

Yhteenveto

Databricks Data Engineer, joka keskittyy teollista lakehousea ja ML-ominaisuuksia. Rakennan luotettavia datatuotteita työkaluilla Databricks, Delta Lake, DLT, PySpark, Unity Catalog ja yhdistän tekniset valinnat mitattaviin tuloksiin.

Työkokemus

Databricks Data Engineer

TehdasFokus

1/2022 - Nykyinen

- Rakensin Databricks- ja Delta Lake-putket teollista lakehousea ja ML-ominaisuuksia-dataa varten yrityksessä TehdasFokus nykyisessä roolissa; työskentelin tuote- ja alustatiimien kanssa 5,2 TB sensoridataa päivässä mittakaavassa ja nostin ajallaan valmistuvien ajojen osuuden 92 prosentista 99,6 prosenttiin.
- Uudistin DLT-mallit ja osioiden projektissa Huollon lakehouse tuotantovastuulla; koordinoituani 9-henkistä insinööriä lyhensin mediaanikyselyn ajan 24 sekunnista 4,1 sekuntiin.
- Mittaroin skeema-, tuoreus- ja lineage-tarkistukset 38 tuotantotaululle teollista lakehousea ja ML-ominaisuuksia-ympäristössä julkaisuprosessin yhteydessä; tein yhteistyötä 14 analyytikon kanssa ja vähensin datapoikkeamia 63 prosenttia.
- Automatisoin backfillit, hälytykset ja runbookit PySpark-ajolle Lahtissa päivittäistä operointia varten; tuimme 12 markkinaa ja säästimme 86 päivystystuntia neljänneksessä.

Analytics Engineer

TehdasFokus Data Studio

4/2019 - 12/2021

- Rakensin Databricks- ja Delta Lake-putket teollista lakehousea ja ML-ominaisuuksia-dataa varten yrityksessä TehdasFokus aiemmassa roolissa; työskentelin tuote- ja alustatiimien kanssa 5,2 TB sensoridataa päivässä mittakaavassa ja nostin ajallaan valmistuvien ajojen osuuden 92 prosentista 99,6 prosenttiin.
- Uudistin DLT-mallit ja osioiden projektissa Huollon lakehouse siirtymävaiheen vastuulla; koordinoituani 6-henkistä insinööriä lyhensin mediaanikyselyn ajan 24 sekunnista 4,1 sekuntiin.
- Mittaroin skeema-, tuoreus- ja lineage-tarkistukset 38 tuotantotaululle teollista lakehousea ja ML-ominaisuuksia-ympäristössä ylläpitovaiheen yhteydessä; tein yhteistyötä 5 sovellustiimin kanssa ja vähensin datapoikkeamia 63 prosenttia.
- Automatisoin backfillit, hälytykset ja runbookit PySpark-ajolle Lahtissa siirrettävää ylläpitoa varten; tuimme 4 liiketoimintayksikköä ja säästimme 86 päivystystuntia neljänneksessä.

Koulutus

Diplomi-insinööri

Lahden ammattikorkeakoulu

9/2015 - 6/2019

Taidot

Teknologia

Databricks, Delta Lake, DLT, PySpark, Unity Catalog

Toimitus

Tietomallinnus, Datan laatu, CI/CD, Observability, Dokumentointi

Yhteistyö

Vaatimusworkshopit, Incident response, Mentorointi, Ketterä toimitus

 Suomi
 Englanti

Databricks Certified Data Engineer Professional

Ammatillinen sertifiointi

2024

Sertifikaatit

Data Reliability Workshop

Data Engineering Guild

2023

<>Hankkeet

Huollon
lakehouse

Rakensin dokumentoidun datatuotteen teollista lakehousea ja ML-ominaisuuksia-tarpeeseen, testatut muunnokset ja analytiikkaa tukevan toimitusmallin.