

✦

Samenvatting

Als PySpark Data Engineer bouw ik dataproducten voor robotsensoren. Ik verbind PySpark, Delta Lake, Structured Streaming, Kafka, MLflow aan stabiele operatie en meetbare resultaten.

🏢

Werkervaring

PySpark Data Engineer

Twente Robotics · Enschede

Jan 2022 - Cadeau

- Ontwierp productie-pipelines voor robotsensoren bij Twente Robotics met PySpark en Delta Lake; een team van 15 beheerde 204 bronnen en verhoogde tijdige levering van 91% naar 99,5%.
- Verbeterde in 'Robotfeature-store' het Structured Streaming-partitioneringsmodel met 8 developers; de mediane querytijd daalde van 28 naar 4,2 seconden.
- Automatiseerde schema- en versheidscontroles voor 36 robotsensoren-tabellen met 8 analisten; datamismatches namen 70% af.
- Herbouwde herverwerking, alerts en runbooks voor Kafka in een operationeel model; incidentrespons daalde met 86 uur per kwartaal.

Data Platform Engineer

Twente Robotics Data Studio · Enschede

Apr 2019 - Dec 2021

- Ontwierp migratie-pipelines voor robotsensoren bij Twente Robotics met PySpark en Delta Lake; een team van 15 beheerde 204 bronnen en verhoogde tijdige levering van 91% naar 99,5%. In de migratiefase.
- Verbeterde in 'Robotfeature-store' het Structured Streaming-partitioneringsmodel met 8 developers; de mediane querytijd daalde van 28 naar 4,2 seconden. In de migratiefase.
- Automatiseerde schema- en versheidscontroles voor 36 robotsensoren-tabellen met 8 analisten; datamismatches namen 70% af. In de migratiefase.
- Herbouwde herverwerking, alerts en runbooks voor Kafka in een migratie model; incidentrespons daalde met 86 uur per kwartaal. In de migratiefase.

🌐

Projecten

Robotfeature-store

Dataprodukt voor robotsensoren, inclusief geteste transformaties en operationele procedures.

PySpark

Delta Lake

Structured Streaming

Kafka

☆

Vaardigheden

Technologie

PySpark

Delta Lake

Structured Streaming

Kafka

MLflow

Kwaliteit en operatie

Datamodellering

Datakwaliteit

CI/CD

Observability

Documentatie

Samenwerking

Requirementsworkshops

Incidentrespons

Mentoring

Agile ontwikkeling



Onderwijs

Universiteit Twente

Bachelor Informatica at Data engineering en gedistribueerde systemen

Sep 2015 - Jun 2019



Certificeringen

- Databricks Spark Developer**
Professionele certificeringsinstantie
May 2024
- Data Reliability Workshop**
Dutch Data Guild
Aug 2023



Talen

Nederlands • Moedertaal

Engels • Zakelijk