

예린 조

Databricks 데이터 엔지니어



예린.조@example.com



+82 10 2407 8107



성남

요약

Databricks 데이터 엔지니어로서 제조 레이크하우스와 ML 피쳐 데이터 제품을 운영해 왔습니다. Databricks, Delta Lake, DLT, PySpark, Unity Catalog을 활용해 안정적인 파이프라인과 측정 가능한 개선을 연결합니다.

업무 경험

- Databricks 데이터 엔지니어**
판교팩토리웍스
1/2022 - 현재의
 - 판교팩토리웍스의 제조 레이크하우스와 ML 피쳐를 위해 Databricks-Delta Lake 기반 운영 파이프라인을 설계하고 10명 팀과 169개 소스를 관리해 정시 완료율을 91%에서 99.5%로 높였습니다.
 - '설비 예지보전 레이크하우스'에서 DLT 파티션과 증분 모델을 개선하고 개발자 8명과 신규 릴리스를 진행해 중앙값 쿼리 시간을 28초에서 5.2초로 줄였습니다.
 - 제조 레이크하우스와 ML 피쳐 테이블 31개의 스키마-신선도 검사를 자동화하고 분석가 9명과 운영 규칙을 합의해 데이터 불일치를 65% 낮췄습니다.
 - PySpark 작업의 재처리-알림-런북을 운영 체계로 정비해 6개 사업부와 운영하고 분기 장애 대응 시간을 76시간 절감했습니다.
- 분석 엔지니어**
판교팩토리웍스 데이터스튜디오
4/2019 - 12/2021
 - 판교팩토리웍스의 제조 레이크하우스와 ML 피쳐를 위해 Databricks-Delta Lake 기반 이관 파이프라인을 설계하고 10명 팀과 169개 소스를 관리해 정시 완료율을 91%에서 99.5%로 높였습니다.
 - '설비 예지보전 레이크하우스'에서 DLT 파티션과 증분 모델을 개선하고 개발자 8명과 이관 릴리스를 진행해 중앙값 쿼리 시간을 28초에서 5.2초로 줄였습니다.
 - 제조 레이크하우스와 ML 피쳐 테이블 31개의 스키마-신선도 검사를 자동화하고 분석가 9명과 마이그레이션 규칙을 합의해 데이터 불일치를 65% 낮췄습니다.
 - PySpark 작업의 재처리-알림-런북을 전환 체계로 정비해 6개 사업부와 운영하고 분기 장애 대응 시간을 76시간 절감했습니다.

<>프로젝트

설비 예지보전 레이크하우스
제조 레이크하우스와 ML 피쳐 문제를 해결하는 데이터 제품을 설계하고 테스트된 변환과 운영 절차를 구축했습니다.

Databricks
Delta Lake
DLT
PySpark

교육

데이터사이언스 석사

한양대학교

3/2015 - 2/2019

기술

기술

Databricks, Delta Lake, DLT, PySpark, Unity Catalog

품질과 운영

데이터 모델링, 데이터 품질, CI/CD, 관측성, 문서화

협업

요구사항 워크숍, 장애 대응, 멘토링, 애자일 개발

한국어 원어민
영어 비즈니스

Databricks Certified Data Engineer Professional

전문 인증기관

2024

Data Reliability Workshop

Korea Data Guild

2023