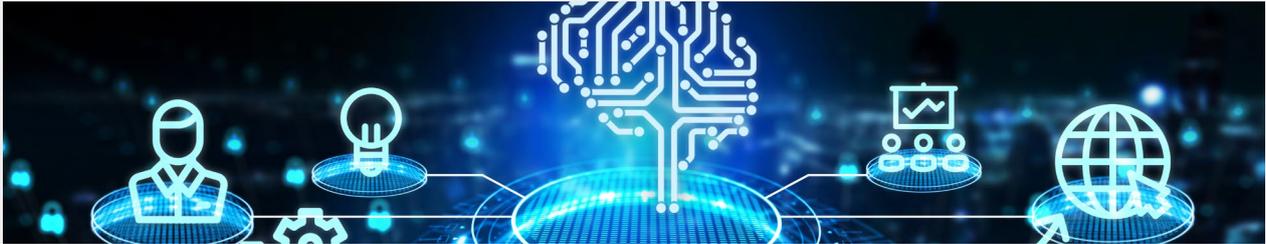


EIN KI-ENGINEERING-WORKFLOW FÜR ERFOLGREICHE KI-PROJEKTE



WAS IST EIN KI-ENGINEERING-WORKFLOW?

Ein KI-Engineering-Workflow beschreibt ein systematisches Vorgehen für die Entwicklung von KI-gestützten Systemen, um die Komplexität der beteiligten Rollen und Disziplinen besser zu beherrschen. Es gibt sieben Disziplinen, die jeweils mehrere Feedbackschleifen enthalten können: Anforderungserhebung, Datenerhebung, Datenaufbereitung, Modelltraining, Modellevaluierung, Modellbereitstellung und Modellüberwachung. Diese werden, je nach Unternehmensgröße, auf mehrere Rollen wie Data Engineers, Data Scientists und MLOps Engineers verteilt. Zwischen den Rollen und Disziplinen ist eine enge Zusammenarbeit zentral.

NUTZEN

- » Transparentes Prozessverständnis mithilfe von klaren Schritten und Schnittstellen
- » Messbarkeit des Erfolgs der entwickelten Modelle durch Evaluierung basierend auf Geschäftsmetriken
- » Systematische Zusammenarbeit und Kommunikation durch integrierte Rollen und Abläufe
- » Gute Anwendbarkeit auf unterschiedliche Datenbestände und Anwendungsfälle
- » Wiederverwendbarkeit von Ergebnissen und Komplexitätsreduktion durch separate, in sich abgeschlossene Teilschritte



SO FUNKTIONIERT'S!

Im ersten Schritt des KI-Engineering-Workflows erfolgt die Anforderungserhebung. Hier werden Geschäftsmetriken festgelegt, durch die der reale Erfolg des Modells gemessen werden kann. Danach identifiziert der:die Data Engineer:in im Schritt der Datenerhebung relevante Datenquellen und extrahiert die benötigten Daten. Im dritten Schritt erfolgt die Datenaufbereitung, um die Qualität der gesammelten Daten zu verbessern. Danach geht es an das Modelltraining, bei dem das Modell mit einem Trainings-Datensatz durch mathematische Verfahren und in einem iterativen Prozess gelernt wird. Im fünften Schritt erfolgt die Modellevaluierung mithilfe eines Test-Datensatzes. Hier wird die gewünschte Leistung des Modells durch Abgleich mit den zu Beginn definierten Leistungs- und Geschäftsmetriken geprüft und sichergestellt.

Die Schritte drei bis fünf werden von dem:der Data Scientist:in ausgeführt. Im sechsten Schritt wird das geprüfte Modell der Rolle MLOps Engineer:in übergeben, die dieses in den definierten Anwendungsfall integriert und im laufenden Betrieb im siebten Schritt der Modellüberwachung sicherstellt, dass die gewünschte Leistung in der produktiven Umgebung erzielt wird. Das lässt sich direkt anhand der Geschäftsmetriken ablesen.

KI-ENGINEERING-WORKFLOW IN DER PRAXIS



^ Damit KI-Projekte auch in der Anwendung erfolgreich sind, müssen Feedbackschleifen zwischen den Rollen und Disziplinen eingehalten werden.



GUT ZU WISSEN

2017 schafften es 54 % der gestarteten Data-Science-Projekte nicht oder nur teilweise in die reale Anwendung in der Produktion (Gartner 2017). Ein Grund dafür ist bis heute, dass der Erfolg von KI-Anwendungen mehr als nur das Erstellen eines Modells mittels sterilen Trainingsdaten bedarf. Vielmehr erfordert es ein systematisches Vorgehen, um die Komplexität der beteiligten Rollen und Disziplinen zu beherrschen. Der größte Unterschied zum klassischen Software-Engineering ist, dass neben Code auch Daten und Modelle organisiert werden müssen.



PRAXISBEISPIEL

Ein produzierendes Unternehmen hat sich als Ziel gesetzt, softwarenahe Geschäftsbereiche in ihren Entscheidungen zu unterstützen, wenn es um die Auswahl von zukunftssträchtigen Softwaretechnologien geht. Dafür wurden in einer Anforderungserhebung die zu analysierenden Technologien und eine Geschäftsmetrik (z. B. Entwicklungsproduktivität) identifiziert. Anschließend hat der:die Data Engineer:in begonnen, Datenquellen der Technologien (GitHub Events, StackOverflow Diskussionen, Stellenanzeigen) zu finden und zu extrahieren. Diese Daten wurden dem:der Data Scientist:in übergeben, um die Qualität der Daten zu verbessern, damit verschiedene Modelle mit verschiedenen Konfigurationen trainiert und evaluiert werden konnten. Das Modell mit der besten Leistung wurde dem:der MLOps Engineer:in übergeben und mithilfe eines Microservices der Außenwelt zur Verfügung gestellt. Damit das Modell auch für neuauftretende Technologien gerüstet ist, wird die produktive Leistung laufend mit der Geschäftsmetrik überwacht. Bei abfallender Leistung wird der KI-Engineering-Workflow bei dem:der Data Engineer:in neu gestartet, um ein Modell mit Daten inklusive neuauftretenden Technologien zu erlernen.

» Besuchen Sie uns auf: www.digitalzentrum-augsburg.de

IMPRESSUM

Verleger: Fraunhofer-Institut für Gießerei-, Composite- und Verarbeitungstechnik IGCV | Am Technologiezentrum 2 | 86159 Augsburg | Als rechtlich nicht selbstständige Einrichtung der Fraunhofer-Gesellschaft zur Förderung der angewandten Forschung e. V. | Hansastraße 27c | 80686 München | Tel.: 0821 90678-0 | E-Mail: info@igcv.fraunhofer.de | Vertretung: Präsident des Vorstandes: Prof. Dr.-Ing. Reimund Neugebauer | **Text/Inhalt:** Alexandros Tsakpinis, fortiss GmbH | **Bildnachweis:** Vorderseite: © photon_photo - stock.adobe.com, Rückseite: © Egor - stock.adobe.com