mimecast

Datenreihe: Die Zukunft der Arbeit

Fragen an KI-Anbieter

Bei Mimecast haben wir schon früh die Entscheidung getroffen, ein KI-natives Unternehmen zu sein. Diesen Leitfaden haben wir gemeinsam mit unseren Datenwissenschaftlern erstellt, damit Unternehmenseinkäufer KI-Anwendungsfälle für ihre Unternehmen eigenständig und ohne Fachkenntnisse im Bereich der Datenwissenschaft evaluieren können. Unabhängig davon, welche Anbieter Sie sich ansehen, sollten diese in der Lage sein, die nachfolgenden Fragen zu beantworten. Warum diese Fragen so wichtig sind, wird in diesem Leitfaden ebenfalls erklärt.

INHALTSVERZEICHNIS

Die Grundlagen

Bei der Prüfung eines KI-Anbieters ist es hilfreich zu verstehen, welche Arten von KI-Modellen dieser verwendet. Diese Informationen können Ihnen Aufschluss über Kosten, Genauigkeit und andere geschäftlich relevante Faktoren dieser KI geben.

Welche Arten von ML/KI-Modellen setzt Ihre Kerntechnologie ein?

In ihrer Antwort verweisen Anbieter möglicherweise auf Trainingsmodalitäten, Modellfamilien und Modelltypen. Hier geht es nicht darum, ob Anbieter eine bestimmte Art von KI bereitstellen, sondern sie sollten in der Lage sein zu erklären, welche Modelle sie verwenden, wie diese funktionieren und warum dies die am besten geeignete KI für Ihren Anwendungsfall ist.

In vielen Fällen müssen Modelle nur einfache Anforderungen erfüllen können, sodass es schneller und kostengünstiger ist, ein relativ einfaches, hochgradig zielgerichtetes Modell einzusetzen, statt in ein großes, komplexes neuronales Netzwerk zu investieren.

Welche Infrastruktur ist für den Betrieb der Modelle erforderlich? Ist es der Kunde (selbst gehostet) oder der Anbieter (SaaS), der die notwendige Hardware bereitstellt?

KI-Modelle verschlingen mit zunehmender Größe immer mehr Ressourcen. Wie würden Kosten- oder Ressourcenknappheit (z. B. fehlender Zugriff auf GPUs) unsere Fähigkeit beeinträchtigen, die KI für unsere Unternehmensworkloads zu skalieren?

Die Daten

Ohne qualitativ hochwertige, relevante Eingabedaten können KI-Modelle keine genauen oder praxistauglichen Ergebnisse liefern. Wenn Ihr KI-Anbieter Ihnen nicht sagen kann, mit welchen Daten seine Modelle trainiert wurden, können Sie auch nicht herausfinden, wie seine Modelle funktionieren oder welche Faktoren die Ergebnisse beeinflussen.

Beschreiben Sie den Prozess, mit dem Ihr Modell erstellt und trainiert wurde.

In den Antworten potenzieller Anbieter sollten Trainings-, Validierungs- und Testdatensätze erwähnt werden. Diese Daten-Batches werden dazu verwendet, KI- und Machine-Learning-Modelle zu trainieren, ihre Ergebnisse zu bewerten und zu optimieren und schließlich die Endergebnisse zu testen. Indem diese Daten für jede Funktion in verschiedene Datensätze aufgeteilt werden, können Datenwissenschaftler entsprechende Benchmarks festlegen, um die Performance und Verbesserung der Modelle genau zu bewerten.

Wie überwachen Sie die Datenqualität, bevor das Modell entwickelt wird?

Anbieter sollten in der Lage sein, ihren Prozess der Datenerfassung und -validierung zu erklären. Falls relevant sollten sie auch erläutern können, wie die Qualität der gekennzeichneten Daten überprüft wird.

Welche Art, Quelle und Menge von Daten sind erforderlich, um Ihre Modelle zu trainieren?

Diese Frage hilft Ihnen zu verstehen, welche Datenqualität und -quantität das KI-Modell erfordert. Komplexere Ausgaben erfordern mehr Trainingsdaten.

