Heart-a-Tech Podcast by MaibornWolff

Folge 02:

Viel mehr als Prompts:
Wie LLM funktionieren und was du damit machen kannst

Mit Experte Fabian Hertwig

Intro Fabian Hertwig: Diese Technologie gibt es noch nicht so lange. Aber man sieht durch das, was sie kann, dass sie auf jeden Fall enormes Potenzial hat.

Brigitte Streibich: Herzlich willkommen bei Heart-a-Tech, dem Podcast rund um alles, was du wissen musst, um neue IT-Trends und Technologien erfolgreich in deinem Unternehmen zu implementieren. Mein Name ist Brigitte Streibich und wir sprechen heute über Prompts, wie LLM funktioniert und was du damit machen kannst.

Der Titel verspricht schon viel und ich habe mir heute wieder einen ganz besonderen Gast eingeladen, der uns ein bisschen die Welt der künstlichen Intelligenz erklären wird. Es ist Fabian Hertwig. Fabian ist studierter Informatiker, kennt sich gut mit Machine Learning und Data Analytics aus, hat im Bereich Data Science die Projektleitung bei verschiedenen Projekten, beschäftigt sich aktuell hauptsächlich mit technischen Kompetenzen und leite Teams, wenn es um das Thema Large Language Models geht. Hallo Fabian, herzlich willkommen!

Fabian Hertwig: Hallo.

Brigitte Streibich: Fabian, du hast mir im Vorgespräch erzählt, dass du dich selbst so als Typ Explorer bezeichnen würdest.

Fabian Hertwig: Genau, ja.

Brigitte Streibich: Was darf ich mir denn darunter vorstellen? Und vor allem, wie wirkt sich das auf deine Arbeit im Bereich Künstliche Intelligenz aus?

Fabian Hertwig: Ich habe vor Kurzem gelernt, dass es verschiedene Typen von Menschen gibt, für die Arbeit wie Spaß ist. Einer von diesen Typen ist der Explorer, der auch ich bin, und diese Menschen macht es besonders Spaß, neue Dinge zu erfahren, neue Dinge zu lernen. So ist es auch stark bei mir. Deswegen bin ich in dem AI-Bereich ganz stark aktiv und interessiert, denn da gibt es immer neue Sachen zu entdecken. Es entwickelt sich ständig weiter und ständig können Probleme gelöst werden, die davor noch nicht gelöst werden konnten. Und ja, es gibt immer was Neues zu explorieren.

Brigitte Streibich: Ja, perfekt, da bist du wirklich sehr gut aufgehoben, und ich glaube, da ist auch in den kommenden Jahren kein Ende in Sicht. Da wird es noch viel zu entdecken geben. Jetzt ist es so, dass man in der öffentlichen Diskussion ganz viel mitbekommt von Open AI, vor allem jetzt mit dem ganzen ChatGPT-Thema. Man spricht viel über Large Language Models, abgekürzt LLM, aber so richtig verstehen, was dahintersteckt tun die wenigsten. Magst du uns mal eine kurze Einordnung geben, was die Unterschiede zwischen diesen Begrifflichkeiten sind, vor allem auch im Unternehmenskontext?

Fabian Hertwig: Fangen wir mit OpenAI an. Das ist ein Research-Unternehmen, das sich vorgenommen hat, starke Künstliche Intelligenz zu erforschen, mit dem Ziel, diese wirklich zu erschaffen. Aber das auf einem sicheren Weg, sodass die AI uns nicht alle umbringt und alle unsere Jobs eliminiert, sondern so dass es im Einklang mit der Gesellschaft ist. Die haben ChatGPT vor ein paar Monaten veröffentlicht.

