신도 현

빅데이터 분석가

big-data-analyst@example.com 010-0000-0000 대한민국 서울

요약

대용량 로그와 거래 데이터 분석을 6년간 수행했습니다. 분산 처리 환경에서 분석 성능과 데이터 품질을 함께 관리하고, 수십억 건의 이벤트에서 사업에 의미 있는 패턴을 찾아냅니다.

업무 경험

빅데이터 분석가 01/2020 - 현재의

스케일데이터 서울

- 수십억 건의 이벤트를 Spark 세션 모델로 집계해 고객 행동 분석 배치 시간을 9시간에서 2시간으로 단축했습니다.
- Kafka·Delta Lake 데이터 품질 체크를 구축해 지연·중복 이벤트를 60% 줄였습니다.
- 고객 행동과 상품 노출 데이터를 결합해 추천 후보군의 클릭률을 14% 높였습니다.
- 분산 쿼리와 데이터 파티션 사용 가이드를 만들어 분석팀의 공용 데이터셋 활용을 확산했습니다.

데이터 분석 엔지니어 04/2017 - 12/2019

플랫폼인사이트 서울

- 로그·거래·고객지원 데이터를 Hive 테이블로 정리해 운영 분석을 표준화했습니다.
- Spark SQL 튜닝으로 일일 리포트 배치 시간을 50% 단축했습니다.
- 데이터 샘플링과 집계 검증으로 대용량 처리 결과의 정확성을 확인했습니다.
- 분석가와 엔지니어 간 스키마 변경 프로세스를 문서화해 장애 재발을 줄였습니다.

프로젝트

대규모 고객 행동 분석

수십억 건의 이벤트와 거래를 세션·고객 단위로 집계해 이탈과 추천 기회를 분석했습니다.

기술: Spark, Delta Lake, Python

기술

핵심 역량: SQL, Spark, Hadoop, Python, Kafka, 대용량 데이터 처리

교육

컴퓨터과학 학사 03/2013 - 02/2017

연세대학교 대한민국

인증

Databricks Certified Data Analyst Associate 01/2024

Professional Certification

언어

한국어: 모국어

영어: 업무 가능