

Praxisprobleme?

WIE ÜBERFÜHRE ICH MACHINE-LEARNING-MODELLE PROFESSIONELL IN DIE PRODUKTIVE PHASE?

„Ich habe drei Wochen gebraucht, um das Machine-Learning-Modell zu entwickeln. Inzwischen ist ein Jahr vergangen und es ist immer noch nicht in Produktion.“ Dies oder ähnliche Aussagen kommen immer wieder in Projekten vor. Diese Klage eines anonymen KI-Entwicklers beschreibt das Dilemma, in dem viele Unternehmen stecken, die ML-Projekte verfolgen und eigentlich die Vorteile von Künstlicher Intelligenz und Machine Learning (KI / ML) in größerem Umfang nutzen möchten.

Auch die Mitglieder des Cross-Business-Architecture Lab (CBA Lab) sehen bei der Übernahme der als Prototypen entwickelten ML-Lösungen in die Produktion noch einige unbewältigte Hürden. Bei ML ist man noch weit entfernt von einem reifen und etablierten DevOps-Ansatz, der bei Entwicklung und Produktion von Nicht-KI/ML-Projekten inzwischen zum Mainstream geworden ist. DevOps beschreibt die Verzahnung des Entwicklungsprozesses mit dem IT-Betrieb. Deshalb befasste sich der KI/ML-Workstream des CBA Lab genau mit diesem Übergang zwischen Entwicklung und Produktion.

Qualität und Automatisierungsgrad

Ziel war es einen Ansatz zu entwickeln, der den Anforderungen in ML-Projekten an Qualität und Automatisierungsgrad entsprechen kann. Der Workstream hat sich deshalb intensiv mit dem sogenannten MLOps- Ansatz auseinandergesetzt. MLOps steht für eine auf Machine Learning ausgerichtete Vorgehensweise, die die Tugenden des Development-and-Operations-Modells (DevOps) nutzt.

Der Workstream führt einige der Aufgaben aus, die der ML-Einsatz mit sich bringt:

➤ Der Betrieb von ML-Systemen ist aufwändiger als bei klassischer Software, weil Training, Deployment und das Monitoring sowie die regelmäßige Anpassung (Retraining) der ML-Modelle mehr Aufwand bringen. Auch die Versionierung ist anspruchsvoll, weil bei ML die zugehörigen Modelle, Trainings, Validierungs- und Testdaten mit versioniert werden müssen, um die Nachvollziehbarkeit zu gewährleisten.

➤ Datenschutz ist in Bezug auf ML oft nicht klar. Dürfen zum Beispiel Bilder von Personen für das Training von ML-Modellen genutzt werden? Wenn Entscheidungen von einer ML getroffen werden, zum Beispiel bezüglich eines Kredites, ist nicht klar geregelt, wie detailliert die Entscheidung gegenüber den Betroffenen nachvollziehbar gemacht werden muss.

➤ Erfahrene Data Science- und ML-Engineering-Spezialisten sind am Arbeitsmarkt eine rare Ressource. Sie werden aber für kompetente Entwicklungsteams gebraucht – genauso wie Expertise im Bereich Software-Engineering und Betrieb. Den eingesetzten Teams fehlt es außerdem an Kompetenzdiversität. Das kann dazu führen, dass Pilotprojekte in kleinem Rahmen funktionieren, dann aber technisch und organisatorisch nicht skalieren (cross-funktionale Teams können hier eine Lösung darstellen).

➤ Die Kosten werden häufig unterschätzt, weil mehr Aufwand als in einem klassischen Softwareprojekt berücksichtigt werden muss (etwa. höhere Personalkosten oder Kosten für Datenaufbereitung, Spezialhardware, Modelltraining, Wartezeiten der Fachseite, Überführung in die Produktion).





”

DAS MLOPS-PRINZIP FUNKTIONIERT NUR DANN, WENN DIE ORGANISATION AUCH ÜBER DIE NÖTIGEN FÄHIGKEITEN VERFÜGT, DIE DER WORKSTREAM IN EINEM CAPABILITY FRAMEWORK ZUSAMMENFASST.

Dr. Jürgen Klein,
Digital Technology Portfolio Manager,
Carl Zeiss AG, www.zeiss.de

- Fehlende Nachvollziehbarkeit der Entscheidungen.
- Unbekannte Abhängigkeiten der Daten, die für das Trainieren der Modelle verwendet werden.
- Fehlende oder nicht verlässliche Daten.