GPT steht für Generative Pre Train Transformer. Das ist eine bestimmte Modellarchitektur von Large Language Models. Wenn wir bei Large Language Models anfangen: Das sind neuronale Netze, die davon inspiriert sind, wie das Gehirn funktioniert. Aber nur bei den Neuronen, nicht im Aufbau oder wie das Gehirn lernt. Ein Language Model ist ein neuronales Netz, das das nächste Wort vorhersagen kann. Beginnt man einen Satz, zum Beispiel: "Der Fuchs springt über…", kann es vorhersagen: Die nächsten Worte sind wahrscheinlich "…den Zaun". Diese Modelle sind auf jeden Text trainiert, den man bekommen kann. Im Prinzip wird das ganze Internet gescrabt und diesem Modell als Training gegeben, einfach nur um zu lernen, was das nächste Wort ist, das man vorhersagen muss. Dadurch lernt das Modell viel über die Sprache, wie sie aufgebaut ist, über Grammatik und auch über verschiedene Sprachen. Es lernt aber auch viele Dinge über die Welt, also wie Dinge in dieser Welt funktionieren, etwa Physik.

Je größer diese Modelle sind, also je mehr von diesen Neuronen sie haben, desto besser funktionieren sie, desto besser können sie die Wörter vorhersagen und desto mehr lernen sie über die Welt.

Bei ChatGPT ist der große Trick, das nicht einfach nur das nächste Wort vorhergesagt wird. Das Modell wurde trainiert, Instruktionen zu beantworten. Es wurden Datensätze erstellt, bei denen Menschen Instruktionen bekommen haben. Zum Beispiel: "Fasse diesen Wikipedia-Artikel zusammen." Das haben sie gemacht, das wurde dem Modell als Training gegeben und dadurch hat es gelernt, nicht einfach nur das nächste Wort in dem Text vorherzusagen, sondern wirklich Aufgaben zu bearbeiten.

Brigitte Streibich: So wie es ein Mensch tun würde.

Fabian Hertwig: Genau!

Brigitte Streibich: Und diese Instruktionen, das sind dann in dem Fall Prompts.

Fabian Hertwig: Genau. Das nennt man Prompts.

Brigitte Streibich: Das heißt das, was wir so in der Konsumerwelt als ChatGPT kennen, das ist im Prinzip nur ein Tool, das das Thema anwendet oder auf dieser Technologie basiert.

Fabian Hertwig: Genau, das ist ein bestimmtes Modell, was auf der GPT-Architektur basiert und eben auch ein Large Language Model ist.

Brigitte Streibich: Und wenn wir das jetzt mal so in den Unternehmenskontext heben, also beispielsweise dorthin, wo ihr mit euren Kunden unterwegs seid. Wie würde GPT ganz konkret im Unternehmen zur Anwendung kommen?

Fabian Hertwig: Das lernen wir gerade mit unseren Kunden zusammen. Diese Technologie gibt es noch nicht so lange. Aber man sieht durch das, was sie kann, dass sie auf jeden Fall enormes Potenzial hat. Dieses Modell hat schon viel Wissen über die Welt - teilweise auch Expertenwissen. Ich frage oft Sachen zu Softwareengineering und kriege da Antworten auf einem Expertenlevel zurück. Es ist einfach extrem hilfreich, diesen Experten immer an seiner Seite zu haben.

Zum anderen kann das Modell auch sehr gut mit Text umgehen. Es kann Text verstehen, zusammenfassen, erweitern oder von einer Form in eine andere bringen. Oder es übersetzt von einer Sprache in eine andere, von einer natürlich sprachlichen Aufgabe in Code oder auch Code in eine natürlich sprachliche Beschreibung. Das ist einfach enorm nützlich. Das kann man, glaube ich, in fast allen Unternehmensbereichen einsetzen.

Brigitte Streibich: Wie verlässlich sind diese Daten? Man hört, dass manchmal gerade ChatGPT nicht unbedingt die besten Quellen nutzt oder sogar selbst irgendwie Quellen erfindet, wenn die KI nicht weiterweiß. Wenn man ein Referat in seinem Studium vorbereitet, ist das nicht so wichtig. Aber wenn man mit sensiblen Unternehmensdaten umgeht, wenn es um große Summen geht oder vielleicht sogar um Deals, die man mit einem Kunden abschließen möchte, dann ist natürlich so eine Verlässlichkeit und Vertraulichkeit wichtig. Wie schätzt du da die Lage ein?

Fabian Hertwig: Das nennt man dann Halluzinationen, wenn das Modell was ausgibt, was richtig klingt, aber nicht richtig ist.

Brigitte Streibich: Wie eine Halluzination halt so funktioniert.

