

Daube, Carl Heinz; Krivenkov, Vladislav

Working Paper

Generative AI Tools zur Prognose von Leitzins-Entscheidungen: eine Fallstudie am Beispiel der Leitzinsentscheidungen der Federal Reserve

IUCF Working Paper, No. 6/2024

Suggested Citation: Daube, Carl Heinz; Krivenkov, Vladislav (2024) : Generative AI Tools zur Prognose von Leitzins-Entscheidungen: eine Fallstudie am Beispiel der Leitzinsentscheidungen der Federal Reserve, IUCF Working Paper, No. 6/2024, ZBW - Leibniz Information Centre for Economics, Kiel, Hamburg

This Version is available at:

<https://hdl.handle.net/10419/293992>

Standard-Nutzungsbedingungen:

Die Dokumente auf EconStor dürfen zu eigenen wissenschaftlichen Zwecken und zum Privatgebrauch gespeichert und kopiert werden.

Sie dürfen die Dokumente nicht für öffentliche oder kommerzielle Zwecke vervielfältigen, öffentlich ausstellen, öffentlich zugänglich machen, vertreiben oder anderweitig nutzen.

Sofern die Verfasser die Dokumente unter Open-Content-Lizenzen (insbesondere CC-Lizenzen) zur Verfügung gestellt haben sollten, gelten abweichend von diesen Nutzungsbedingungen die in der dort genannten Lizenz gewährten Nutzungsrechte.

Terms of use:

Documents in EconStor may be saved and copied for your personal and scholarly purposes.

You are not to copy documents for public or commercial purposes, to exhibit the documents publicly, to make them publicly available on the internet, or to distribute or otherwise use the documents in public.

If the documents have been made available under an Open Content Licence (especially Creative Commons Licences), you may exercise further usage rights as specified in the indicated licence.

Carl Heinz Daube und Vladislav Krivenkov¹

Generative AI Tools zur Prognose von Leitzins-Entscheidungen

Eine Fallstudie am Beispiel der Leitzinsentscheidungen der Federal Reserve

Zusammenfassung

Dieses Working Paper untersucht den Einsatz von Generative AI Anwendungen zur Prognose von Leitzinsentscheidungen der Federal Reserve. Es bewertet, ob diese Anwendungen eingesetzt werden können, um Leitzinsänderungen vorherzusagen, und vergleicht ihre Vorhersagegenauigkeit mit den Markterwartungen über einen Zeitraum von sechs Monaten.

Abstract

This working paper investigates the use of generative AI applications for predicting interest rate decisions by the Federal Reserve. It assesses whether these applications can be utilized to forecast changes in interest rates and compares their predictive accuracy with market expectations over a period of six months.

¹ **Prof. Dr. Carl Heinz Daube**, Professor für Finanzierung und ProRektor Forschung an der NBS Northern Business School - University of Applied Sciences. Er ist Studiengangsleiter für den M.Sc. Controlling & Finance und lehrt und forscht am IUCF Institut für Unternehmensrechnung, Controlling und Finanzmanagement. Er ist außerdem Partner bei MGRP Management Group Dr. Röser & Partner.

Vladislav Krivenkov, B.A. Portfolio-Manager bei der nordIX AG und Dozent für Finanzierung an der NBS Northern Business School - University of Applied Sciences.

Inhalt

Abbildungsverzeichnis	II
Tabellenverzeichnis	III
1. Generative AI im Asset Management	1
2. Gang der Untersuchung	2
3. Versuchsbeschreibung und Vorbereitung.....	3
3.1. Fragestellungen der Versuche.....	3
3.2. Dateninput.....	3
3.3. Durchführung der Versuche.....	5
4. Ergebnisse der Versuche	7
5. Bewertung der Ergebnisse	11
6. Fazit und kritische Würdigung	13
7. Quellen- und Literaturverzeichnis	16
Anhang	18

Abbildungsverzeichnis

Abb. 1: Beispiel Prompt aus Versuch Nr. 1	8
Abb. 2: Reaktion von Gemini auf Prompt	9

Tabellenverzeichnis

Tab. 1: Fragestellungen der Versuche	3
Tab. 2: Verwendete Daten und Informationen.....	4
Tab. 3: Ergebnisse der Versuche	10
Tab. 4: Auswertung der Konfusionsmatrix.....	11

1. Generative AI im Asset Management

Die Methoden der künstlichen Intelligenz (KI) finden zunehmend ihren Einzug im Asset Management. Begründet ist das Interesse nach Lösungsangeboten mit KI-Bezug unter anderem durch das Streben nach Wettbewerbsvorteilen, Kosteneffizienz und Produktivitätssteigerungen. Darüber hinaus spielt die Verbesserung der Qualität von Produkten oder Dienstleistungen eine zentrale Rolle. Diese Qualitätssteigerungen könnten beispielsweise durch Verbesserungen im Risikomanagement und in der Verbesserung der Qualität von Investitionsentscheidungen und Handelssignalen realisiert werden.²

Mit dem Release des Chat-basierten large language model (LLM) ChatGPT durch das Unternehmen OpenAI im November 2022 kam es zu einer Demokratisierung eines bis dahin als Nischenthema geltenden Aufgabenfeldes. Belegt wird diese These mit der explosionsartig angestiegenen Nutzerzahl, welche in den ersten fünf Tagen nach Veröffentlichung die 1-Mio. Marke überschritt. Wenige Monate später im Januar 2023 wurde vermeldet, dass 100 Mio. User die Anwendung nutzen.³ Trotz der Komplexität der dahinter liegenden Modellen, wie beispielsweise tiefen neuronalen Netzen, ermöglicht ChatGPT Anwendern komplexe Aufgaben durch die Verarbeitung natürlicher Sprache zu lösen.⁴

Die Anwendungsfelder von natural language processing (NLP) sind vielfältig. Unter anderem lassen sich durch ChatGPT Text-Klassifikations-Aufgaben, Sentiment Analysen oder die Generierung von natürlicher Sprache automatisieren.⁵ Schon heute

² Vgl.: OECD: Artificial Intelligence, Machine Learning and Big Data in Finance – Opportunities, Challenges, and Implications for Policy Makers, Paris 2021.

