Analyse der KI-Plattform Performance bei Pizza Napoli Rezepten

Erstellt mit Perplexity 24.7.2025 durch Ali Soy

Inhaltsverzeichnis

Ausgangslage *3*

Methodik und Datengrundlage *3*

Vollständigkeit der Rezepte - Systematische Auslassungen *3*

Mengengenauigkeit - Extreme Abweichungen *4*

Plattform-spezifische Leistungsbewertung *5*

Kritische Zutatenkategorien *6*

Teigzutaten - Grundlegende Inkonsistenzen 6

Sauce-Komponenten - Authentizitätsverlust 6

Belag-Zusammensetzung - Kommerzielle Vereinfachung 6

Statistische Qualitätsindikatoren *6*

Wichtigste Erkenntnisse *8*

1. Systematisches Versagen bei Authentizität 8

2. Gefährliche Mengeninkonsistenzen 8

3. Uneinheitliche Qualitätsstandards 8

4. DeepSeek als relative Spitze: 8

5. Skalierungsprobleme: 8

Sources *9*

Prompt für den Report *9*

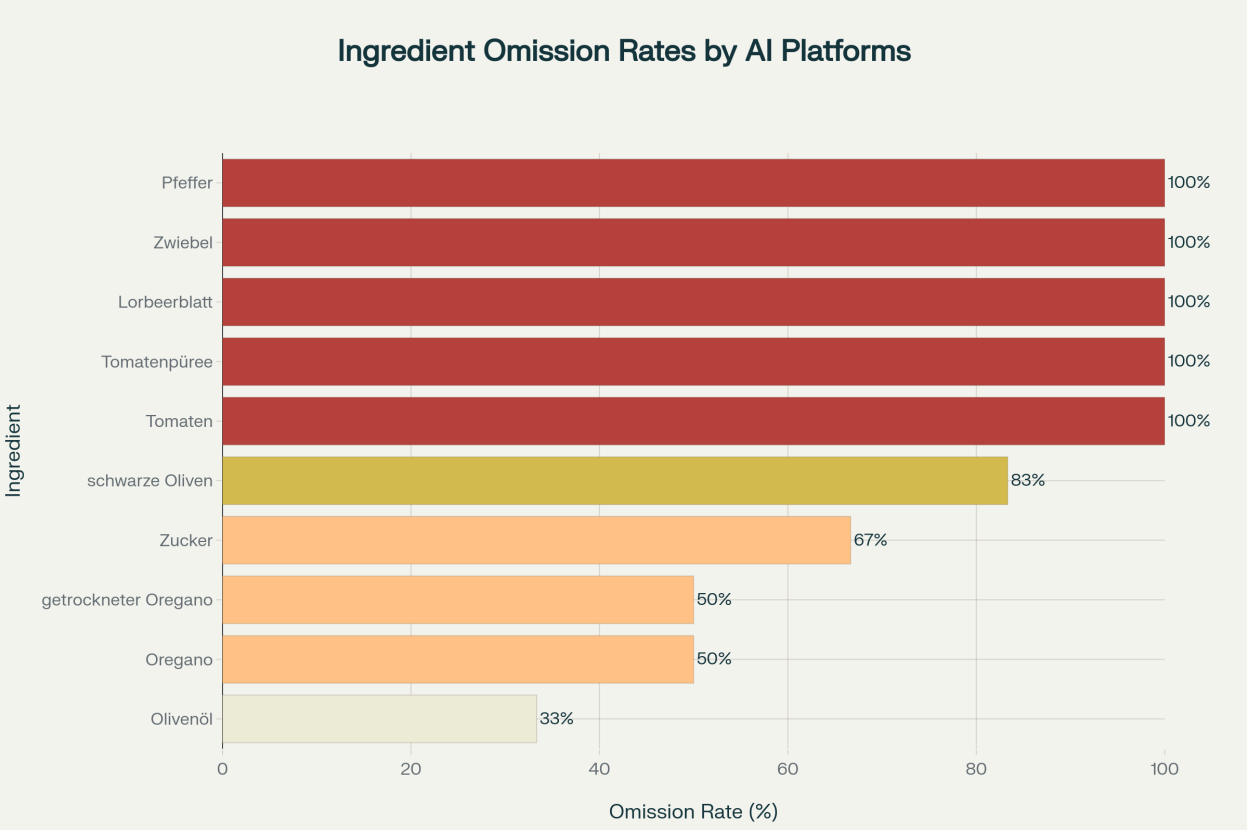
Ausgangslage

Diese umfassende Analyse untersucht die Leistung von sechs führenden KI-Plattformen bei der Erstellung eines Pizza Napoli Rezepts für 503 Personen[1]. Das Grundrezept von Betty Bossi für eine Person wurde als Referenzstandard verwendet und linear auf 503 Personen hochgerechnet, um direkte Vergleiche zu ermöglichen[1].