Fabian Hertwig: Genau, das ist noch eins der großen Probleme bei diesen Modellen. Man kann den Outputs nicht unbedingt vertrauen und sollte immer von einem Menschen noch überprüft werden, vor allem, wenn es irgendwie um sicherheitsrelevante Dinge geht.

Es gibt verschiedene Techniken, wie man diese Halluzinationen verringern kann. Zum Beispiel, wenn man noch extra Daten mit dazugibt in den Prompt, verringert es diese Halluzination, da

das Modell einfach mehr Kontextwissen hat. Man auch das Modell speziell prompten, sodass es eher sagt, "Ich weiß es nicht.", anstatt zu halluzinieren.

Brigitte Streibich: Was meistens wahrscheinlich besser wäre, als irgendwas Falsches auszuspucken.

Fabian Hertwig: Genau, aber ganz gelöst dieses Problem noch nicht und muss momentan auf jeden Fall extrem beachtet werden. In der Wissenschaft gibt es die ersten Anzeichen, dass es ein lösbares Problem ist und vielleicht passiert das in den nächsten Jahren nicht mehr.

Brigitte Streibich: Du kommst gerade aus einem Workshop, hast du mir vorhin erzählt. Ihr habt ein internes Projekt laufen, wo ihr auch schon das Thema LLM einsetzt oder versucht, das in unterschiedliche Bereiche hier bei MaibornWolff zu integrieren. Was habt ihr heute gemacht? Magst du mal ein bisschen erzählen?

Fabian Hertwig: Bei uns, bei MaibornWolff, sind wir in technische Bereiche aufgeteilt. Jeder Bereich beschäftigt sich mit einer Technologie oder mit einem Technologie-Schwerpunkt. In meinem Bereich sind wir die AI-Leute. Ein Problem, das wir immer haben: Wir kennen die Technologie sehr gut, wissen aber nicht automatisch, was die ganzen Probleme sind, die man damit lösen könnte. Und jetzt gibt es diese neue Technologie, die nicht nur unsere Kunden beschäftigt, sondern auch uns selbst und alle unsere Projekte. Deswegen haben wir eine Initiative gestartet, in der wir durch unser Unternehmen durch die verschiedenen technischen Bereiche gehen. Gemeinsam erarbeiten wir, wie man diese Technologie für die Probleme, die bei denen auftreten, am besten einsetzen kann.

Brigitte Streibich: Also, ihr nutzt euch selbst als Testlabor, um zu gucken, ob das gut funktioniert, und wenn es gut funktioniert, dann übertragt ihr es auch auf den Kunden.

Fabian Hertwig: Ja, so ungefähr. Wir überlegen, welche Probleme sind auch interessant für unsere Kunden und versuchen diese in Projekten anzugehen, sodass sofort ein Mehrwert für das Projekt entsteht. Gleichzeitig lernen wir, was man machen kann und was nicht so gut funktioniert.

Brigitte Streibich: Und wo steht ihr da gerade? Was habt ihr heute gemacht?

Fabian Hertwig: Heute war unser erster Workshop, bei dem wir uns die Probleme, die es so gibt, angeschaut und priorisiert haben. Dann haben wir uns Lösungen überlegt und diese auch wieder priorisiert.

Brigitte Streibich: Arbeitet ihr mit Design Thinking Techniken oder Ideation Techniken?

Fabian Hertwig: Design Thinking ist das Format, was wir heute verwendet haben. Bei uns im Team gibt es eine Digital Designerin und die hat uns heute sehr gut durch moderiert.

Brigitte Streibich: Wir haben vorher drüber gesprochen, dass die Daten, die so ein Tool ausspuckt, vielleicht nicht immer ganz verlässlich sind auf der inhaltlichen Ebene. Die Frage ist natürlich auch, wie vertraulich diese Daten sind. Inwiefern kann man sich sicher sein, dass die Daten und die Chats geschützt sind, dass das alles irgendwie im Unternehmen verbleibt? Beschäftigt ihr euch damit auch?