Als Referenz: Ein kleines Klassifizierungsmodell benötigt etwa 20.000 hochwertige Beispiele pro Klasse. Ein Large Language Model (LLM) benötigt 20–30 Token (Wörter, Codesegmente usw.) pro Parameter im Modell. Selbst ein kleines LLM wie Llama-2 hat 7 Milliarden Parameter, was bedeutet, dass es zum Trainieren mindestens 140 Milliarden Token benötigte.

Die Daten (Fortsetzung)

Wie oft werden Daten zum Training und zur Aktualisierung der Modelle erfasst?

In dem Moment, in dem ein KI-Modell verfügbar gemacht wird, ist es bereits nicht mehr aktuell. Daher ist eine sorgfältige Abwägung zwischen den Kosten und der Komplexität der Modellaktualisierung einerseits und dem Grad der Veralterung andererseits unabdingbar. Einige Modelle müssen nur jährlich oder sogar noch seltener aktualisiert werden, um ihren Nutzen beizubehalten, andere viel häufiger.

Wie werden überwachte Modelle mit Labeln gekennzeichnet? Woher stammen diese Labels?

Beim überwachten Lernen wird eine KI anhand von Datensätzen trainiert, die mit entsprechenden Labels gekennzeichnet wurden. Diese Aufgabe kann entweder das interne Team übernehmen, welches das Modell trainiert, oder sie kann ausgelagert oder automatisiert bzw. durch Crowdsourcing erledigt werden. Manche Labels basieren auf synthetischen Daten, die entsprechend der erforderlichen Kriterien künstlich generiert werden. Eine unzulängliche Kennzeichnung führt zu unzulänglichen Ergebnissen, weswegen Labels optimiert und auf Klarheit und Einheitlichkeit überprüft werden sollten.

Wer ist für das Training und die Validierung von Modellen verantwortlich? Haben Sie ein internes Team für maschinelles Lernen, und wie groß ist es?

Wenn der Kunde für die Feinabstimmung der Modelle verantwortlich ist, verfügt er dann über ein Team für maschinelles Lernen mit dem entsprechenden Fachwissen, das diese Aufgabe übernehmen kann?

Die Modelle

Damit KI-Modelle ihre Funktionalität bewahren, muss ihre Genauigkeit kontinuierlich gewährleistet werden. Hat Ihr Unternehmen hingegen keinen Plan für die regelmäßige Aktualisierung und Erneuerung Ihrer Modelle, treffen Sie möglicherweise Entscheidungen auf der Grundlage fehlerhafter Daten.

Wie überwachen Sie die Genauigkeit und Leistung des Modells? Wer ist dafür verantwortlich?

Modelle sollten kontinuierlich auf Faktoren wie Genauigkeit, Ressourcenverbrauch, API-Nutzung und Anforderungsvolumen hin geprüft werden. Dashboards und Alarm-Trigger können diese Prüfprozesse unterstützen, indem sie eine ganze Reihe von Überwachungsprozessen automatisieren. Es ist jedoch wichtig nachzufragen, wer für die Überwachung verantwortlich ist. Wenn etwa ein SRE- oder Infrastrukturteam zuständig ist, beschränken diese sich möglicherweise auf die Überwachung der Betriebszeit und vernachlässigen andere Faktoren wie etwa die Ausgabequalität.

Können Sie testen, wie genau Ihre Modelle sind?

Durch die Überprüfung der Testergebnisse erhalten Sie ein gutes Verständnis dafür, welche Faktoren wie oft überwacht werden und ob dabei regelmäßig Probleme mit der KI aufgedeckt werden.

Kann Ihre KI an die individuellen Bedürfnisse von Kunden angepasst werden?

Je nachdem, aus welchem Grund Sie eine KI-Lösung erwerben möchten, kann eine maßgeschneiderte Lösung zu weitaus besseren Ergebnissen führen als ein generisches Modell.

Wie gehen Sie mit Datenabweichungen um?