³ Vgl.: Hu, Krystal: ChatGPT sets record for fastest growing user base, auf: Reuters, 01.02.2023, 12:18 Uhr, <https://www.reuters.com/technology/chatgpt-sets-record-fastest-growing-user-base-analyst-note-2023-02-01>

⁴ Vgl.: Reinking, Ernst; Becker, Marco: Einsatzmöglichkeiten von KI in Unternehmen – Zeitnah erreichbare Personalisierung von Large Language Models (wie ChatGPT) in Zeiten der Industrie 5.0, Kiel, Hamburg 2023.

⁵ Vgl.: Zaremba, Adam; Demir, Ender: ChatGPT: Unlocking the future of NLP in finance, in: Modern Finance, 1. Jahrgang 2023, Heft 1, S. 93-98.

gibt es eine Vielzahl von Anwendungsmöglichkeiten in der Wirtschaft⁶ im Allgemeinen und in der Finanzindustrie⁷ im Besonderen.

Im Rahmen dieser Fallstudie sollen makroökonomische Daten und Dokumente mit den Generative AI Anwendungen ChatGPT und Google Gemini ausgewertet werden, die für die Leitzinsentscheidungen der Federal Reserve eine zentrale Rolle spielen. Im zweiten Schritt sollen diese Anwendungen dann eine Prognose zur Leitzinsentwicklung auf Basis dieser Datengrundlage abgeben. Die Ergebnisse werden im Rahmen dieser Fallstudie gegen die Markterwartung und die tatsächlich eingetretenen Leitzinsentscheidungen in rund sechs Monaten verglichen. Die Ableitung einer Prognose der US-Leitzinsentwicklung durch Generative AI Anwendungen geht einen Schritt weiter als bisherige Untersuchungen, welche sich häufig auf die Text-Klassifizierungsfähigkeit von Generative AI Anwendungen beschäftigen.⁸ Damit komplementiert dieses Working Paper den bestehenden Forschungsstand.

2. Gang der Untersuchung

Im nachfolgenden Abschnitt werden die Parameter der durchgeführten Versuche, des Versuchsaufbaus, sowie der verwendeten Daten und deren Quellen erläutert.

Die Ergebnisse und die Beobachtungen der durchgeführten Versuche werden im Abschnitt 4 beschrieben und in einer Tabelle zusammengefasst.

Die Bewertung der Ergebnisse auf Basis der im Abschnitt 4 formulierten Fragestellungen findet sich im darauffolgenden Abschnitt. Abschließend erfolgt im Abschnitt 6 ein Fazit und eine kritische Würdigung.

⁶ Vgl.: z.B. Reinking, Ernst; Becker, Marco: Einsatzmöglichkeiten von KI in Unternehmen – Zeitnah erreichbare Personalisierung von Large Language Models (wie ChatGPT) in Zeiten der Industrie 5.0, Kiel, Hamburg 2023;

Reinking, Ernst; Becker, Marco: Large Language Modelle und unternehmenseigene Daten – Genauere Abfrageergebnisse dank effizientem Datenmanagement und verbesserten technischen Verfahren, Kiel, Hamburg 2024.;

Becker, Marco; Daube, Carl Heinz; Peskes, Markus; Reinking, Ernst: Zukunft der Unternehmensdaten: Chancen und Herausforderungen von KI-Anwendungen im Spannungsfeld der Corporate Governance, Kiel, Hamburg 2024.

⁷ Vgl.: z.B. Daube, Carl Heinz: Künstliche Intelligenz in der Finanz- und Investitionsentscheidung, Kiel, Hamburg 2024.

⁸ Vgl.: Hansen, Anne Lundgaard; Kazinnik, Sophia: Can ChatGPT Decipher Fed speak?, Richmond 2023 und Lee, Justina: ChatGPT can decode Fed speak, predict stock moves from headlines, auf: Bloomberg, 21.04.2024, 12:18 Uhr, <https://www.bloomberg.com/news/articles/2023-04-17/chatgpt-can-decode-fed-speak-predict-stock-moves-from-headlines>.

3. Versuchsbeschreibung und Vorbereitung

3.1. Fragestellungen der Versuche

Im Rahmen dieser Fallstudie sollen nachfolgende Fragestellungen beantwortet werden.

#	Fragestellung
1	Können die verwendeten generative AI Tools die unterschiedlichen Dateiformate richtig verarbeiten und die notwendigen Daten extrahieren?
2	Können die verwendeten generative AI Tools die Entwicklung von den verwendeten makroökonomischen Kennzahlen richtig <i>beschreiben</i> ?
3	Können die verwendeten generative AI Tools die Entwicklung von den verwendeten makroökonomischen Kennzahlen richtig <i>bewerten</i> ?
4	Können die verwendeten generative AI Tools die zur Verfügung gestellten Unterlagen auslesen und aus dem Kontext die implizierte Meinung des Zentralbank Rats zur Entwicklung der Leitzinsen in den nächsten sechs Monaten auffassen?
5	Wie ist die Prognose Qualität der generative AI Tools im Hinblick auf die Prognose der kurzfristigen Leitzins-Entwicklung?

Tab. 1: Fragestellungen der Versuche⁹

Die Ergebnisse sollen für die Fragestellungen 1 bis 4 aus einer qualitativen Perspektive bewertet werden.

3.2. Dateninput

Der Untersuchungszeitraum umfasst die Entscheidungen der *Federal Open Market Committee* (FOMC) Sitzungen unter der (aktuell laufenden) Amtszeit von Jerome Powell als Vorsitzender der Federal Reserve¹⁰. Im Versuch blieben die FOMC Sitzungen sechs Monate vor Ausbruch der Corona-Pandemie mangels Prognostizierbarkeit unberücksichtigt.

