

Zum Klienten:

Weitere Informationen zum Klienten und laufend neue Jobchancen erhalten Sie nach Ihrer Anmeldung und dem Onboarding durch Ihren Fachberater per Mail oder online. Wir freuen uns, Sie persönlich kennenzulernen!

Arbeitsort:

Bern (X Kilometer und X Minuten von Ihrer Wohnadresse entfernt)

Jobbeschreibung:

Lead AI/ML Engineer (w/m/d)

Ihr Aufgabengebiet:

- Enge Zusammenarbeit mit funktionsübergreifenden Projektteammitgliedern zur Sicherstellung eines integrierten Ansatzes bei der Anwendung von KI-/ML-Modellen
- Leitung von Design und Entwicklung von Generative-AI-Anwendungen unter Einsatz von LLMs, Embeddings, Vektordatenbanken, RAG und Multi-Agenten-Frameworks
- Zusammenarbeit mit AI Engineers, KI-Experten und weiterem technischen Personal zur Bereitstellung durchgängiger Lösungen, die das Nutzererlebnis verbessern und überzeugende Ergebnisse liefern
- Mitwirkung bei der Definition von Standards für die Entwicklung und Anwendung von KI-/ML-Modellen
- Einsatz kreativer und wegweisender Denkansätze, um Anwenderanforderungen mit technischer Machbarkeit und wirtschaftlicher Tragfähigkeit in Einklang zu bringen
- Zusammenarbeit mit Modellentwicklern und Empfehlung von Architekturen zur Sicherstellung einer termingerechten und qualitativ hochwertigen Leistungserbringung
- Gezieltes Feedback und Peer-Review für KI-/ML- und verwandte Teammitglieder
- Nutzung aktuellen Wissens aus KI/ML und Softwareentwicklung, um Ideen in konkrete Modell-, Entwicklungs- und Produktempfehlungen zu überführen

Was Sie ausmacht:

- BS/BE-Abschluss mit mindestens 8 Jahren einschlägiger Erfahrung
- Fundierte praktische Erfahrung mit Large Language Models (LLMs) und deren Einsatz in agentenbasierten und autonomen Systemen
- Solides Verständnis von Agentic-AI-Konzepten, einschliesslich Planung, Reasoning und Multi-Agenten-Zusammenarbeit
- Solides Verständnis und praktische Erfahrung mit Agent-Orchestrierungs-Frameworks und KI-Tooling-Ökosystemen (z. B. CrewAI)
- Erfahrung mit Azure AI Services, einschliesslich Azure OpenAI, Azure AI Doc Intelligence, Translator etc.
- Fundiertes Verständnis von Azure OpenAI GPT-Modellen, einschliesslich Fähigkeiten, Limitierungen, Prompt Engineering, Function Calling, Tool Use und strukturierten Outputs
- Starke Programmierkenntnisse in Python, mit Erfahrung in der Integration von KI-Services, APIs und Datenpipelines
- Erfahrung in der Backend-Entwicklung mit Python und/oder Golang, mit Fokus auf API-basierte Architekturen
- Erfahrung mit Datenbanken wie PostgreSQL sowie Kenntnisse in Vektordatenbanken
- Arbeitskenntnisse in mindestens einem Frontend-Framework wie React oder Angular

- Praktische Erfahrung mit Cloud-nativer Entwicklung, Containern und Microservices
- Nachgewiesene Erfahrung im Cluster-Betrieb mit Kubernetes, Docker, Helm und Plattformen wie Rancher
- Vertrautheit mit UI/UX-Tools wie Figma für schnelles Prototyping und Design-Zusammenarbeit
- Cloud-Zertifizierungen (Azure Developer, Azure AI Engineer oder gleichwertig) von Vorteil

Was diese Stelle ausmacht:

- Zahlreiche individuelle Benefits