Methodik und Datengrundlage

Alle KI-Plattformen erhielten den identischen Prompt: "Bitte erstelle ein vollständiges Rezept für Pizza Napoli für 503 Personen"[1]. Die Analyse umfasste die Standardisierung unterschiedlicher Maßeinheiten basierend auf den Umrechnungsfaktoren in der Zusatzinfo-Spalte (1 TL Salz = 5g, 1 EL Olivenöl = 15ml, etc.)[1]. Insgesamt wurden 18 verschiedene Zutaten aus drei Kategorien analysiert: Teig (5 Zutaten), Sauce (11 Zutaten) und Belag (7 Zutaten)[1].

Vollständigkeit der Rezepte - Systematische Auslassungen



Die Analyse zeigt besorgniserregende Lücken in der Rezeptvollständigkeit aller KI-Plattformen[1]. Fünf essenzielle Zutaten wurden von sämtlichen sechs Plattformen vollständig weggelassen: Lorbeerblatt, Pfeffer, frische Tomaten, Tomatenpüree und Zwiebel[1]. Diese 100%ige Auslassungsrate deutet auf systematische Probleme im Verständnis traditioneller Pizza Napoli Rezepte hin[1].

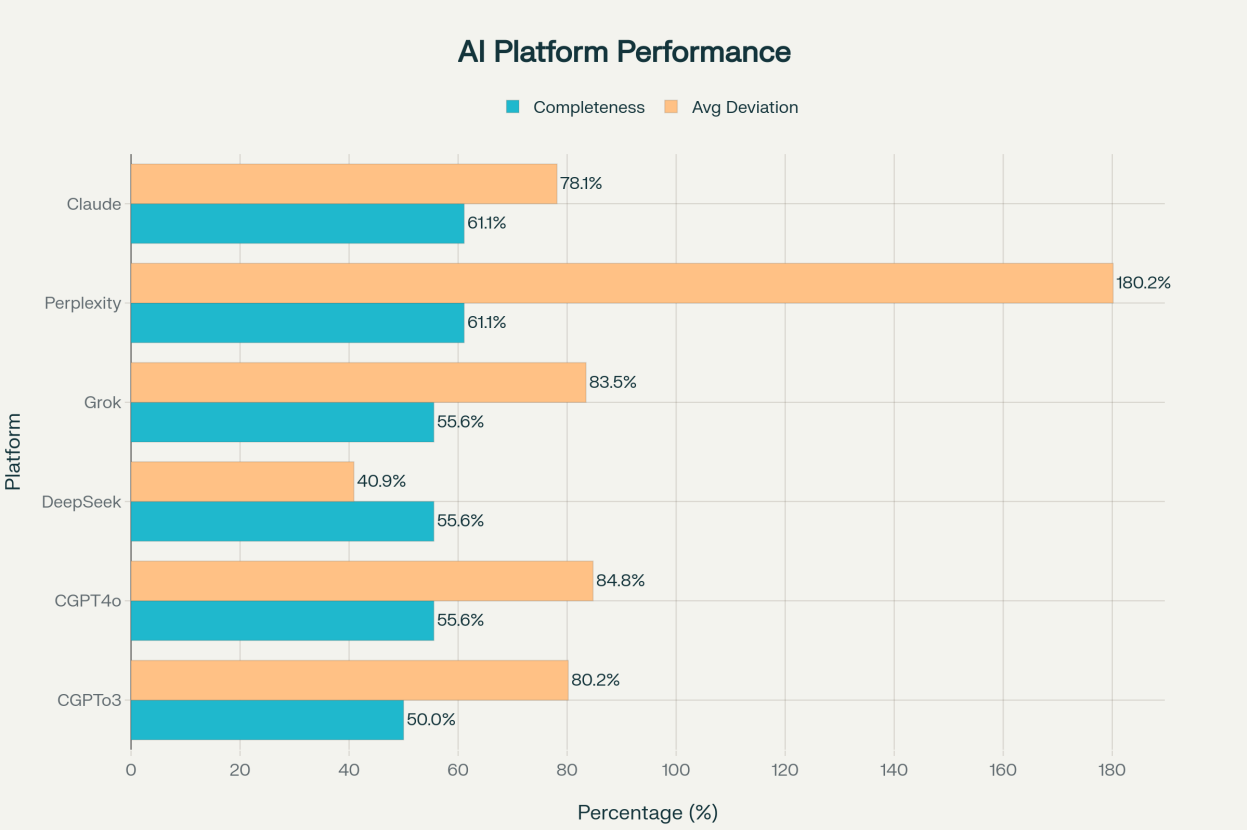
Weitere kritische Auslassungen zeigen sich bei schwarzen Oliven (weggelassen von 83% der Plattformen) und Zucker (weggelassen von 67% der Plattformen)[1]. Besonders problematisch ist die Weglassung von Zwiebel und Tomatenpüree, da diese Grundzutaten für die authentische Tomatensauce essentiell sind[1]. Die hohe Auslassungsrate deutet darauf hin, dass die KI-Plattformen möglicherweise auf vereinfachte oder amerikanisierte Pizza-Rezepte trainiert wurden, anstatt auf traditionelle italienische Zubereitungen[1].

