

Nurul Aminah

PySpark Data Engineer

 nurul.aminah@example.com

 +60 114 24000012

 Seremban

Summary

Sebagai PySpark Data Engineer, saya mengurus produk data untuk pemprosesan sensor robot. Saya menghubungkan PySpark, Delta Lake, Structured Streaming, Kafka, MLflow dengan operasi yang stabil dan hasil yang boleh diukur.

kerja Experience

PySpark Data Engineer

Negeri Robotics · Seremban

Jan 2022 - Sekarang

- Untuk pemprosesan sensor robot di Negeri Robotics, bina saluran produksi menggunakan PySpark dan Delta Lake; pasukan 15 orang mengurus 204 sumber dan menaikkan kadar siap tepat masa daripada 91% kepada 99.5%.
- Dalam projek "Stor ciri robot", baik pulih partition Structured Streaming bersama 8 pembangun dan melancarkan binaan baharu, lalu kurangkan masa pertanyaan median daripada 28 saat kepada 4.2 saat.
- Automatikkan semakan skema dan kesegaran untuk 36 jadual pemprosesan sensor robot, bekerjasama dengan 8 penganalisis dan peraturan operasi, lalu mengurangkan ketidakpadanan data sebanyak 70%.
- Susun semula pemprosesan semula, amaran dan runbook Kafka dalam model operasi untuk 6 unit perniagaan; masa tindak balas insiden suku tahunan turun 86 jam.

Jurutera Platform Data

Negeri Robotics Data Studio · Seremban

Apr 2019 - Dec 2021

- Untuk pemprosesan sensor robot di Negeri Robotics, bina saluran migrasi menggunakan PySpark dan Delta Lake; pasukan 15 orang mengurus 204 sumber dan menaikkan kadar siap tepat masa daripada 91% kepada 99.5%.
- Dalam projek "Stor ciri robot", baik pulih partition Structured Streaming bersama 8 pembangun dan melancarkan binaan pemindahan, lalu kurangkan masa pertanyaan median daripada 28 saat kepada 4.2 saat.
- Automatikkan semakan skema dan kesegaran untuk 36 jadual pemprosesan sensor robot, bekerjasama dengan 8 penganalisis dan peraturan migrasi, lalu mengurangkan ketidakpadanan data sebanyak 70%.
- Susun semula pemprosesan semula, amaran dan runbook Kafka dalam model peralihan untuk 6 unit perniagaan; masa tindak balas insiden suku tahunan turun 86 jam.

Projects

 **Stor ciri robot**

Mereka bentuk produk data untuk pemprosesan sensor robot, lengkap dengan transformasi diuji dan prosedur operasi.

PySpark

Delta Lake

Structured Streaming

Kafka

Skills

Teknologi

PySpark

Delta Lake

Structured Streaming

Kafka

MLflow

Kualiti dan operasi

Pemodelan data

Kualiti data

CI/CD

Pemerhatian

Dokumentasi

Kerjasama

Bengkel keperluan

Tindak balas insiden

Mentoran

Pembangunan tangkas



Education

Universiti Sains Islam Malaysia

Ijazah Sarjana Muda Sains Komputer di Kejuruteraan data dan sistem teragih

Mar 2015 - Feb 2019



Certifications

- Databricks Associate Developer for Spark**
Badan pensijilan profesional
 - Data Reliability Workshop**
Malaysia Data Guild

May 2024

Aug 2023



Languages

- Bahasa Melayu

• Asli

Inggeris

• Perniagaan