Die nachfolgenden Datensätze und Dokumente bildeten die Grundlage der unterschiedlichen Versuche.

⁹ Eigene Darstellung

¹⁰ Vgl.: o.V.: About The Fed - Board Members, auf: Federal Reserve, 21.04.2024, 12:18 Uhr, <https://www.federalreserve.gov/aboutthefed/bios/board/powell.htm>

#	Kennzahl	Quelle
1	Kursdaten des S&P 500	https://fred.stlouisfed.org/series/SP500
2	Federal Funds Target Range - Upper Limit	https://fred.stlouisfed.org/series/DFEDTARU
3	US-Arbeitslosenquote	U.S. Bureau of Labor Statistics
4	Non-Farm-Payrolls	U.S. Bureau of Labor Statistics
5	Consumer Price Index for All Urban Consumers	U.S. Bureau of Labor Statistics
6	Consumer Price Index for All Urban Consumers less food and energy	U.S. Bureau of Labor Statistics
7	Personal Consumption Expenditures Price Index	U.S. Bureau of Economic Analysis
8	Personal Consumption Expenditures Price Index excluding food and energy	U.S. Bureau of Economic Analysis
9	Federal Open Market Committee Economic projections	Federal Reserve System
10	FOMC statements	Federal Reserve System
11	FOMC statements Press Conference Transcript	Federal Reserve System
12	Implizierte Leitzinserwartungen	Bloomberg

Tab. 2: Verwendete Daten und Informationen¹¹

Die in dem Versuch berücksichtigten US Zentralbank Sitzungen erstrecken sich vom März 2018 bis zum Juli 2019 und dann wieder vom April 2020 bis zum Juli 2023. Der Zeitraum vor März 2018 blieb unberücksichtigt um zu gewährleisten, dass unterschiedliche persönliche Stile in der Kommunikation durch die Federal Reserve Vorsitzenden das Ergebnis nicht verfälschen. Entsprechend beginnt der Versuch mit der ersten Zentralbank Sitzung unter dem Vorsitz von Jerome Powell. Der Zeitraum zwischen Juli 2019 und April 2020 blieb aufgrund der fehlenden Prognostizierbarkeit der Corona-Pandemie unberücksichtigt. Der Untersuchungszeitraum des Versuchs endet im Juli 2023, da die Prognosen gegen das US-Leitzins-Niveau in den darauffolgenden sechs Monaten geprüft wurden. Die FOMC Sitzung im Januar 2024 stellt die letzte

¹¹ Eigene Darstellung

verfügbare Sitzung vor Redaktionsschluss dar. Somit gingen in die Stichprobe insgesamt 39 Leitzinsentscheidungen (n = 39) des FOMC Boards ein.

3.3. Durchführung der Versuche

Im Rahmen der Versuche werden die Generative AI Tools mit Eingabeaufforderungen (*Prompts*) befüllt. Diese enthalten zum einen die für die Auswertung notwendigen Daten und zum anderen Fragestellungen, die die Tools zur Beschreibung und Bewertung des Inputs auffordern. Anschließend werden die Generative AI Tools zur Generierung einer Prognose aufgefordert.

Versuch Nr. 1: ChatGPT-4 Prompts auf Basis der FED Dokumente (im PDF Format)

Im ersten Versuch wurden für jede Federal Open Market Committee (FOMC) Sitzung die relevanten Output-Dokumente der Entscheidungen bei *ChatGPT-4* hochgeladen. Zu diesen gehören die aktuellsten *Economic projections of Federal Reserve Board members* (erscheinen quartalsweise), die *Presse Mitteilung* über die FOMC Entscheidung sowie das *Transkript der anschließenden Pressekonferenz*.

Nach dem Hochladen der Dokumente im PDF-Format wurde folgende Eingabeaufforderung (Prompt) eingegeben:

Es ist der [Datum der FOMC Sitzung]

Die Federal Reserve hat auf Ihrer Sitzung eine Reihe von Entscheidungen getroffen. Die Entscheidungen und die Gründe für diese findest du in den angehängten Dokumenten:
- *Press Release der Entscheidung*
- *die aktuellsten Projektionen volkswirtschaftlicher Kennzahlen*
- *Transkript der Pressekonferenz*

Auf Basis dieser Informationen:
Lege dich auf eine konkrete Entscheidung fest. Werden die Leitzinsen im Verlauf der nächsten sechs Monate entweder gesenkt, konstant gehalten oder erhöht? Schreibe deine Entscheidung in die Überschrift und begründe deine Entscheidung in maximal zwei Sätzen.

Die Ergebnisse bzw. die Antworten wurden zu einer der Kategorien *Hike*, *None* oder *Cut* zugeordnet und tabellarisch dokumentiert. Die Kategorien spiegeln die Erwartung von steigenden, gleichbleibenden oder fallenden US Leitzinsen in den folgenden sechs Monaten wider.

Versuch Nr. 2: Gemini Prompts auf Basis der FED Dokumente (Hyperlinks zu den Dokumenten)

Im zweiten Versuch wurden die Dokumente bei *Google Gemini* mangels Möglichkeit nicht hochgeladen, sondern verlinkt. Der Untersuchungszeitraum ist derselbe wie im

ersten Versuch. Ebenso die verwendeten Dokumente. Analog zum ersten Versuch wurden auch in diesem Versuch die Ergebnisse den Kategorien *Hike*, *None* oder *Cut* zugeordnet.

Für diesen Versuch wurde folgender Prompt verwendet:

Es ist der [Datum der FOMC Sitzung]

Die Federal Reserve hat auf Ihrer Sitzung eine Reihe von Entscheidungen getroffen.