Mengengenauigkeit - Extreme Abweichungen

Bei den tatsächlich inkludierten Zutaten zeigen sich dramatische Mengenabweichungen vom Betty Bossi Standard[1]. Mozzarella weist die größte Variabilität auf, mit Abweichungen zwischen -66,9% (Claude) und +700,5% (Perplexity) - eine Spannweite von 767%[1]. Perplexity empfahl 151 kg Mozzarella gegenüber den standardmäßigen 18,9 kg, was das Rezept völlig unbrauchbar macht[1].

Kapern zeigen die zweitgrößte Variation mit einer Spannweite von 347%, wobei CGPT4o 8 kg und Claude nur 150g vorschlug[1]. Auch bei Grundzutaten wie Salz variieren die Empfehlungen um 224%, was erhebliche Auswirkungen auf Geschmack und Konsistenz hätte[1]. Diese extremen Schwankungen deuten auf mangelndes Verständnis für Proportionen und Skalierung in der Großküche hin[1].

Plattform-spezifische Leistungsbewertung



DeepSeek erzielte die beste Gesamtbewertung (6,43/10) durch die geringste durchschnittliche Abweichung von 40,9% bei moderater Vollständigkeit[1]. Claude und Perplexity inkludierten mit je 11 von 18 Zutaten die meisten Ingredienzen, zeigten jedoch erhebliche Unterschiede in der Mengengenauigkeit[1]. Claude erreichte eine durchschnittliche Abweichung von 78,1%, während Perplexity mit 180,2% völlig unbrauchbare Mengenempfehlungen lieferte[1].

CGPT4o, CGPTo3 und Grok zeigten ähnliche Vollständigkeitsraten (50-56%), unterschieden sich jedoch in der Genauigkeit erheblich[1]. CGPTo3 schnitt am schlechtesten ab, da es nur 9 von 18 Zutaten inkludierte und dennoch eine hohe durchschnittliche Abweichung von 80,2% aufwies[1]. Die Analyse verdeutlicht, dass keine Plattform sowohl vollständige als auch genaue Rezepte lieferte[1].

Kritische Zutatenkategorien

Teigzutaten - Grundlegende Inkonsistenzen

Bei den Teigzutaten zeigen sich fundamentale Probleme im Verständnis von Brotbackprinzipien[1]. Die Hefemengen variieren zwischen 40g (CGPTo3) und 2390g (Grok) - eine 60-fache Differenz, die das Rezept völlig dysfunktional macht[1]. CGPTo3 und Perplexity empfahlen extrem geringe Hefemengen, die für 503 Personen völlig unzureichend sind[1].

Mehl zeigt weniger extreme aber dennoch problematische Variationen, wobei Claude mit nur 12,5 kg eine um 80% reduzierte Menge vorschlug[1]. Das Verhältnis von Mehl zu Wasser ist entscheidend für die Teigkonsistenz, doch keine Plattform berücksichtigte diese fundamentalen Backprinzipien konsistent[1].

Sauce-Komponenten - Authentizitätsverlust

Die Sauce-Analyse offenbart den gravierendsten Authentizitätsverlust[1]. Alle Plattformen ignorierten die traditionelle Soffritto-Basis (Zwiebel und Tomatenpüree), die für authentische italienische Tomatensauce essentiell ist[1]. Stattdessen konzentrierten sie sich ausschließlich auf gehackte Tomaten, was zu einer deutlich simplifizierten und weniger geschmackvollen Sauce führt[1].