Fabian Hertwig: Ja, damit beschäftigen wir uns sehr stark. Sowohl für unsere Kunden als auch für uns ist das super wichtig. OpenAI stellt diese Modelle bereit, aber auch Microsoft hat in OpenAI investiert und kann daher die Modelle auf Azure bereitstellen. In der Azure Cloud gibt es bereits Datensicherheitsversprechen, und die gelten auch für die Modelle, die dort bereitgestellt werden. Deswegen ist momentan unsere Empfehlungen, die Azure Cloud zu verwenden, wenn es geht. Früher hat OpenAI für Trainingszwecke Daten gespeichert, wenn man mit der API interagiert hat. Benutzt man aber das Webinterface, das normale ChatGPT …

Brigitte Streibich: Das Öffentliche.

Fabian Hertwig: … dann werden dort für Trainingszwecke normalerweise Daten gespeichert. Man kann auch ein Opt-out machen, aber die meisten Leute wissen das nicht. Deswegen sollte man sehr darauf achten, dass man dort nicht interne Daten eingibt und damit chattet.

Brigitte Streibich: Zum Thema Sicherheit werden wir noch mal in einer anderen Folge sprechen. Wir haben einen Gast eingeladen, mit dem wir tiefer reingehen in die Themen Datenschutz und Datensicherheit. Du hast gerade schon das Thema Azure angesprochen. Azure und GPT spielen ja auch irgendwie zusammen. Ihr habt dazu auch ein Whitepaper veröffentlicht. Da kann man noch mal alles ganz genau nachlesen und auch der Punkt "Was sind eigentlich die Unterschiede zwischen GPT und LLMU?" ist da noch mal erklärt. Das packen wir den Zuhörerinnen und Zuhörern auch gerne in die Shownotes rein, dass sie das schwarz auf weiß nachlesen können.

Aber wenn wir gerade drüber sprechen: Azure und GPT - wie ist da das Zusammenspiel, gerade wenn es um den Unternehmenskontext geht, also mal raus aus diesem klassischen OpenAI und ChatGPT, was ich eigentlich auch als Privatperson irgendwie machen kann. Wie funktioniert das Ganze, wenn man das in den Businesskontext hebt?

Fabian Hertwig: Also in Azure werden die Modelle als API bereitgestellt. Das heißt, man kann einfach aus Software heraus mit diesen Modellen kommunizieren. Azure übernimmt für einen das ganze Hosting und Scaling und die Operations. Darum muss man sich zum Glück nicht kümmern. Man kann einfach Anfragen an das Modell schicken und kriegt Antworten zurück. Wenn man schon in der Cloud bei Azure ist, macht es das einfach, damit zu interagieren. Aber auch wenn man in einer anderen Cloud ist, zum Beispiel ABS, lässt sich die API aufrufen und man interagiert normal über das Internet.

Brigitte Streibich: Und dann kann man theoretisch eigentlich loslegen.

Fabian Hertwig: Genau, das ist relativ einfach. Man muss allerdings von Microsoft Zugriff auf die Modelle bekommen. Ich glaube, momentan bekommen den nur ausgewählte Partner und Use Cases. Man kann sich die deployen und hat dann den APN-Point und einen APT und kann anfangen, dort Anfragen zu stellen.

Brigitte Streibich: Wenn ihr kommt, als Team, als MaibornWolff, was könnt ihr dem Kunden noch für einen Mehrwert bieten, wenn er sowieso eigentlich schon mit Azure und GPT arbeitet?

Fabian Hertwig: Wir fangen an bei der Ideation dafür haben wir verschiedene Workshop Formate. Einer ist vier Stunden lang, wo wir erst mal erklären: Was sind Large Language Models, was kann man damit machen, was sollte man nicht damit machen und was muss man beachten? Im weiteren Workshop geht es um Ideation, also wie kann ich das im Unternehmen einsetzen? Was für Probleme habe ich dort, wo diese Technologie gut drauf passt?

Und dann gehts schon ins Prototyping oder in MVPs erstellen, also so schnell wie möglich Tools erstellen, die nützlich sind für die Mitarbeiter. Wenn es schon darüber hinausgeht, dann haben wir natürlich unser Standard Software Engineering Angebot, wo wir Featureteams bereitstellen, die die ganze Software drum herum bauen können.

