittelten Tonalität besteht.

Die Aussagen unterscheiden sich zudem in einer weiteren Dimension. Beispiele wie „Auf dem Rentenmarkte hat sich die Stimmung für Italiener auch heute nicht gebessert“ beschreiben die Börsenstimmung selbst, während Aussagen wie „Die rheinisch-westfälischen Bahnen büßten durchschnittlich nur 3 pCt. ein“ zwar auch Tonalität enthalten, allerdings eher eine (positive, neutrale oder negative) finanzwirtschaftliche Information wiedergeben. Während beispielsweise Takala et al. (2014) solchen Aussagen ein positives Sentiment bescheinigen, würden wir sie als Informationsaussagen mit positiver Tonalität definieren, da es aus unserer Sicht zwischen Sentiment im Sinne der Stimmung an der Börse und Sentiment als Ausdruck einer Informationsaussage zu differenzieren gilt. Beide Unterscheidungen, Aussageebene und Aussagegegenstand, scheinen uns für die Bestimmung eines repräsentativen Sentiment Scores als sehr wichtig. Dementsprechend erfordern die Börsenberichte der BBZ eigentlich eine Aspekt-basierte Sentimentanalyse, bei der neben der Entitätskategorie (Börse, Teilmarkt, Einzeltitel) auch die Aussageart (Informations- vs. Stimmungsaussage) berücksichtigt wird.

Fazit

Die Erkenntnis, dass der Zero-Shot-Ansatz bei einfachen Klassifizierungsaufgaben im Falle einer sehr spezifischen und komplexen Domäne bereits eine hohe Datenqualität liefert, stiftet Zuversicht, da in diesem Fall bereits geringe Mengen von Annotationsdaten ausreichen, um die Datenqualität zu evaluieren und gegebenenfalls durch entsprechendes Nachtrainieren zu erhöhen. Nun stellt sich die weitergehende Frage, inwieweit dieser Ansatz auch für komplexere Aufgaben wie Aspect-Based-Sentiment geeignet ist. Erste Analysen sowie die Arbeit von Shu et al. (2022) stimmen hier optimistisch. Dies wäre insofern eine relevante Erkenntnis, als die meisten praxisbezogenen Aufgaben eher komplexerer Natur sind. Komplexere Aufgaben lassen sich im Bereich des maschinellen Lernens häufig durch eine größere Menge an annotierten Trainingsdaten lösen. Ein Vorteil des Zero-Shot Ansatzes läge hier darin, dass das Verfahren auf den Zieldaten zunächst direkt evaluiert werden kann und nur im Bedarfsfall weitere Daten für das Nachtrainieren manuell erstellt werden müssen. Insgesamt scheint dieser neuartige Ansatz aus dem Bereich des Transfer Learning also sehr vielversprechend, auch für sehr heterogene Textkorpora, wie sie in den Digital Humanities häufig vorliegen.

Fußnoten

1. Siehe Raimondo (2019) für einen Literaturüberblick.
2. Siehe Projekt-Homepage <https://media-sentiment.uni-leipzig.de>. Während dieses Zeitraums entwickelte sich die Berliner Börse zu einem international bedeutsamen Handelsplatz (Pohl 2002; Buchner 2019).
3. Zur Geschichte der BBZ, siehe die Beiträge in Bertkau (1930).
4. Siehe <https://zefys.staatsbibliothek-berlin.de/list/>.
5. Trotz dieser Datenlücken haben wir aufgrund ihrer finanzhistorischen Bedeutung an der BBZ als Untersuchungsobjekt festgehalten, zumal nicht ohne Weiteres zu nachvollziehbar ist, ob die *Vossische Zeitung* nicht ähnliche, zeitlich versetzte Lücken aufweist.
6. Daher sehen wir von einer detaillierten Evaluation des Lexikonansatzes gegenüber dem des Maschinellen Lernens an dieser Stelle ab. Zudem soll im weiteren Projektverlauf (siehe Ausblick) eine Aspekt-basierte Sentiment Analyse umgesetzt werden, die sich aus unserer Sicht wesentlich leichter in einen neuronalen Ansatz integrieren lässt als dies bei lexikonbasierten Ansätzen der Fall ist.
7. <https://huggingface.co/svalabs/gbert-large-zero-shot-nli>.
8. An dieser Stelle liegt ein systematischer Vergleich dieser Reihen mit historischen Finanzmarktdaten nahe. Da ein solcher Vergleich eine Vielzahl methodischer Fragen aufwirft, vertagen wir diesen Schritt aus Platzgründen auf eine künftige Publikation. Natürlich gilt es bei der vorläufigen Interpretation das Risiko eines Confirmation Bias im Hinterkopf zu halten.