Die Entscheidungen und die Gründe für diese findest du in den angehängten Dokumenten:

- Press Release der Entscheidung

[URL zum Press Release auf [federalreserve.gov](https://www.federalreserve.gov)]

- die aktuellsten Projektionen volkswirtschaftlicher Kennzahlen

[URL zu den Projektionen volkswirtschaftlicher Kennzahlen auf [federalreserve.gov](https://www.federalreserve.gov)]

- Transkript der Pressekonferenz

[URL zum Transkript auf [federalreserve.gov](https://www.federalreserve.gov)]

Auf Basis dieser Informationen:

Lege dich auf eine konkrete Entscheidung fest. Werden die Leitzinsen im Verlauf der nächsten sechs

Monate entweder gesenkt, konstant gehalten oder erhöht? Schreibe deine Entscheidung in die

Überschrift und begründe deine Entscheidung in maximal zwei Sätzen.

Versuch Nr. 3: ChatGPT-4 Leitzins-Entscheidung Prognose auf Basis von volkswirtschaftlichen Daten

Während die ersten zwei Versuche sich auf eine sprachliche Auswertung der Kommunikationsdokumente der Federal Reserve fokussierten, geht der dritte Versuch dem Ansatz nach, ChatGPT-4 eine datenbasierte Prognose zur kurzfristigen Leitzinsentwicklung abgeben zu lassen. Die verwendeten Daten und die Quellen dieser wurden im Abschnitt *Dateninput* aufgelistet. Im dritten Versuch wurden die zum Zeitpunkt der FOMC Sitzung verfügbaren volkswirtschaftlichen Kennzahlen der vergangenen sechs Veröffentlichungen mitberücksichtigt. Bei den Schlusskursen des S&P 500 wurden die monatlichen Schlusskurse der zu dem Zeitpunkt vergangenen zwölf Monate miteinbezogen.

Für diesen Versuch wurde folgender Prompt verwendet:

Es ist der [Datum der FOMC Sitzung]

Die Federal Reserve hat in Ihrer Sitzung eine Leitzins-Entscheidung getroffen.

*Die aktuelle und die bisherigen Leitzins-Entscheidungen sahen so aus:
[...]*

*Die Consumer Price Inflation entwickelte sich bisher wie folgt:
[...]*

*Die Consumer Price Inflation ex Food and Energy entwickelte sich bisher wie folgt:
[...]*

*Die Personal Consumption Expenditure Inflation entwickelte sich bisher wie folgt:
[...]*

*Die Personal Consumption Expenditure Inflation ex Food and Energy entwickelte sich bisher wie folgt:
[...]*

*Die Non-Farm-Payrolls entwickelten sich bisher wie folgt:
[...]*

*Die US-Arbeitslosenquote entwickelte sich bisher wie folgt:
[...]*

*Der S&P500 entwickelte sich bisher wie folgt:
[...]*

Erstelle eine Tabelle. In eine Spalte sollen die Namen der volkswirtschaftlichen Kennzahlen eingefügt werden. In eine weitere Spalte sollst du eine Bewertung des Trends der Kennzahl abgeben. Gebe in eine weitere Spalte eine Bewertung ab ob die Entwicklung und der aktuelle Stand der jeweiligen Kennzahl für eine Leitzinserhöhung, eine Beibehaltung der Leitzinsen oder eine Zinssenkung spricht.

*Auf Basis dieser Informationen:
Lege dich auf eine konkrete Entscheidung fest. Werden die Leitzinsen im Verlauf der nächsten sechs Monate entweder gesenkt, konstant gehalten oder erhöht? Schreibe deine Entscheidung in die Überschrift und begründe deine Entscheidung in maximal zwei Sätzen.*

Analog zu den vorangegangenen Versuchen wurden die Ergebnisse den Kategorien *Hike*, *None* oder *Cut* zugeordnet.

4. Ergebnisse der Versuche

In Versuch Nr. 1 erfüllte ChatGPT-4 die Eingabeaufforderungen in der beschriebenen Art und Weise. Die Antworten umfassten zum einen die Prognose zur kurzfristigen Leitzinsentwicklung und zum anderen die Begründung dieser. Ein Beispiel von Prompt und Ergebnis zeigt die nachfolgende Grafik.

ChatGPT 4 ▾

You

2022-07-27_PC.pdf
PDF

2022-07-27_PR.pdf
PDF

2022-06-15_Proj.pdf
PDF

Es ist der 2022-07-27

Die Federal Reserve hat in Ihrer Sitzung eine Reihe von Entscheidungen getroffen.

Die Entscheidungen und die Gründe für diese findest du in den angehängten Dokumenten:

- Press Release der Entscheidung
- die aktuellsten Projektionen volkswirtschaftlicher Kennzahlen
- Transcript der Pressekonferenz

Auf Basis dieser Informationen:

Lege dich auf eine konkrete Entscheidung fest. Werden die Leitzinsen im Verlauf der nächsten sechs Monate entweder gesenkt, konstant gehalten oder erhöht?

Schreibe deine Entscheidung in die Überschrift und begründe deine Entscheidung in maximal zwei Sätzen.

ChatGPT

Entscheidung: Leitzinsen werden im Verlauf der nächsten sechs Monate erhöht.

Begründung: Die Federal Reserve hat den Zielbereich für den Federal Funds Rate auf 2-1/4 bis 2-1/2 Prozent erhöht und betont, dass weitere Erhöhungen des Zielbereichs angemessen sein werden, um die Inflation auf ihr 2 Prozent Ziel zurückzuführen [†]. Die Entscheidung reflektiert die starke Verpflichtung der Federal Reserve, die Inflation zu bekämpfen, was durch die anhaltend hohen Inflationsraten und robuste Arbeitsmarktdaten unterstrichen wird [†].



Abb. 1: Beispiel Prompt aus Versuch Nr. 1¹²

In der zweiten Versuchsdurchführung mit *Google Gemini* fielen die Antworten des Generative AI Tools, trotz Aufforderung zur Begrenzung, deutlich umfangreicher aus. Wie im Abschnitt *Durchführung der Versuche* dargestellt, wurde bei der Handlungsaufforderung bei Gemini derselbe Wortlaut verwendet wie bei ChatGPT-4. Trotz desselben Wortlauts fiel die Antwort von Google Gemini deutlich umfangreicher aus als aufgefordert – wie in der nachfolgenden Grafik zu sehen ist.

