

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 (w/m/d)

Ihr Aufgabengebiet:

- Entwicklung anspruchsvoller KI-Lösungen, darunter Large Language Models (LLMs), Retrieval-Augmented Generation (RAG) Systeme und weitere generative KI-Anwendungen
- Engineering und Optimierung von Prompts für verschiedene Business-Anwendungsfälle mit Fokus auf Zuverlässigkeit, Genauigkeit und Performance
- Aufbau end-to-end KI-Pipelines von der Datenaufnahme bis zum Modell-Deployment und 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 wie Modellversionierung, A/B-Testing sowie Continuous Integration/Deployment für KI-Systeme
- Beobachtung neuer Technologien und Beurteilung ihrer Anwendbarkeit auf Herausforderungen im Versicherungs- und Rückversicherungsgeschäft
- Aufbau und Pflege von Beziehungen zu Business Leadern und Kunden, um deren Bedürfnisse zu verstehen und als vertrauenswürdiger Partner in der KI-Innovation aufzutreten

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 dem Deployment von AI/ML-Lösungen in einem Unternehmens- oder Produktionsumfeld
- Fundierte Expertise in modernen KI-Techniken wie Prompt Engineering, RAG-Architekturen, Vektordatenbanken und Embedding-Modellen sowie konzeptionelles Know-how im Fine-Tuning von Foundation-Modellen und KI-Agenten/Orchestrierungs-Frameworks
- Sicherer Umgang mit Python inklusive solider Software-Engineering-Praktiken (Versionskontrolle, Testing, CI/CD, Code Reviews)
- Verständnis von MLOps-Tools und -Praktiken (z. B. MLflow, Weights & Biases)
- SQL- und Datenverarbeitungskenntnisse, PySpark von Vorteil
- Erfahrung mit ML-Frameworks (PyTorch, TensorFlow, Hugging Face Transformers) sowie Cloud-Plattformen (Azure, AWS oder GCP)
- Erfahrung mit dem Palantir-Ökosystem sowie spezifische Erfahrung in der Versicherungsbranche von Vorteil

Was diese Stelle ausmacht:

- Attraktive leistungsorientierte variable Vergütung

- Vielfältige globale und standortspezifische Zusatzleistungen
- Hybrides Arbeitsmodell mit Präsenz an mindestens drei Tagen pro Woche im Büro
- Inklusive und flexible Arbeitsumgebung mit Fokus auf persönliche und berufliche Entwicklung