Bibliographie

- Akerlof, George A. und Robert J. Shiller.** 2009. "Animal Spirits." Princeton.
- Bertkau, Friedrich (Hg.)** 1930. "75 Jahre Berliner Börsen-Zeitung." Berlin.
- Buchner, Michael.** 2019. "Die Spielregeln der Börse: Institutionen, Kultur und die Grundlagen des Wertpapierhandels in Berlin und London, ca. 1860-1914." Tübingen.
- Brown, Tom, Benjamin Mann, Nick Ryder, et al.** 2020. "Language models are few-shot learners." NeurIPS.
- Chang, Ming-Wei, Lev Ratinov, Dan Roth und Vivek Srikumar.** 2008. "Importance of Semantic Representation: Dataless Classification." Proceedings of the Twenty-Third AAAI Conference on Artificial Intelligence.
- Conneau, Alexis, Ruty Rinott, Guillaume Lample, Adina Williams, Samuel Bowman, Holger Schwenk und Veselin Stoyanov.** "XNLI: Evaluating Cross-lingual Sentence Representations." Proceedings of the 2018 Conference on Empirical Methods Natural Language Processing.
- Devlin, Jacob, Ming-Wei Chang, Kenton Lee und Kristina Toutanova.** 2019. "BERT: Pre-training of Deep Bidirectional Transformers for Language Understanding." NAACL-HLT (1): 4171-4186.
- Ehrmann, Maud, Ahmed Hamdi, Elvys Linhares Pontes, Matteo Romanello und Antoine Doucet.** 2021. "Named Entity Recognition and Classification on Historical Documents: A Survey". ACM Computing Survey.
- García, Diego.** 2013. "Sentiment during Recessions." The Journal of Finance 68 (3): 1267-1300.
- Hanna, Alan J., John D. Turner und Clive B. Walker.** 2020. "News Media and Investor Sentiment during Bull and Bear Markets." The European Journal of Finance 26 (14): 1377-95.
- Kautsch, Jacob.** 1912. "Handbuch des Bank- und Börsenwesens für Kaufleute, Industrielle, Kapitalisten, Bankiers und Bankbeamte. Mit besonderer Berücksichtigung Deutscher, Österreichischer und Schweizerischer Verhältnisse und den in Deutschland, Österreich und der Schweiz Geltenden Bank- und Börsengesetzen." Berlin.
- Keynes, John Maynard.** 1936. "The General Theory of Employment, Interest and Money." London.
- Krupke, Franz.** 1904. "Krupkes Konversations-Lexikon der Börse und des Handels und praktischer Führer für Kapitalisten." Berlin.
- Liebl, Bernhard und Manuel Burghardt.** 2020. "From Historical Newspapers to Machine-Readable Data: The Origami OCR Pipeline." Proceedings of the 1st Workshop on Computational Humanities Research (CHR).
- Liu, Bing.** 2015. "Sentiment analysis: mining opinions, sentiments, and emotions". New York.
- Loughran, Tim und Bill McDonald.** 2011. "When Is a Liability Not a Liability? Textual Analysis, Dictionaries, and 10-Ks." The Journal of Finance 66 (1): 35-65.
- Malo, Pekka, Ankur Sinha, Pekka Korhonen, Jyrki Wallenius und Pyy Takala.** 2014. "Good Debt or Bad Debt: Detecting Semantic Orientations in Economic Texts." Journal of the Association for Information Science and Technology 65 (4): 782-96.
- Mishev, Kostadin, Ana Gjorgjevikj, Irena Vodenska, Lubomir T. Chitkushev und Dimitar Trajanov.** 2020.

"Evaluation of Sentiment Analysis in Finance: From Lexicons to Transformers," IEEE Access 8: 131662-131682.

Pohl, Hans (Hg.) 2002. "Geschichte des Finanzplatzes Berlin." Frankfurt am Main.

Raimondo, Carlo. 2019. "The Media and the Financial Markets: A Review." Asia-Pacific Journal of Financial Studies 48 (2): 155-84.

Shiller, Robert J. 2015. "Irrational Exuberance." Revised and expanded third edition. Princeton.

Shu, Lei, Hu Xu, Bing Liu und Jiahua Chen. 2022. "Zero-Shot Aspect-Based Sentiment Analysis." arXiv preprint arXiv:2202.01924.

Takala, Pyry, Pekka Malo, Ankur Sinha und Oskar Ahlgren. 2014. "Gold-Standard for Topic-Specific Sentiment Analysis of Economic Texts." Proceedings of the Ninth International Conference on Language Resources and Evaluation (LREC'14).

Tetlock, Paul C. 2007. "Giving Content to Investor Sentiment: The Role of Media in the Stock Market." The Journal of Finance 62 (3): 1139-68.

van Atteveldt, Wouter, Mariken A. C. G. van der Velden und Mark Boukes. 2021. "The Validity of Sentiment Analysis: Comparing Manual Annotation, Crowd-Coding, Dictionary Approaches, and Machine Learning Algorithms". Communication Methods and Measures, 15 (2): 121-140

Veeranna, Sappadla Prateek, Jinseok Nam, Eneldo Loza Mencía und Johannes Fürnkranz. 2016. "Using Semantic Similarity for Multi-Label Zero-Shot Classification of Text Documents." European Symposium on Artificial Neural Networks, Computational Intelligence and Machine Learning.

Xing, Franz Z., Lorenzo Malandri, Yue Zhang und Erik Cambria. 2020. "Financial Sentiment Analysis: An Investigation into Common Mistakes and Silver Bullets." Proceedings of the 28th International Conference on Computational Linguistics, 978-87.

Yin, Wenpeng, Jamaal Hay und Dan Roth. 2019. "Benchmarking Zero-shot Text Classification: Datasets, Evaluation and Entailment Approach." EMNLP

Zhuang, Fuzhen, Zhiyuan Qi, Keyu Duan, Dongbo Xi, Yongchun Zhu, Hengshu Zhu, Hui Xiong und Qing He.(2020). "A comprehensive survey on transfer learning." Proceedings of the IEEE, 109(1), 43-76.