¹² Vgl.: Screenshot ChatGPT, 21.04.2024, 12:18 Uhr; OpenAI, Inc.

Gemini



◆ **Leitzinsen werden deutlich erhöht**

Die Federal Reserve wird die Leitzinsen im Verlauf der nächsten sechs Monate **deutlich erhöhen**.

Begründung:

1. Die Fed hat in der Sitzung am 15.06.2022 den Leitzins um 0,75 Prozentpunkte angehoben, die größte Erhöhung seit 1994.
2. Die Fed signalisierte in der Pressemitteilung und der Pressekonferenz, dass sie weitere Leitzinserhöhungen im Laufe des Jahres 2022 vornehmen wird, um die Inflation einzudämmen.
3. Die Projektionen volkswirtschaftlicher Kennzahlen deuten auf ein moderates Wirtschaftswachstum und eine weiterhin hohe Inflation im Jahr 2022 hin.
4. Die Fed ist besorgt über die Inflationsentwicklung und deutet an, dass sie die Leitzinsen im Jahr 2022 **deutlich** erhöhen muss, um die Preissteigerungen einzudämmen.

Hinweis:

Es ist wichtig zu beachten, dass die Entscheidung der Fed von der Entwicklung der Wirtschaftsdaten in den nächsten sechs Monaten abhängt.

Weitere Informationen:



- <https://www.federalreserve.gov/monetarypolicy/files/monetary20220615a1.pdf>
- <https://www.federalreserve.gov/monetarypolicy/files/fomcprojtabl20220615.pdf>
- <https://www.federalreserve.gov/mediacenter/files/FOMCpresconf20220615.pdf>

Zu beachten ist:

- Die beigefügten Projektionen volkswirtschaftlicher Kennzahlen sind vom 15.06.2022. Es ist wichtig, die aktuellen Projektionen zu berücksichtigen, die im September 2022 veröffentlicht werden.
- Die Entscheidung der Fed ist komplex und basiert auf einer Vielzahl von Faktoren. Die oben genannte Zusammenfassung ist stark vereinfacht und sollte nicht als vollständige Analyse der Situation betrachtet werden.

Zusätzliche Hinweise:

- Die Märkte erwarten, dass die Fed die Leitzinsen im Jahr 2022 mehrmals anheben wird.
- Die EZB hat im Dezember 2021 ihr Anleihekaufprogramm PEPP beendet und deutet an, dass sie die Leitzinsen im Jahr 2022 noch nicht erhöhen wird.

Es ist wichtig, die Entwicklung der Wirtschaftsdaten und die geldpolitischen Entscheidungen der Fed und der EZB in den nächsten sechs Monaten genau zu beobachten.

Abb. 2: Reaktion von Gemini auf Prompt¹³

¹³ Vgl.: Screenshot Gemini, 21.04.2024, 12:18 Uhr; Google LLC.

Die Ergebnisse bzw. die Prognosen der jeweiligen Versuche finden sich zusammengefasst in der nachfolgenden Tabelle.

Decision by FED (H, N; C)	Futures Implied Decision			ChatGPT-4 Forecast			Gemini Forecast			ChatGPT-4 data-based		
	Hike	None	Cut	Hike	None	Cut	Hike	None	Cut	Hike	None	Cut
Hike	11	2	0	10	3	0	10	3	0	8	5	0
None	5	14	4	7	14	2	8	14	1	2	20	1
Cut	0	3	0	0	3	0	0	3	0	0	3	0
Sum	39											

Tab. 3: Ergebnisse der Versuche¹⁴

Von den insgesamt 13 Leitzinserhöhungen im Untersuchungszeitraum wurden in Versuch Nr. 1 und Nr. 2 jeweils 10 richtig prognostiziert. Der datenbasierte Ansatz prognostizierte dagegen nur acht Leitzinserhöhungen richtig. Im Gegenzug prognostizierte der datenbasierte Ansatz die Beibehaltung des Leitzinsniveaus mit insgesamt 20 Treffern deutlich besser, als die sprachbasierten Ansätze, welche jeweils 14 richtige Treffer zählten. Die für viele Marktteilnehmer überraschenden Leitzinssenkungen in 2019¹⁵ wurden von keinem der drei Ansätze prognostiziert.

Als zusätzliche Dimension zum Vergleich der Ergebnisse wurden mit dem Bloomberg Terminal die in Futures auf US Zinsen implizierten Zinserwartungen in sechs Monaten jeweils zwei Werktage nach jeder FOMC Sitzung abgetragen und tabellarisch dokumentiert. Die dort verwendete Methodik wird von Keasler/Goff (2007) einleuchtend beschrieben.¹⁶ Die Ergebnisse finden sich in Tabelle 3. Auffällig ist, dass die Ergebnisse der sprachbasierten Ansätze und der Futures-implizierten Schätzung starke Ähnlichkeiten aufweisen.

Insgesamt lässt sich eine deutliche Ungleichverteilung der Leitzinsentscheidungen feststellen.

¹⁴ Eigene Darstellung

¹⁵ Vgl.: o.V.: Federal Reserve cuts interest rates by 0.25% - its first in a decade, auf: The Guardian, 31.07.2019, 12:18 Uhr, <https://www.theguardian.com/business/2019/jul/31/federal-reserve-cuts-interest-rates-by-025-its-first-in-a-decade>

¹⁶ Vgl.: Keasler, Terrill R.; Goff, Delbert C.: Using Fed Funds Futures to Predict a Federal Reserve Rate Hike, in: Journal of Economics and Finance Education, 6. Jahrgang 2007, Heft 2, S. 9-14.

5. Bewertung der Ergebnisse

Die Fragestellungen 1 bis 4 im Abschnitt *Fragestellungen der Versuche* können bejahend beantwortet werden. Die Generative AI Anwendungen waren in der Lage die verlinkten bzw. hochgeladenen PDF Dokumente zu verarbeiten und die notwendigen Daten zu extrahieren. Selbiges gilt für die im Prompt aufgeführten Daten. Auch gelang es ChatGPT-4 die präsentierten makroökonomischen Kennzahlen zutreffend zu beschreiben und zu bewerten, wie am Screenshot in Anhang 2 zu sehen ist.

