

1. Ask a same question both in ChatGPT and Gemini, then compare the results.

Frage

Was sind die Aufgaben eines Data Scientists?

ChatGPT:

Die Aufgaben eines Data Scientists umfassen typischerweise:

1. **Datenbeschaffung und -bereinigung:** Sammeln von Daten aus verschiedenen Quellen und Reinigen der Daten, um sicherzustellen, dass sie für die Analyse geeignet sind.
2. **Datenanalyse und -exploration:** Anwenden von statistischen Methoden, um Muster, Trends und Zusammenhänge in den Daten zu identifizieren.
3. **Modellierung und Vorhersage:** Entwicklung von Datenmodellen und Algorithmen, um Vorhersagen über zukünftige Ereignisse oder Trends zu treffen.
4. **Maschinelles Lernen und künstliche Intelligenz:** Anwendung von maschinellen Lern- und KI-Techniken, um automatisierte Entscheidungen zu treffen oder Muster in den Daten zu erkennen, die für menschliche Analysten schwer zu erkennen wären.
5. **Visualisierung und Kommunikation:** Präsentation der Analyseergebnisse mithilfe von Visualisierungen und Berichten, um komplexe Daten für nicht-technische Stakeholder verständlich zu machen.
6. **Optimierung und Verbesserung:** Kontinuierliche Überprüfung und Verbesserung von Modellen und Prozessen, um bessere Vorhersagen und Ergebnisse zu erzielen.
7. **Ethik und Datenschutz:** Berücksichtigung ethischer Grundsätze und Sicherstellung der Einhaltung von Datenschutzrichtlinien bei der Verwendung von Daten.
8. **Zusammenarbeit mit anderen Teams:** Zusammenarbeit mit Teams wie Softwareentwicklung, Marketing, und Produktmanagement, um Datenstrategien zu entwickeln und umzusetzen, die die Geschäftsziele unterstützen.

Die genauen Aufgaben können je nach Organisation und spezifischem Projekt variieren, aber im Allgemeinen sind Data Scientists für den gesamten Lebenszyklus der Datenanalyse verantwortlich, von der Datenerfassung bis zur Bereitstellung von Erkenntnissen und Empfehlungen.

Gemini:

Die Aufgaben eines Data Scientists im Überblick:

Datenerhebung und -vorbereitung:

Identifizieren und Sammeln relevanter Daten aus verschiedenen Quellen (z.B. Datenbanken, Sensoren, Social Media)

Bereinigen und Transformieren von Daten um Fehler und Inkonsistenzen zu beseitigen

Zusammenführen und Integrieren von Daten aus verschiedenen Quellen

Erstellen von explorativen Datenanalysen um die Datenstruktur und Merkmale zu verstehen

Datenanalyse und Modellierung:

Anwenden statistischer und maschineller Lernverfahren um Muster und Trends in den Daten zu erkennen

Entwickeln von prädiktiven Modellen um zukünftige Ereignisse oder Trends vorherzusagen

Evaluieren und Optimieren von Modellen um deren Genauigkeit und Leistung zu verbessern

Interpretieren von Modellergebnissen und Kommunizieren der Erkenntnisse an Stakeholder

Kommunikation und Präsentation:

Visualisieren von Daten und Analyseergebnissen in verständlichen Formaten (z.B. Diagramme, Dashboards)

Präsentieren von Ergebnissen und Erkenntnissen an Fach- und Nicht-Fachpublikum

Verfassen von Berichten und Dokumentationen der Analyseprozesse und Ergebnisse

Zusammenarbeiten mit anderen Teams um datenbasierte Entscheidungen zu unterstützen

Zusätzliche Aufgaben:

Bleiben auf dem Laufenden mit neuen Entwicklungen in den Bereichen Data Science, Statistik und maschinelles Lernen

Entwickeln neuer Analysemethoden und Algorithmen zur Lösung spezifischer Probleme

Fördern einer datenbasierten Kultur im Unternehmen durch Schulungen und Beratungen

Ethische Aspekte der Datennutzung und des Datenschutzes beachten

Die konkreten Aufgaben eines Data Scientists können je nach Branche, Unternehmen und Projekt variieren.

Einige Beispiele für Anwendungsgebiete von Data Science:

Marketing: Kundenverhalten analysieren, Marketingkampagnen optimieren, neue Produkte und Dienstleistungen entwickeln

Vertrieb: Risikomanagement verbessern, Kundenakquise und -bindung steigern, Preisgestaltung optimieren

Finanzen: Betrugsbekämpfung, Risikomanagement, Portfoliooptimierung

Produktion: Predictive Maintenance, Optimierung von Produktionsprozessen, Qualitätskontrolle

Personalwesen: Mitarbeiterzufriedenheit analysieren, Fluktuationsraten reduzieren, Talent Management verbessern

Data Science ist ein wachsendes Feld mit vielfältigen Einsatzmöglichkeiten. Die Nachfrage nach qualifizierten Data Scientists ist hoch und die Karriereaussichten sind ausgezeichnet.

Vergleich:

Zuerst zu den Gemeinsamkeiten der beiden künstlichen Intelligenzen. In beiden Varianten kann die gegebene Antwort als „inkorrekt“ oder „schlecht“ bewertet werden. Beide KI's bieten einen Button zum sofortigen Kopieren der Antwort, oder zum Neugenerieren einer Antwort. Gemini liegt jedoch im Punkt Benutzerfreundlichkeit und Usability weiter vorne und stellt eine Vielzahl von Funktionen bereit, die jedenfalls in ChatGPT's Gratis-Version nicht vorhanden zu sein scheinen.

