



Data Engineering für Data Science

Spannende Vorträge im Januar!

Wir konnten mit der ORDIX AG (www.ordix.de), einem mittelständischen IT-Dienstleister, eine Firma gewinnen, die im Januar zwei spannende Vorträge mit Einblick in reale Anwendungsfälle halten wird. Im Anschluss an die Vorträge wird es die Möglichkeit geben, die Experten der ORDIX AG mit Fragen zu löchern. Die Veranstaltung findet im Rahmen unserer Lehrveranstaltung „Data Engineering für Data Science“ statt, jedoch sind auch alle anderen Interessierten der FernUniversität herzlich eingeladen!

 26.01. & 02.02.2024

 FernUniversität in Hagen



 **Neuer Termin**
1 **02. Februar 2024** 4  **14:00 Uhr**

Hadoop und Spark

Apache Hadoop ist ein Open-Source-Framework für die zuverlässige, verteilte Verarbeitung großer Datensätze. Es ist so konzipiert, dass es von einzelnen Servern bis hin zu Tausenden von Rechnern skaliert werden kann. Apache Spark wird häufig im Zusammenhang mit Hadoop für die Ausführung von Data Engineering, Data Science und Machine Learning eingesetzt. Die Architektur von Hadoop und das Zusammenspiel mit Spark werden in diesem Vortrag anhand eines realen Anwendungsfalls beleuchtet.

 **26. Januar 2024**  **14:00 Uhr**

Data Augmentation und Datenqualität

Data Scientists stehen in Projekten häufig vor dem „Cold-Start-Problem“. Dabei kann sich nicht direkt mit dem Erstellen und Tunen von Machine-Learning-Modellen befasst werden, weil keine ausreichenden Daten vorliegen oder die vorliegenden Daten keine ausreichende Qualität vorweisen. In diesem Vortrag wird der Umgang mit dem Cold-Start-Problem anhand eines realen Anwendungsfalls aufgezeigt.



Teilnahme über Zoom:
<https://e.feu.de/deds-ordix>