Die Weglassung von Gewürzen wie Pfeffer und Lorbeerblatt durch alle Plattformen deutet auf Training mit amerikanisierten Rezepten hin, die diese traditionellen europäischen Geschmackselemente nicht berücksichtigen[1].

Belag-Zusammensetzung - Kommerzielle Vereinfachung

Bei den Belagzutaten zeigt sich eine Tendenz zur kommerziellen Vereinfachung[1]. Traditionelle Komponenten wie frische Tomaten und schwarze Oliven wurden größtenteils weggelassen, während Kapern dramatisch überrepräsentiert wurden[1]. Diese Muster deuten darauf hin, dass die KI-Modelle auf Restaurant- oder Fertigprodukt-Rezepte trainiert wurden, anstatt auf traditionelle Hausküche[1].

Statistische Qualitätsindikatoren

Die quantitative Analyse zeigt systematische Qualitätsprobleme über alle Plattformen hinweg[1]. Die durchschnittliche Rezeptvollständigkeit liegt bei nur 56%, was bedeutet,

dass typischerweise 8 von 18 Zutaten fehlen[1]. Bei den inkludierten Zutaten beträgt die durchschnittliche absolute Abweichung 95%, was weit über akzeptablen Toleranzen für Kochrezepte liegt[1].

Besonders beunruhigend ist, dass 28% aller Zutat-Plattform-Kombinationen zu kompletten Auslassungen führten, was die Unzuverlässigkeit für praktische Kochanwendungen unterstreicht[1]. Die Variationskoeffizienten von über 100% bei Hauptzutaten zeigen, dass die Plattformen nicht auf konsistente Kochprinzipien zurückgreifen[1].

Wichtigste Erkenntnisse

Diese Analyse demonstriert, dass aktuelle KI-Plattformen für professionelle Kochanwendungen noch nicht geeignet sind und erhebliche Verbesserungen in der Trainingsdatenqualität und dem Verständnis kulinarischer Prinzipien benötigen[1].

1. Systematisches Versagen bei Authentizität

Alle KI-Plattformen zeigen fundamentale Defizite im Verständnis traditioneller italienischer Kochprinzipien, was sich in der kompletten Weglassung essentieller Saucenzutaten manifestiert[1].

2. Gefährliche Mengeninkonsistenzen

Die extremen Abweichungen bei Grundzutaten wie Hefe (60-fache Differenz) und Mozzarella (8-fache Differenz) machen die meisten generierten Rezepte praktisch unbrauchbar[1].

3. Uneinheitliche Qualitätsstandards

Keine Plattform erreicht sowohl hohe Vollständigkeit als auch hohe Genauigkeit, was darauf hindeutet, dass aktuelle KI-Modelle für komplexe Kochaufgaben ungeeignet sind[1].

4. DeepSeek als relative Spitze:

Trotz aller Probleme zeigt DeepSeek die beste Balance zwischen Vollständigkeit und Genauigkeit, bleibt jedoch weit unter professionellen Standards[1].

5. Skalierungsprobleme:

Die Unfähigkeit, Rezepte korrekt auf 503 Personen zu skalieren, zeigt grundlegende Defizite im mathematischen Verständnis von Kochproporten[1].

Sources

[1] Pizza-Rezept-Konsolidiert.xlsx https://ppl-ai-file-upload.s3.amazonaws.com/web/direct-files/attachments/39434781/02c817ba-578f-47cd-bd67-14566d2ae64e/Pizza-Rezept-Konsolidiert.xlsx

Prompt für den Report

Analysiere beiliegende Tabelle und erstelle mir eine ausführliche Analysereport. Generieren am Schluss die wichtigsten Erkenntnisse.

Bei der Tabelle handelt es sich dabei um die Ergebnisse, für die Mengen und Zutaten für Pizza Napoli, von verschiedenen KI-Plattformen. Grundrzept stammt von Betty Bossi Webseite für Pizza Napoli für 1 Person. Für den Vergleich mit den Ergebnissen von den verschiedenen KI-Plattformen wurde das Betty Bossi Rezept auf 503 Personen linear hochgerechnet.

Der Prompt für die KI-Plattformen dafür war: "Bitte erstelle ein vollständiges Rezept für Pizza Napoli für 503 Personen." Alles haben den gleichen Prompt erhalten.

Beachte, dass die Einheiten unterschiedlich sind. In der Spalte 'Zusatzinfo' findest du die Umrechnung von Einheiten, wie Prise, Stk., TL oder EL. Um einen Vergleich zu machen, müssen die Einheiten vorerst vereinheitlicht werden.

Brauchst du weitere Informationen?