Die Bewertung der Fragestellung 5 erfolgt quantitativ mit Hilfe einer Konfusionsmatrix. Im Zuge dessen wurden die Ergebnisse der vorhergesagten und eingetretenen Entscheidungen der vier Ansätze in einer Konfusionsmatrix ausgewertet und für die jeweiligen Klassen (*Hike*, *None* und *Cut*) die Metriken *Sensitivität (Trefferquote)*, *Präzision*, *F1 Score* und *Balanced Accuracy* mit einem R Script errechnet. Die Ergebnisse sind in der nachfolgenden Tabelle zu erkennen.

	marktbasiert	sprachbasiert		datenbasiert
	Futures Implied Decision	ChatGPT-4 Forecast	Gemini Forecast	ChatGPT-4 data-based
Sensitivität (Trefferquote)				
Hike	84,6%	76,9%	76,9%	61,5%
None	60,9%	60,9%	60,9%	87,0%
Cut	0,0%	0,0%	0,0%	0,0%
Präzision				
Hike	68,8%	58,8%	55,6%	80,0%
None	73,7%	70,0%	70,0%	71,4%
Cut	0,0%	0,0%	0,0%	0,0%
F1 Score				
Hike	75,9%	66,7%	64,5%	69,6%
None	66,7%	65,1%	65,1%	78,4%
Cut	NA	NA	NA	NA
F1 Score Mittelwert	47,5%	43,9%	43,2%	49,3%
Balanced Accuracy				
Hike	82,7%	75,0%	73,1%	76,9%
None	64,8%	61,7%	61,7%	68,5%
Cut	44,4%	47,2%	48,6%	48,6%
Balanced Accuracy Mittelwert	64,0%	61,3%	61,1%	64,7%

Tab. 4: Auswertung der Konfusionsmatrix¹⁷

¹⁷ Eigene Darstellung

Die Formeln der einzelnen Metriken finden sich zum Beispiel in Grandini/Bagli/Visani (2020) oder auch in Beleites/Salzer/Sergo (2013).

Die Trefferquoten der drei verschiedenen Prognose Ansätze weisen ein uneinheitliches Bild aus. In der Prognose von **US-Leitzinserhöhungen** (*Hike*) war der marktbasierter Ansatz mit den in Futures Preisen implizierten Leitzinsänderungen mit einer Trefferquote von knapp 85 % (11 Fälle) deutlich besser als der datenbasierte Ansatz, welcher nur 8 richtige Prognosen zählte und so auf eine Trefferquote von rund 62 % kam.

Dem gegenüber war der datenbasierte Ansatz scheinbar jedoch deutlich besser in der Prognose von **gleichbleibenden Leitzinsen**. Mit einer Trefferquote von 87 % wies dieser Ansatz ein deutlich besseres Ergebnis als der markt- und der sprachbasierte Ansatz auf. Dieses Ergebnis ist jedoch kritisch zu betrachten. In Anbetracht der Tatsache, dass Leitzinsentscheidungen auf Basis der Interpretation der vorliegenden Daten durch die Mitglieder des Federal Reserve Boards getroffen werden¹⁸, könnte die Trefferquote auch in der Entscheidungsunsicherheit des Generative AI Tools zu finden sein. In 28 der 39 Fälle schätzte ChatGPT-4 auf Grund der vorliegenden Datenbasis, dass es zu keiner Leitzinsanpassung kommt. Das liegt deutlich über den Prognosen des markt- und sprachbasierten Ansatzes, welche 19 bzw. 20-mal keine Leitzinsanpassung prognostizierten. Diese Entscheidungsunsicherheit könnte unter anderem darin begründet sein, dass die Gewichtung und die Relevanz der jeweiligen Kennzahlen unbekannt sind und darüber hinaus sich im Zeitverlauf verändern können. Zudem wurden im datenbasierten Modell nur Kennzahlen verwendet, die dem Mandat der Federal Reserve zu zuordnen sind.¹⁹ Die Leitzinssenkungen in 2019 sind ein gutes Beispiel dafür, dass die Entscheidung des Zentralbank Rats auch durch andere Argumentationen begründet sein könnte.

In der Stichprobe kam es in Summe zu drei aufeinander folgenden **Leitzinssenkungen**. Diese wurden von keinem der drei Modelle prognostiziert. Entsprechend liegt die Trefferquote bei allen vier Versuchsreihen bei 0 %.

Die schwache Trefferquote bei Leitzinssenkungen führt unter anderem auch zu einer deutlichen Verschlechterung der allgemeinen Prognose-Qualität der Modelle wie sich an den durchschnittlichen **F1 Scores** von zwischen 40 % und 50 % erkennen lässt. Durch

¹⁸ Vgl.: Bernanke, Ben: The Taylor Rule: A Benchmark for Monetary Policy, auf: Brookings, 21.04.2024, 12:18 Uhr, <https://www.brookings.edu/articles/the-taylor-rule-a-benchmark-for-monetary-policy/>

¹⁹ Vgl.: o.V.: The Fed Explained, auf: Federal Reserve, 21.04.2024, 12:18 Uhr, <https://www.federalreserve.gov/aboutthefed/the-fed-explained.htm>

die Berücksichtigung von Trefferquote und von der Präzision der Treffer eignet sich der F1 Score als Metrik um die Klassifizierungsqualität des jeweiligen Modells zu bewerten.²⁰ Den besten F1 Score kann dabei der **datenbasierte Ansatz** mit einem Wert von 49,3 % erzielen, was jedoch mit Vorsicht zu bewerten ist. Den schlechtesten Wert erzielte der **sprachbasierte Ansatz mit Google's Gemini** mit einem Wert von 43,2 %. Beim Vergleich zwischen dem marktbasieren und den sprachbasierten Ansätzen zeigt sich die deutlich bessere Prognose-Qualität des marktbasieren Ansatzes, welcher sich in einem F1 Score von knapp 50 % widerspiegelt.