Brigitte Streibich: Jetzt kann ich die unterschiedlichen Begrifflichkeiten auseinanderhalten. Wenn wir jetzt über Use Cases nachdenken: also wie kommt das alles ganz konkret zur Anwendung in Unternehmen? Kannst du da vielleicht schon einen kleinen Ausblick geben? Was wird denn möglich sein in Zukunft oder vielleicht schon sehr bald?

Fabian Hertwig: Der interessanteste Use Case für die meisten Unternehmen ist der "Chat with your data" Use Case. Der Nutzer kann Anfragen an das Modell schicken und das Modell sucht dann in den Unternehmensdokumenten, wo die Antwort steht, holt sich die Texte und formuliert damit die Antwort. Das ist extrem hilfreich. Man spart sich zum einen das Suchen durch teilweise verschiedene Dokumentquellen – das ist schon schwer genug - und zum anderen wird genau die Frage, die man hat, beantwortet. Man findet nicht nur eine Textstelle, die ungefähr dazu passen könnte.

Brigitte Streibich: Das ist natürlich auch ein großer Effizienzgewinn, nehme ich an.

Fabian Hertwig: Ja, auf jeden Fall.

Brigitte Streibich: Super Fabian, vielen Dank. Du hast mir auf jeden Fall einen sehr guten Einblick gegeben. Ich verstehe jetzt schon viel mehr, wie die Technologien aufeinander aufbauen und wie die auch tatsächlich dann zur Anwendung kommen. Jetzt sind wir ja hier im Podcast Heart-a-Tech, wenn du einen Herzenswunsch hättest oder etwas, was dein Herz höher schlagen lässt für die Zukunft von LLM im Unternehmenskontext, was wäre das? Was wünschst du dir?

Fabian Hertwig: Das ist eine schwierige Frage. Also ich glaube, LLMs haben momentan zwei Riesenprobleme. Das eine haben wir schon angesprochen: Halluzinationen. Das andere ist die Kontextgröße: Wie viel Text kann ich in den Prompt reinstecken oder kann das Modell auf einmal verarbeiten. Und ich glaube, wenn diese beiden Probleme gelöst sind, dann geht es noch mal richtig ab. Mein Herzenswunsch ist, dass die gelöst werden.

Brigitte Streibich: Das man da auf eine neue Stufe, auf ein neues Level kommt. Aber gerade wenn wir noch mal über die Prompts sprechen, habe ich doch noch eine allerletzte Frage. Müssen die Technologien trainiert werden, um die Prompts der Menschen besser zu verstehen? Oder geht es vielleicht auch darum, die Menschen zu trainieren, dass sie einfach ihre Prompts besser schreiben, dass das System sie verarbeiten kann?

Fabian Hertwig: Ganz sicherlich beides. Dass die Modelle die Menschen besser verstehen, darum kümmern sich solche Research Unternehmen wie OpenAi. Die gerade die Modelle so trainiert haben, dass die Instruktionen verstehen. Aber auch wie man die Modelle effektiv prompte ist so eine - ich würde jetzt nicht unbedingt Wissenschaft für sich sagen - aber man kann auf jeden Fall viel lernen. Manche behaupten, es wird bald den Job des Prompt Engineers geben, also Leute, die sich darauf spezialisiert haben, diese Modelle effektiv abzufragen.

Brigitte Streibich: Ich nutze selbst auch ab und zu ChatGpt und schreibe immer sehr höflich. Ich schreibe dann immer, "Könntest du bitte das und das machen", und irgendwann fällt mir dann auf, naja, eigentlich ist es eine KI. Ich muss jetzt hier nicht bitten und nicht höflich fragen. Vielleicht ist das ja auch mal eine Idee.

Fabian Hertwig: Ja, ich mache das lustigerweise auch. Und immer, wenn ich ab und zu mal im Befehlston schreibe, fühle ich mich schlecht.

Brigitte Streibich: Ja, genau, so geht mir tatsächlich auch. Gut, also dann mehr Menschlichkeit bei den Prompts. Damit würde ich sagen, vielen Dank, dass du bei uns warst, und alles Gute!

Fabian Hertwig: Sehr gerne, vielen Dank.