

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:

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

Jobbeschreibung:

AI Engineer (Hybrid) (w/m/d)

Ihr Aufgabengebiet:

- Design und Entwicklung anspruchsvoller KI-Lösungen, einschliesslich Large Language Models (LLMs), Retrieval-Augmented Generation (RAG) Systemen und weiteren generativen KI-Anwendungen
- Engineering und Optimierung von Prompts für diverse Business Use Cases zur Sicherstellung von Zuverlässigkeit, Genauigkeit und Performance der KI-Systeme
- Aufbau von End-to-End-KI-Pipelines von der Datenaufnahme über das Modell-Deployment bis zum Monitoring in Produktionsumgebungen
- Zusammenarbeit mit Data Scientists, Software Engineers und Business Stakeholdern zur Übersetzung von Geschäftsanforderungen in skalierbare KI-Lösungen
- Umsetzung von MLOps Best Practices, u.a. Modellversionierung, A/B-Testing sowie Continuous Integration/Deployment für KI-Systeme
- Verfolgung neuer Technologien und Beurteilung ihrer Anwendbarkeit auf Fragestellungen im Versicherungs- und Rückversicherungsgeschäft
- Aufbau und Pflege von Beziehungen zu Business Leadern und Kunden, um deren Bedürfnisse zu verstehen und als vertrauenswürdiger Partner für KI-Innovation zu agieren

Was Sie ausmacht:

- Höherer Abschluss in Informatik, künstlicher Intelligenz, Machine Learning, Ingenieurwesen oder einem verwandten quantitativen Fachgebiet
- 3-5 Jahre praktische Erfahrung in der Entwicklung und Bereitstellung von KI/ML-Lösungen in einem Unternehmens- oder Produktionsumfeld
- Fundierte Kenntnisse moderner KI-Techniken: Prompt Engineering und Optimierung für LLMs, RAG-Architekturen, Vektordatenbanken und Embedding-Modelle, konzeptionelles Know-how zu Fine-Tuning und Anpassung von Foundation Models sowie KI-Agenten und Orchestrierungs-Frameworks
- Sicherer Umgang mit Python und guten Software-Engineering-Praktiken (Versionskontrolle, Testing, CI/CD, Code Reviews)
- Verständnis von MLOps-Tools und -Praktiken (z.B. MLflow, Weights & Biases)
- SQL- und Datenmanipulationskenntnisse, PySpark von Vorteil
- Erfahrung mit ML-Frameworks (PyTorch, TensorFlow, Hugging Face Transformers) und Cloud-Plattformen (Azure, AWS oder GCP)
- Erfahrung mit dem Palantir-Ökosystem und spezifische Erfahrung in der Versicherungsbranche von Vorteil

Was diese Stelle ausmacht:

- Attraktive leistungsbezogene variable Vergütung zusätzlich zum Basisgehalt

- Diverse globale und standortspezifische Zusatzleistungen
- Hybrides Arbeitsmodell mit Präsenz von mindestens drei Tagen pro Woche im Büro
- Zusammenarbeit in einem globalen, vielfältigen Team an mehreren internationalen Standorten