Ähnliche Verhältnisse lassen sich auch in der Metrik **Balanced Accuracy** erkennen. Zum einen zeigen die Werte von größer als 60 %, dass die Modelle eine bessere Performance aufweisen als die Zufallswahl und zum anderen lässt sich wieder dieselbe Rangfolge der Qualität der Prognose-Modelle erkennen. Analog zu den Ergebnissen des F1 Scores zeigt sich auch hier der datenbasierte Ansatz als bestes Klassifikationsmodell. Darauf folgt der marktbasierte Ansatz. Die sprachbasierten Ansätze weisen auch bei dieser Metrik schwache Ausprägungen auf. Unter den sprachbasierten Ansätzen zeigt ChatGPT-4 im Rahmen des Versuchs eine leicht bessere Prognosequalität als Google's Gemini.

6. Fazit und kritische Würdigung

Im Rahmen dieser Fallstudie wurde die Prognose-Qualität von Generative AI Anwendungen im Hinblick auf die kurzfristige US-Leitzinsentwicklung untersucht. Die verwendeten Ansätze waren sowohl sprach- als auch datenbasiert. Die Ergebnisse dieser Ansätze wurden dann gegen die Einschätzung des Kapitalmarkts verglichen.

Zunächst lässt sich festhalten, dass die von den KI Tools im sprachbasierten Ansatz prognostizierten Leitzinsänderungen eine ähnliche Qualität haben wie die des Kapitalmarkts. Erneut zeigte sich, dass die verbale Forward Guidance letztendlich bedeutender für die Abschätzung der zukünftigen Leitzinsentwicklung ist als die Vielzahl der verfügbaren makroökonomischen Daten, welche im Zweifel die Prognoseunsicherheit verstärken können, wie der datenbasierte Ansatz der

²⁰ Vgl.: Grandini, Margherita; Bagli, Enrico; Visani, Giorgio: Metrics for Multi-Class Classification – An Overview, 2020, S. 6.

Untersuchung andeuten könnte. Damit bestätigt sich, dass die sogenannte Forward Guidance ein ganz wesentliches geldpolitisches Steuerungsinstrument ist.

Bei unseren Untersuchungen stellten wir fest, dass sowohl die sprach- als auch die marktbasierten Ansätze einen Bias aufweisen. Im Ergebnis kommt es auf die Interpretation der makroökonomischen Daten an. Dabei waren die Übereinstimmungen zwischen sprach- und marktbasierten Ansätzen überraschend hoch. Marktbasierte Ansätze berücksichtigen die Erwartungen, Einschätzungen und Prognosen einer großen Zahl von Marktteilnehmern, die daraus wiederum Transaktionen ableiten. Diese Transaktionen basieren schlussendlich auf Angebot und Nachfrage: In der logischen Sekunde einer abgeschlossenen Transaktion beobachten wir einen Gleichgewichtspreis.²¹ Gleichwohl bietet generative AI eine wertvolle Ergänzung und in diesem in gewissem Maße eine qualitative Verbesserung von geldpolitischen Prognosen. Damit bestätigt die Fallstudie auch das Fazit der Untersuchung von Hansen/Kazinnik (2023).

An dieser Stelle kann zusammengefasst werden, dass die Ergebnisse unserer Versuche einen ersten Eindruck davon vermitteln, in welchem Umfang generative AI die Prognose der künftigen Leitzinsentwicklung verbessern könnte. Zu berücksichtigen ist dabei, dass die Stichprobe mit $n=39$ vergleichsweise gering war. Hinzu kommt, dass jeder Zentralbankvorsitzende seinen eigenen persönlichen Kommunikationsstil pflegt, weshalb die herangezogenen Presstexte stets eine gewisse Tendenz aufweisen, den wiederum die angewendete KI *lernt* und auf dieser *tendenziösen* Basis Aussagen liefert.

In zukünftigen Untersuchungen könnten, deutlich mehr FED Dokumente ausgewertet werden und die Analyse auch auf die Europäische Zentralbank ausgeweitet werden. Als Startzeitpunkt könnte sich die Gründung der EZB anbieten, damit stünden Daten ab dem 01.06.1998 zur Verfügung. Der Zeitraum von über 25 Jahren scheint verlockend zu sein. Allerdings ist dabei zu berücksichtigen, dass in diesen Zeitraum Ereignisse (u.a. Zerfall Jugoslawiens mit anschließenden kriegerischen Auseinandersetzungen, Finanzmarktkrise, Staatsschuldenkrise, Corona, Krieg in der Ukraine und im Gazastreifen, und weitere) fielen, die die Ergebnisse verzerren könnten. Gleichwohl erscheint es reizvoll, die Versuche einmal deutlich auszudehnen.

²¹ Nicht umsonst wird im IFRS-Regime immer wieder auf den *fair value* Ansatz auf Basis von am Markt stattgefundenen Transaktionen mit einer Vielzahl von Marktteilnehmern abgestellt.

