

NADIA PRATIWI

Expert Data Engineer

Jakarta, Indonesia | Phone: 08726156210 | Email: nadiapradiwi@email.com

SUMMARY

Highly skilled data engineer with over 12 years of experience in designing, building, and maintaining scalable data pipelines and analytics platforms. Specializes in big data technologies, including Hadoop, Spark, and Kafka. Proficient in data modeling, ETL processes, and real-time data processing. Demonstrates expertise in architecting data solutions to support business objectives and drive data-driven decision-making. Adept at leading data engineering projects and collaborating with cross-functional teams to deliver actionable insights.

EXPERIENCE

PT DATA INOVASI TERDEPAN - Jakarta, Indonesia

Senior Data Engineer (January 2015 – Present)

- Designed and implemented distributed data processing systems using Hadoop and Spark to ingest, process, and analyze large volumes of data.
- Developed and maintained data pipelines to extract, transform, and load (ETL) data from various sources into data lakes and data warehouses.
- Built real-time streaming pipelines using Kafka and Spark Streaming for processing and analyzing streaming data.
- Collaborated with data scientists and analysts to deploy machine learning models into production environments for real-time inference.

EDUCATION

UNIVERSITAS INDONESIA (2008 - 2012)

Bachelor of Computer Science

SKILLS

- Big data technologies (Hadoop, Spark, Kafka)
- Data modeling
- ETL processes
- Real-time data processing
- Distributed systems
- Project management
- Team leadership

PROJECT

Real-Time Fraud Detection Platform (2018-2019)

- Led the development of a real-time fraud detection platform using Kafka and Spark Streaming to monitor and analyze transaction data in real-time.
- Implemented machine learning algorithms for detecting fraudulent activities and generating alerts for immediate response.
- Reduced fraud losses and improved fraud detection accuracy, saving millions of dollars for the organization.

Data Lake Architecture and Implementation Project (2016-2017)

- Designed and implemented a scalable data lake architecture using Hadoop ecosystem technologies, including HDFS, Hive, and Impala.
- Ingested and processed structured and unstructured data from multiple sources, providing a centralized repository for analytics and reporting.
- Enabled data scientists and analysts to perform ad-hoc queries and analysis on large datasets with improved efficiency.