Beispielsweise kann zu Beginn 1 von 3 möglichen Antworten gewählt werden. Diese unterscheiden sich in Struktur, Sprache sowie Inhalt. Auch das Generieren der Antworten geschieht in einer viel schnelleren Zeitspanne. Über einen Schieberegler lassen sich Antworten auch individuell anpassen, darunter Optionen zum Einschränken der Länge der Antwort, oder Optionen zum Einstellen des Schreibstils bzw. der Sprache der KI. Antworten können ebenfalls über einen Button sofort geteilt, exportiert oder per Mail versendet werden.

Doch die wohl praktischste Funktion ist das automatische Überprüfen und Bewerten der jeweiligen Antwort. Hier sucht Gemini automatisch auf Google nach den produzierten Inhalten der generierten Antwort. Gibt es Übereinstimmungen, werden Inhalte grün markiert. Werden Abweichungen entdeckt oder keine relevanten Inhalte gefunden, so werden Passagen auffällig orange markiert. Bei nicht hervorgehobenen Aussagen liegen entweder nicht genügend Informationen zur Bewertung vor, oder es liegen keine faktenbasierten Informationen vor.

Der große Vorteil hiervon ist, dass der Nutzer direkt nachvollziehen kann wie die KI die Antwort generiert hat und woher das Wissen extrahiert wurde. So können mögliche Fehler identifiziert und vermieden werden. ChatGPT bietet hier in der Gratis-Version keine Funktion zum Überprüfen von generierten Antworten, lediglich einen Hinweis, dass diese möglicherweise fehlerbehaftet sein können.

Auch der Inhalt unterscheidet sich ein wenig. Während ChatGPT einzig und alleine auf die Fragestellung eingeht, gehen die Antworten von Gemini auch weit über diese hinaus. Neben der bloßen Nennung der Tätigkeiten unterteilt Gemini diese ebenfalls in die verschiedenen zugehörigen Anwendungsgebiete (Marketing, Vertrieb, Produktion etc.). Auch das Hinweisen darauf, dass die

Nachfrage nach qualifizierten Data Scientists hoch ist und die Karrierechancen ausgezeichnet sind, lässt vermuten, dass Gemini Fragestellungen hinterfragt und potentiell interessante Fakten hinzufügt. ChatGPT kann hier zwar die gleiche Antwort liefern, muss dabei jedoch explizit danach gefragt werden. Gemini erwähnt ebenfalls dass Data Scientists die Aufgabe haben ständig auf dem Laufenden zu bleiben im Bereich ML, DS und Statistik. Auch hier geht die KI über die trivialen Grundtätigkeit hinaus und vertieft die Beantwortung der Frage.

Final schlägt Gemini weitere interessante Folgefragestellungen vor, die der Nutzer per Klick nahtlos fragen kann.

Ansonsten gibt es bei Gemini ebenfalls die Möglichkeit Bilder hochzuladen und diesbezügliche Fragen zu stellen. Zusätzlich kann der Nutzer über ein Mikrofon mit der KI sprechen.

Die Unterschiede sind deutlich und ich habe eine neue Lieblings-KI für mich entdeckt 😊

2. You are a CEO from a business/startup with around 50 employees. You are considering to hire one or several data positions, what kind of data employee are you going to hire and why. Please describe the business, product/service first, and explain the reason of hiring data experts. (you can also not hire anyone)

Unternehmen:

„NextGen Gaming“

Produkt/Dienstleistung:

NextGen Gaming ist eine Firma, die innovative Spieleplattformen und Technologien entwickelt, die die Grenzen des Gamings neu definieren. Das Unternehmen fokussiert sich auf die Integration von künstlicher Intelligenz in Spiele, um Spielern noch immersivere und realistischere Spielerlebnisse zu bieten, die auch individuell auf Spielertypen eingehen können.

Grund für die Einstellung:

Die Technologie von NextGen Gaming generiert eine immense Menge an Daten über verschiedene Aspekte wie Spielerverhalten, Spielerinteraktionen und Systemleistung. Damit die Produkte diese Daten nutzen können um die Technologie kontinuierlich zu verbessern, werden Data Analysts, Data Scientists sowie Data Engineers benötigt.

Data Analyst

Der Data Analyst analysiert Spieldaten um Präferenzen, Balance und Systemleistung zu optimieren.

Data Scientist

Der Data Scientist entwickelt ML-Modelle um Spielverhalten vorherzusehen, damit KI-gestützte Spielmechaniken verbessert werden (bessere Vorhersagen treffen).

Data Engineer

Der Data Engineer führt den Aufbau und die Wartung von Datenpipelines aus. Er kümmert sich um die Infrastruktur zur Speicherung, Verarbeitung und Visualisierung von Spieldaten.

Warum?

Datenmengen werden immer größer und gerade bei erfolgreichen Videospielen kann die Anzahl der Spieler sehr stark steigen. Wenn man dann noch berücksichtigen muss, dass jeder Spieler individuelle Spielweisen und Präferenzen an den Tag legt, wird deutlich, dass Experten für das Analysieren und Auswerten von Rohdaten benötigt werden.

Im Falle von NextGen Gaming werden alle 3 Rollen benötigt. Der Data Analyst verschafft sich besonders einen Überblick über die ganzen Daten. Diese müssen auch effektiv kommuniziert und visualisiert werden. Der Data Scientist kann dadurch effizientere ML-Modelle implementieren, die einen größeren Bereich von verschiedenen Spielertypen abdecken. Hier werden spezielle Muster erkannt die zur Verbesserung des Produktes führen können. Data Engineers werden benötigt um an die verschiedenen repräsentativen Daten eines Videospiels zu gelangen. Auch eine Skalierbarkeit der Dateninfrastruktur muss gegeben sein, für den Fall, dass das Spiel an Popularität gewinnt.