7. Quellen- und Literaturverzeichnis

- Becker, Marco; Daube, Carl Heinz; Peskes, Markus; Reinking, Ernst:** Zukunft der Unternehmensdaten: Chancen und Herausforderungen von KI-Anwendungen im Spannungsfeld der Corporate Governance, Kiel, Hamburg 2024.
- Bernanke, Ben:** The Taylor Rule: A Benchmark for Monetary Policy, auf: Brookings, 21.04.2024, 12:18 Uhr, <https://www.brookings.edu/articles/the-taylor-rule-a-benchmark-for-monetary-policy/>
- Daube, Carl Heinz:** Künstliche Intelligenz in der Finanz- und Investitionsentscheidung, Kiel, Hamburg 2024.
- Grandini, Margherita; Bagli, Enrico; Visani, Giorgio:** Metrics for Multi-Class Classification – An Overview, 2020, S. 6.
- Hansen, Anne Lundgaard; Kazinnik, Sophia:** Can ChatGPT Decipher Fed speak?, Richmond 2023.
- Hu, Krystal:** ChatGPT sets record for fastest growing user base, auf: Reuters, 01.02.2023, 12:18 Uhr, <https://www.reuters.com/technology/chatgpt-sets-record-fastest-growing-user-base-analyst-note-2023-02-01>
- Keasler, Terrill R.; Goff, Delbert C.:** Using Fed Funds Futures to Predict a Federal Reserve Rate Hike, in: Journal of Economics and Finance Education, 6. Jahrgang 2007, Heft 2, S. 9-14.
- Lee, Justina:** ChatGPT can decode Fed speak, predict stock moves from headlines, auf: Bloomberg, 21.04.2024, 12:18 Uhr, <https://www.bloomberg.com/news/articles/2023-04-17/chatgpt-can-decode-fed-speak-predict-stock-moves-from-headlines>.
- o.V.:** About The Fed - Board Members, auf: Federal Reserve, 21.04.2024, 12:18 Uhr, <https://www.federalreserve.gov/aboutthefed/bios/board/powell.htm>
- o.V.:** Federal Reserve cuts interest rates by 0.25% - its first in a decade, auf: The Guardian, 31.07.2019, 12:18 Uhr, <https://www.theguardian.com/business/2019/jul/31/federal-reserve-cuts-interest-rates-by-025-its-first-in-a-decade>
- o.V.:** The Fed Explained, auf: Federal Reserve, 21.04.2024, 12:18 Uhr, <https://www.federalreserve.gov/aboutthefed/the-fed-explained.htm>

OECD: Artificial Intelligence, Machine Learning and Big Data in Finance – Opportunities, Challenges, and Implications for Policy Makers, Paris 2021.

Reinking, Ernst; Becker, Marco: Einsatzmöglichkeiten von KI in Unternehmen – Zeitnah erreichbare Personalisierung von Large Language Models (wie ChatGPT) in Zeiten der Industrie 5.0, Kiel, Hamburg 2023.

Reinking, Ernst; Becker, Marco: Large Language Modelle und unternehmenseigene Daten – Genauere Abfrageergebnisse dank effizientem Datenmanagement und verbesserten technischen Verfahren, Kiel, Hamburg 2024.

Screenshot ChatGPT, 21.04.2024, 12:18 Uhr; OpenAI, Inc.

Screenshot Gemini, 21.04.2024, 12:18 Uhr; Google LLC.

Zaremba, Adam; Demir, Ender: ChatGPT: Unlocking the future of NLP in finance, in: Modern Finance, 1. Jahrgang 2023, Heft 1, S. 93-98.

Anhang

Anlage 1: R Script

```
# Notwendige Libraries
library(caret)

# Eingetretene Leitzinsentscheidungen
actual <-
factor(c('None','Hike','Hike','Hike','Hike','Hike','Hike','Hike','Hike','Hike','None','Cut','Cut','Cut','Hike','None','Hike','
Hike','Hike','None','None','None','None','None','None','None','None','None','None','None','None','None','None','N
one','None','None','None','None','None'))

# Vorhergesagte Leitzinsentscheidungen
predicted <-
factor(c('Cut','Hike','Hike','Hike','Hike','Hike','Hike','Hike','Hike','Hike','None','Hike','None','None','None','None','Hike','N
one','None','None','None','None','None','None','None','None','None','None','None','None','None','None','None','None','N
one','None','None','None','None','None'))

# Erstellen der Konfusionsmatrix
confMatrix <- confusionMatrix(predicted, actual)
Metrics <- as.data.frame(confMatrix$byClass)

print(confMatrix$table)

# Berechnung und Ausgabe der Metriken: Recall (Sensitivity), Precision, BalancedAccuracy
recall <- Metrics["Sensitivity"]; recall
precision <- Metrics["Precision"]; precision
BalancedAccuracy <- Metrics["Balanced Accuracy"]; BalancedAccuracy
```


Anlage 2: Screenshot Antwort ChatGPT-4 datenbasierter Ansatz per FOMC Sitzung am 2022-09-21

ChatGPT 4 ▾



ChatGPT

Leitzinsen könnten im Verlauf der nächsten sechs Monate weiter erhöht werden

Angesichts der anhaltend hohen Inflation und der konsequenten Zinserhöhungen durch die Federal Reserve in den letzten Sitzungen ist es wahrscheinlich, dass die Fed ihre Politik der Zinserhöhungen fortsetzen wird, um die Inflation zu bekämpfen und die Wirtschaft zu stabilisieren. Trotz einiger Schwankungen in den Wirtschaftsindikatoren deuten die anhaltend hohen Inflationsraten und die stabilen Arbeitsmarktdaten darauf hin, dass weitere Maßnahmen erforderlich sind, um die Preisstabilität zu gewährleisten.

Volkswirtschaftliche Kennzahl	Trendbewertung	Bewertung für Zinsentscheidung
Consumer Price Inflation (CPI)	Hoch, leicht fallend	Erhöhung
Consumer Price Inflation ex Food and Energy	Hoch, stabil	Erhöhung
Personal Consumption Expenditure Inflation (PCE)	Leicht fallend	Beibehaltung
Personal Consumption Expenditure Inflation ex Food and Energy	Hoch, stabil	Erhöhung
Non-Farm-Payrolls	Leicht fallend	Beibehaltung
US-Arbeitslosenquote	Niedrig, stabil	Beibehaltung
S&P500	Schwankend	Beibehaltung

Die konsistente Erhöhung des Leitzinses in den letzten Sitzungen, gepaart mit der anhaltenden hohen Inflation, deutet darauf hin, dass die Federal Reserve weiterhin geneigt sein wird, die Leitzinsen zu erhöhen, um die Inflation zu kontrollieren und langfristig ein nachhaltiges Wirtschaftswachstum zu fördern.