RAG-Systeme Basis - Vollzeit - Virtuell

Kurs-Nr. DIGIRAG

Beschreibung

Dieser Kurs führt ML Engineers in die Welt der RAG-Systeme ein. Sie verstehen das RAG-Konzept, arbeiten mit Vector Databases wie Pinecone, Weaviate und ChromaDB, wählen passende Embedding-Modelle und entwickeln Chunking-Strategien. Die Evaluation von RAG-Systemen sorgt dafür, dass Ihre Lösung qualitativ hochwertige Ergebnisse liefert.

Zielgruppe

ML Engineer

Inhalt

- Retrieval-Augmented Generation verstehen - RAG-Architektur, Datenfluss und Abgrenzung zu Fine-Tuning und In-Context-Learning
- Vector Databases - Pinecone, Weaviate, ChromaDB und Qdrant: Architektur, Indexierung und Abfrage-Strategien vergleichen
- Embedding-Modelle auswählen - OpenAI, Cohere, Sentence Transformers und lokale Modelle bewerten und benchmarken
- Chunking-Strategien - Fixed-Size, Semantic, Recursive und Document-Level Chunking mit Overlap-Optimierung
- Evaluation von RAG-Systemen - Retrieval-Qualität, Answer-Qualität, Halluzinations-Rate und RAGAS-Framework einsetzen

Abschluss

- IHK-Bescheinigung

Veranstaltung online ansehen

https://www.ihkakademie-sbh.de/weiterbildung/details/rag-systeme-basis-vollzeit-virtuell_119687

Standort und Termin

Termin auf Anfrage
Vollzeit - Virtuell

Termindetails

12 Unterrichtseinheiten

Investition

499,-
inkl. Unterlagen

Ansprechpartnerin

Linda Broghammer

07721 922-161
07721 922-9161
linda.broghammer@vs.ihk.de

Veranstaltungsort

Virtuelles Klassenzimmer

https://lernwelt.ihkakademie-sbh.de

Förderungen

- ESF Förderung

07721 922-200
akademie@vs.ihk.de
IHK Akademie
Albert-Schweitzer-Str. 7
78052 Villingen-Schwenningen

Kofinanziert von der Europäischen Union

Kofinanziert vom Ministerium für Soziales, Gesundheit und Integration Baden-Württemberg

07721 922-200

akademie@vs.ihk.de

IHK Akademie
Albert-Schweitzer-Str. 7
78052 Villingen-Schwenningen