„Drift“ bezieht sich auf Veränderungen im Laufe der Zeit, die die Eigenschaften der zugrunde liegenden Trainings- und Eingabedaten beeinflussen. Ein Modell, das dazu entwickelt wurde, die Stimmung von Reden zu bewerten, kann schnell veraltet sein, wenn sich die Bedeutung von Wörtern ändert. Ein Anbieter sollte in der Lage sein zu erklären, wie und wie schnell Datendrift erkannt wird.

Wie erfassen und integrieren Sie Feedback in Ihre Modelle?

Zwar gibt es KI-Modelle, die menschliche Leistungen übertreffen, doch auch diese Modelle machen Fehler. Anbieter müssen daher in der Lage sein zu erklären, wie die vom Kunden festgestellten Fehler in die zukünftige Nachschulung des Modells einbezogen werden können. Oft erfolgt dies durch einen einfachen Feedbackprozess. Komplexere Modelle erfordern jedoch möglicherweise eine direkte Interaktion zwischen Kunden und Anbietern, um Fehler besser zu verstehen und um zu ermitteln, wie die Modelle aktualisiert werden können

Die Modelle (Fortsetzung)

Wie schnell werden neue und aktualisierte Modelle in der Produktion bereitgestellt? Wie werden Änderungen und Fehlerbehebungen vorgenommen?

Informieren Sie sich im Voraus über potenzielle Probleme, die die rechtzeitige Bereitstellung von Updates und Korrekturen beeinträchtigen könnten. Finden Sie heraus, ob der Zugriff der Endbenutzer auf die KI während Updates beeinträchtigt wird.

Welche Vorteile bieten sich uns, wenn wir unsere Daten zum Training der Modelle nutzen?

Der Anbieter sollte erklären können, wie er Ihre Unternehmensdaten auf verantwortungsvolle Weise nutzt, um seine KI-Modelle zu trainieren und die Ergebnisse so zu verfeinern, dass sie mit der Kommunikationsstrategie und den Abläufen in Ihrem Unternehmen übereinstimmen. Mit einem solchen Ansatz kann der Anbieter präzisere, individuellere und effektivere Lösungen bereitstellen, die auf Ihre spezifischen Anforderungen zugeschnitten sind – was zu einem höheren ROI, wertvollen Insights und einem besseren Schutz führt. Während der Anbieter seine KI ständig anpasst und verbessert, stärkt er damit Ihre Sicherheitslage und bietet Ihnen eine Lösung, die sich mit Ihrem Unternehmen weiterentwickelt.

Beschreiben Sie die Datenverarbeitungspipeline, die es Ihnen ermöglicht, neue und aktualisierte Modelle für Kunden bereitzustellen.

KI-Modell-Pipelines sind oft komplex. Ein Anbieter sollte in der Lage sein, den Prozess zu erläutern, durch den diese Pipelines den Kunden schnell, effizient und mit minimalen Ausfallzeiten bereitgestellt werden können.

Die Kosten

KI-Modelle benötigen oft eine enorme Rechenleistung. Um die richtige KI für Ihre Bedürfnisse zu bestimmen, müssen Sie also den Umfang der angebotenen KI kennen, ebenso wie die Kostenfaktoren.

Wie hoch sind die geschätzten Kosten für den Aufbau eines Modells?

Wie hoch sind die Laufzeitkosten des Modells?

Wie groß sind Ihre Modelle typischerweise?

Bedenken Sie, dass viele KI-Modelle im Laufe der Zeit wachsen, während sie anhand neuer Daten verfeinert werden.

Skalierbarkeit

Als Skalierbarkeit bezeichnen wir die Fähigkeit des Modells, ohne Leistungseinbußen mehr Daten, Benutzer und Aufgaben zu verarbeiten. Diese Fähigkeit ist besonders wichtig für Großunternehmen.

Wie schnell ist die Modellinferenz?

Als Modellinferenz bezeichnet man den Prozess, bei dem eine Vorhersage für einen bestimmten Datensatz erstellt wird (z. B. die Wahrscheinlichkeit, dass ein Kunde nach einem Call-Center-Anruf abwandert). In vielen Fällen können Inferenzen während der Dateneinspeisung nahezu in Echtzeit gezogen werden. Je nach Modelltyp, Datenbedarf und Kostenerwägungen kann es jedoch sinnvoller sein, eine Batch-Verarbeitung durchzuführen. Als Kunde müssen Sie die geschäftlichen Auswirkungen der Batch- gegenüber der Echtzeit-Inferenz berücksichtigen.

