Dissertationsvorhaben von Nicolas Weeger

Thema: AI Engineering Blueprints for practicable Machine Learning development

Hochschule für angewandte Wissenschaften Ansbach, Fakultät Technik

Erstbetreuer: Prof. Dr. Christian Uhl

Zweitbetreuer: Prof. Dr. Stefan Geißelsöder

angestrebter Titel: Dr. rer. nat.

1 Forschungsthema

Künstliche Intelligenz (KI) verändert zahlreiche Industrien und Anwendungsbereiche. Für Unternehmen ist es entscheidend KI-Techniken einzusetzen, um geschäftlichen Erfolg zu erzielen [1], [2]. Vor allem KMU können von der Einführung von Einsatz von KI-Techniken profitieren. Sie können ihre Fähigkeiten Fähigkeiten in bestimmten Bereichen, wie z.B. Kundenerfahrung, Produktionsüberwachung und Entscheidungsprozesse [3]. Die firmeneigene Entwicklung von ML-Modellen zur Verwendung als oder innerhalb eines Produkts, genannt KI-Systeme, in einem organisatorischen Kontext kann zu einer Reihe von Herausforderungen führen [4]-[7]. Diese Dazu gehört das Verständnis für die Feinheiten der KI, einschließlich ihre funktionalen Anforderungen und Einsatzszenarien. Die Integration zusätzlicher Prozesse, wie z. B. Datengenerierung Datengenerierung und -vorverarbeitung oder Modelltraining und -einsatz, in den traditionellen Softwareentwicklungsprozess kann potenziell zu organisatorischen Promblemen führen, insbesondere für KMU [6]. Der ML Modellentwicklungszyklus umfasst zusätzliche Praktiken neben DevOps für die Daten und Modelle. Dazu gehören MLOps und DataOps-Techniken, die eine Kultur, Praktiken und Werkzeuge für den Umgang mit Daten und Modellen. Darüber hinaus muss die Systemarchitektur für KI-Systeme auf die Anforderungen des Anforderungen des zugrunde liegenden Modells abgestimmt sein. Training und Inferenz Trainings- und Inferenzumgebungen, sowie Datenspeicherung und Versionierung der der verschiedenen Artefakte müssen integriert werden, damit das System als KI-System zu funktionieren. Folglich erfordert die effektive Implementierung von KI-Systemen eine gut durchdachte Architektur, die auf die speziellen Anforderungen Anforderungen der beabsichtigten KI-Anwendung zugeschnitten ist. Diese Dissertation beschäftigt sich mit der Entwicklung von Blueprints, die auf die Anforderungen der verschiedenen Arten von KI und Entwicklungsstufen abgestimmt sind. Diese Blueprints verbinden die Prinzipien von AI-Engineering, DevOps, MLOps und DataOps um die Herausforderungen, die mit der Entwicklung von KI-Systemen verbunden sind, zu bewältigen. Folglich werden sie die Blueprints anwenden können, indem sie Referenzarchitekturen Referenzarchitekturen und geeignete Automatisierungsansätze für verschiedene Arten von KI implementieren.

2 Stand der Forschung

2.1 AI Engineering

AI-Engineering ist eine Weiterentwicklung des Bereichs Software Software-Engineering, und aufgrund des raschen Wachstums der ML-Entwicklungen ist es ein aufstrebendes Feld. Die KI-Technik befindet sich derzeit auf dem Höhepunkt der Erwartungen“, wie der von Gartner Gartners AI Hype Cycle für 2024 (https://www.gartner.com/en/articles/hype-cycle-for-artifcial-intelligence) Engineering ist die Grundlage für die unternehmensweite Bereitstellung von KI und GenAI in großem Umfang. Den meisten Unternehmen fehlen die Daten-, Analytik und Software-Grundlagen, um einzelne KI-Projekte in großem Produktion zu bringen - geschweige denn ein Portfolio von KI Lösungen in großem Umfang zu betreiben.“ Mehr als ein Dutzend Projekte wurden untersucht in [8] untersucht, bei denen die Herausforderungen der KI-Technik zu Problemen bei der ML-Modelle zu produzieren. Die Studie ergab, dass die Mehrheit Unternehmen, die Modelle für maschinelles Lernen entwickeln, auf Schwierigkeiten haben, wenn sie versuchen, diese in die Produktion zu überführen.
Sie bieten eine Forschungsagenda und einen Überblick über die Fragen, die die in dieser Richtung angegangen werden müssen. [9] weisen darauf hin, dass zwar der Bereich der Software-Engineering-Forschung bereits ausgiebig diskutiert wurde, KI-Engineering jedoch viel weniger behandelt wurde. Nur eine begrenzte Anzahl von Publikationen präsentieren konkrete Erfahrungen Erfahrungen mit der Anwendung von KI-Engineering-Prinzipien. Sie wählten zehn AI-Engineering-Praktiken in mehreren Kategorien aus der aus der Literatur und wendeten sie auf eine Beispielimplementierung an um die Praktiken und ihre Systemarchitektur zu bewerten. Darüber hinaus haben Gespräche und Fragebögen, insbesondere und Fragebögen, insbesondere mit KMUs, haben gezeigt, dass der Wunsch besteht, KI KI in ihre Systeme zu implementieren. Allerdings hängt der Erfolg ihrer Modell Modells hängt jedoch von der Bewältigung der oben genannten oben genannten Herausforderungen ab. So kann die Anwendung von KI-Engineering kann Unternehmen helfen, die Entwicklung, den Einsatz und den Betrieb von Modellen für maschinelles Lernen zu optimieren.