Von DevOps zu MLOps

Einige dieser Gestaltungsaufgaben können adressiert werden, indem man ML-Anwendungen einem erweiterten DevOps-Ansatz unterwirft, dem sogenannten MLOps. Es erweitert die bekannten DevOps-Prinzipien, um die Entwicklung und den Betrieb von ML-basierten Anteilen der Lösungen spezifisch und optimal zu unterstützen. Das Hinzufügen neuer Datensätze, aber auch die schleichende Degradation der Modellperformanz benötigt ein kontinuierliches Training (CT), um diese stabil zu halten oder gar zu verbessern.

Da ein ML-Modell meistens nur eine kleine, aber sehr kritische Komponente eines Software-Systems darstellt, muss ihre Interaktion mit anderen Komponenten ständig überprüft werden. Das bedeutet die Überprüfung neuer Modelle durch besondere Testverfahren wie Daten- und Modellvalidierung.

Das MLOps-Prinzip funktioniert allerdings nur dann, wenn die Organisation auch über die nötigen Fähigkeiten verfügt, die der Workstream in einem Capability Framework zusammenfasst.

Es besteht aus folgenden sechs Bausteinen:

1. Mensch & Kompetenz: Der Mensch und die benötigten Fertigkeiten sind Grundvoraussetzung für ein erfolgreiches MLOps. Es braucht nicht nur den Data Scientist oder den ML Engineer, sondern eine Vielzahl unterschiedlichster Fähigkeiten. Diese müssen rekrutiert, ausgebildet und an die Organisation gebunden werden.

2. Kultur: Die Organisation muss sich auf die neuen Technologien auch kulturell vorbereiten. Es braucht bei den einzelnen Teilnehmenden einer MLOps-Initiative, aber auch in der gesamten Organisation, eine Bereitschaft, sich auf ML-unterstützte Prozesse einzulassen und diese ständig weiterzuentwickeln. Eine Grundvoraussetzung dafür ist die Unterstützung des Topmanagements. Die Organisation kann sich erst dann zur Einführung von MLOps verpflichten, wenn das Topmanagement klare Support-Signale sendet.

3. Prozesse: Änderungen, die mit der Adaption von ML einhergehen, beeinflussen immer die Prozesse ei-

ner Organisation. Die Prozesse werden aufgrund des systematischen Einbezugs von Datenströmen geändert.

4. Daten: Daten sind der Treibstoff für eine ML-Organisation. Ohne qualitativ hochwertige und korrekte Daten gibt es kein ML. Unternehmen haben häufig Probleme mit der Qualität historischer Daten. Deshalb müssen grundlegende Fähigkeiten wie Datenaufbereitung, Datenverarbeitung und Datenqualitätssicherung verbessert werden, um die Bereitschaft für ML zu erhöhen.

5. Technologie und Infrastruktur: ML basiert auf einem komplexen Technologie-Stack und benötigt eine hoch performante Infrastruktur, die in einem sehr dynamischen Umfeld funktionieren muss. Die stetige technologische Innovation und Pflege sind Grundvoraussetzung für ML. Dafür müssen die notwendigen Ressourcen sowohl finanziell als auch personell bereitgestellt werden.

6. Risiko, Compliance & Ethik: Der Einsatz von Systemen, die potenziell selbstständig Entscheidungen treffen, birgt Gefahren. So können unausgewogene Daten zu tendenziösen Resultaten und unethischen Entscheidungen führen, die im schlimmsten Fall Menschen gefährden und die ganze Organisation bedrohen können.

Für die Beherrschung der Risiken und die Sicherstellung der Compliance ergeben sich damit völlig neue Fragestellungen.

Dr. Jürgen Klein

ÜBER DAS CROSS-BUSINESS-ARCHITECTURE LAB

Das CBA Lab ist ein Anwenderverband von Unternehmen aus allen Wirtschaftszweigen, die gemeinsam neue Best Practices erschließen, erarbeiten und trainieren. Es erarbeitet mit und für seine Mitglieder innovative „Bausteine“ für die Digitale Transformation, die die Architektur prägen und organisieren. Am Cross-Business-Architecture Lab beteiligen sich CIOs, CDOs und Chefarchitekten aus führenden Unternehmen und Organisationen im deutschsprachigen Raum. Die Mitglieder profitieren vom gemeinsamen Netzwerk und dem Vertrauensraum des Verbandes, der sie sehr offen Know-how und Ideen teilen lässt.