Wie skalierbar ist die KI in Bezug auf die Dateneinspeisung?

Wie gelangen Ihre Modelloptimierungsdaten in das System, und können Sie Daten ausschließen, die das Modell nicht aufnehmen soll?

Wie integriert sich die KI in Ihre bestehenden Systeme?

Welche Art von IT-Aufwand ist erforderlich, um die KI zu verbinden? Sind mit der Nutzung von APIs Kosten verbunden? Wer schreibt und pflegt diesen Code? Und welche Infrastruktur wird auf Kundenseite benötigt, damit die Integration funktioniert?

Verantwortlichkeit

KI hat in der Vergangenheit aufgrund unethischer Praktiken bei der Datenerfassung und -verwaltung einige negative Schlagzeilen gemacht. Dies könnte dazu führen, dass die Verantwortlichen für Informationssicherheit und Rechtsangelegenheiten zögern, den Einsatz von KI zu genehmigen. Antworten auf diese Fragen können dabei helfen, solche Bedenken zu beseitigen.

Mit welcher Art von Verpflichtung, Versprechen oder Zusage bei der Nutzung von KI, Daten und Insights fördert Ihr Unternehmen eine vertrauensvolle und transparente Zusammenarbeit?

Welche Richtlinien und Selbstverpflichtungen halten Sie hinsichtlich der Umsetzung verantwortungsvoller KI-/ML-Praktiken, bei denen Datenschutz, Fairness, Transparenz und Interpretierbarkeit im Vordergrund stehen, ein? Gewährleisten Sie die Sicherheit und Rechenschaftslegung, indem Sie eine menschliche Aufsicht im Prozess integrieren? Wir erkennen auch die Bedeutung der Nachhaltigkeit an und integrieren umweltbewusste Praktiken in unsere KI-Initiativen.

Wie geht das KI-Modell mit Fairness und Bias um?

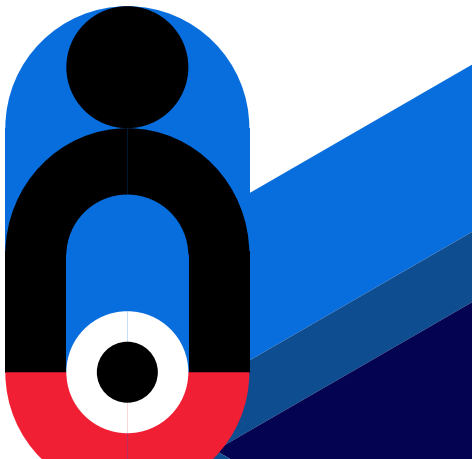
Je nach Anwendungsfall der KI können Bias (Voreingenommenheit) und Fairness sehr unterschiedliche Dinge bedeuten und auch andere Auswirkungen für die Modellausgaben haben. Sie sollten daher die vom Anbieter vorgeschlagenen Maßnahmen zur Vermeidung von Bias im Vergleich zu deren Auswirkungen auf das Endergebnis evaluieren.

Wie werden die Daten gesichert und verwaltet?

Werden die Daten vom Kunden oder vom Dienstleister gehostet? Welche Infrastruktur wird zur Verwaltung der Daten verwendet und wie wird diese geschützt?

Abschließende Gedanken

KI ist ein heiß diskutiertes Thema in der Geschäftswelt – und das zu Recht: Künstliche Intelligenz kann Unternehmen unglaubliche Vorteile erschließen. Bevor Sie jedoch eine KI-Lösung kaufen, müssen Sie sicherstellen, dass diese Technologie auch die gewünschten Ergebnisse erzielen kann. Indem Sie die Antworten auf die obigen Fragen mit Ihren Ansprüchen abgleichen, wird es Ihnen gelingen, bei der Auswahl von KI für Ihr Unternehmen die richtige Entscheidung zu treffen.



mimecast®