2.2 MLOps

Die Idee von MLOps ist die Bereitstellung von Techniken und Werkzeugen für den Einsatz und Betrieb von KI-Systemen bereitzustellen [10]. Das Ziel ist es, eine Strategie für die Lösung von realen Problemen mit den Einsatz von ML-Modellen. Mehrere Studien untersuchen verschiedene Literatur in diesem Bereich und bieten Pipelines, Taxonomien, Werkzeuge, Methoden und Herausforderungen in diesem Bereich [11]-[13]. [14] führen eine systematische Mapping-Studie für MLOps Architekturen durch und zeigen 35 Architekturkomponenten auf, de- beschreiben verschiedene Architekturvarianten für unterschiedliche Anwendungsfälle und stellen gängige Werkzeuge für diese Architekturkomponenten zur Verfügung.

2.3 Weitere relevante Forschungsarbeiten

In [15] wurde eine Referenzarchitektur für die spezifischen Anwendungsfälle in der der Prozessindustrie vorgestellt, die sich mit Edge Devices befasst. Sie demonstrierten die Architektur durch die Implementierung einer Fallstudie Fallstudie für einen realen Anwendungsfall und bewies die Funktionalität mit dieser Anwendung. [16] entwickelte eine Referenzarchitektur zur Erleichterung der Nutzung von

Big Data in Edge Computing ML-Techniken zu erleichtern. Sie verschiedene Ansichten über die Architektur der Modellentwicklung Entwicklung und Einsatz für diesen speziellen Anwendungsfall. Eine andere Studie [17] stellt eine Vision für „disziplinierte, wiederholbare wiederholbare und transparente modellgetriebene Entwicklung und Machine Learning Operations (MLOps) von intelligenten Unternehmensapplikationen cations“. Sie stellen ein dreistufiges Metamodell für die modell modellbasierte Entwicklung von AI/ML-Blueprints auf Basis intelligenter Anwendungsarchitektur.
Mit Blick auf Software und Architektur werden Entwurfsmuster für KI-basierte Systeme in mehreren Studien diskutiert [18]-[21]. Sie geben einen Überblick über Entwurfsmuster, die für KI-Anwendungsfälle angepasst oder für KI-Anwendungsfälle, und zeigen die Anwendung und die und die daraus resultierenden Vorteile bei der Entwicklung von Modellen für maschinelles Modelle.

2.4 Zusammenfassung des Standes der Forschung

Zusammenfassend lässt sich sagen, dass die Literatur Einblicke in die Wichtigkeit Bedeutung, mögliche Architekturen und Prinzipien für KI-Engineering und MLOps-Praktiken. Allerdings ist die Anwendung dieser Erkenntnisse konzentriert sich derzeit auf einige wenige Referenzarchitekturen in bestimmten Bereichen, wie z. B. Big Data oder Edge Devices. Andere Studien konzentrieren sich auf die Definition von Architekturen und Mustern und beweisen ihre Anwendbarkeit in Fallstudien. Die Prinzipien des AI-Engineering bilden die Grundlage die Grundlage für die Entwicklung der in diesem Papier vorgeschlagenen Blaupausen. MLOps-Pipelines und -Tools sowie bestehende Referenzarchitekturen Referenzarchitekturen und Frameworks werden eingesetzt, um die Entwicklung Entwicklung von KI-Systemen zu unterstützen und so den Prozess zu rationalisieren, zu standardisieren und beschleunigen den Prozess. Software- und Architekturentwurfsmuster werden zur Beschreibung der Entwicklung verwendet, um die um die nicht-funktionalen Anforderungen (NFRs) für die verschiedenen Entwürfe zu erfüllen. Der Einsatz in Feldprojekten ermöglicht ein flexibles, hochautomatisierten Einsatz und ressourcenschonenden Betrieb für unterschiedliche Anforderungen in KMUs.

3 Ziele und wissenschaftlicher Beitrag der Promotion

4 Methodik

Um die Entwürfe zu untersuchen und ihre Brauchbarkeit für Unternehmen zu validieren für Unternehmen zu überprüfen, wird die Methode der Design Science Research (DSR) [22], [23], eingesetzt werden. Mit Hilfe von Interviews und einer einer umfassenden Literaturrecherche werden die Herausforderungen und Anforderungen von KMU identifiziert und Verbesserungspotenziale Verbesserungsmöglichkeiten ermittelt. Basierend auf diesen Erkenntnissen können Geschäftsanforderungen in Zusammenarbeit mit den relevanten Interessengruppen festgelegt werden, die sich an deren spezifischen Bedürfnissen orientieren. Durch die Integration der fachlichen Anforderungen mit den Anforderungen der verschiedenen Arten von der eingesetzten KI, z. B. Algorithmen, Datenspeicherung, Kom und NFRs, kann ein umfassendes Rahmenwerk Rahmenwerk entwickelt werden, das die Verifizierung der entworfenen Artefakte zu erleichtern. Anschließend können diese einer iterativen Tests und Validierung unterzogen werden. Schließlich können die Artefakte in den den Projekten der Beteiligten als Feldtest eingesetzt werden. Dieser Prozess wird dann Dieser Prozess wird solange wiederholt, bis die Anforderungen fnalisiert sind und die Artefakte die Anforderungen nachweislich

erfüllen. Der Prozess wird in mehreren Projekten wiederholt, um die Ergebnisse zu verallgemeinern und sie für KMU so anwendbar wie möglich zu machen.

5 Arbeits- und Zeitplan

6 Zuordnung zum Promotionskolleg REDIG

Literatur