Yusril Albi, Bachelor of Computer Science

Data Engineer | Software Engineer | Machine Learning Engineer
yusrilalbi.pro@gmail.com | https://yusrilalbi.my.id | https://www.linkedin.com/in/yusrilalbi
Indonesia, Jakarta | +6281273032294 | https://wa.me/6281273032294

SUMMARY

Software Engineer specializing in **large-scale data platforms and data engineering**, with strong experience designing, modernizing, and operating **high-throughput batch and streaming pipelines** in cloud environments. Currently working at a multinational R&D organization, collaborating with global stakeholders to deliver **reliable, cost-efficient, and scalable data solutions**. Experienced in **data normalization, enrichment, pipeline refactoring, system modernization (EOL replacement)**, and **end-to-end ownership** from design to operations. Strong background in backend engineering, distributed systems, and big data processing, with proven ability to improve **data quality, performance, and operational stability** at scale. **[Software Engineer - Data]**

EXPERIENCE

Samsung R&D | Data Engineer

Apr 2024 - Present

- Took **end-to-end ownership** of multiple **mission-critical data pipelines**, covering system design, development, testing, deployment, and production operations.
- Modernized and refactored **legacy batch and streaming pipelines**, reducing infrastructure cost from **~$1,000/day to ~$50/day (95% reduction)** through architectural redesign, compute optimization, and elimination of inefficient processing paths.
- Designed **scalable, modular pipeline architectures** with zero-downtime deployment, significantly improving **operational stability and maintainability**.
- Processed and normalized **200B+ records per scheduled run**, using **Spark, Hadoop, Airflow**, and **AWS (EKS, EMR, EC2, S3)** to support high-volume ingestion and transformation workloads.
- Improved **data quality and consistency** by enforcing schema validation, structured transformations, and reliable ingestion contracts across multiple data sources.
- Led large-scale **system refactoring and migration projects**, achieving **96% unit test coverage**, improved code readability, and reduced cyclomatic complexity.
- Collaborated closely with global platform teams and stakeholders on **architecture decisions, design reviews, and operational runbooks**.

Tech: Java (data processing concepts), Spark, Hadoop, Airflow, AWS, Kubernetes, SQL, Python, Linux

----

Pertamina & Petronas (University) | Machine Learning Engineer

Feb 2023 - Jan 2024

- Worked as ML Engineer on an **industry-funded data platform project** focused on **large-scale time-series data processing** for oil & gas asset integrity.
- Designed and implemented **end-to-end data preprocessing and transformation pipelines**, handling noisy, high-volume sensor data prior to model inference.
- Developed a **hybrid prediction architecture** using **LSTM-RNN**, resulting in a **patented model** and academic thesis.
- Ensured reproducibility, pipeline reliability, and clear documentation for model training and evaluation workflows.

Tech: Python, SQL, Time-Series Pipelines, Data Preprocessing, ML Pipelines

----

Shopee | Backend Engineer

Aug 2023 - Nov 2023

- Developed and maintained **backend services and RESTful APIs** for transaction and data ingestion systems.
- Built **batch-oriented crawling pipelines**, collecting and normalizing **1M+ product records** including metadata and images.
- Led a **6-member engineering team**, handling sprint planning, task breakdown, code reviews, and technical discussions.

- Produced clear **technical documentation and API specifications**, improving cross-team integration and development velocity.

Tech: Backend APIs, Batch Processing, SQL, System Design, Documentation

----

Banking Industry | Data Engineer

Sep 2022 - Mar 2023

- Designed and implemented **data infrastructure and ETL pipelines** for financial and regulatory reporting systems.
- Built centralized **monitoring and observability platforms** using **ELK Stack**, integrated with **API Gateway (KONG)**.
- Processed structured financial datasets with strict requirements on **data accuracy, consistency, and auditability**.
- Improved pipeline reliability and operational visibility through logging, alerting, and runbooks.

Tech: SQL, ETL, Elasticsearch, Linux, Data Quality Controls

Service Education | Informatics Mentor – Part Time

Jan 2020 - Present

- Mentored 50+ senior high school students in algorithms, C++, problem-solving, and competitive programming.
- Helped students build strong computational thinking foundations, several of whom advanced to national-level competitions.
- Developed structured teaching materials and hands-on exercises to bridge theory and real-world programming skills.

EDUCATIONS

Bangkit Academy - Cert by Google |2022
Machine Learning - TensorFlow Certified 📇

Universitas Pertamina (Pertamina University) |2019
Bachelor of Computer Science (Cum Laude), Full Scholarship
Informatics Competition & Competitive Programming – National Level (Link: 📇)

- **1st Place** in Intercollegiate Olympiad of Informatics: OSM, FUSC, OSSN, FOHKN, BIMAGE, etc.
- **Honorable Mention** in Hology 4.0

SKILLS

BE Engineer: Go (Golang), Java, C++, BaaS, Redis, Kafka, PostgreSQL, MongoDB, API, node.js, NoSQL, etc.
Data Engineer: SQL, AWS, Airflow, ETL, Python, Big-Query, EKS, EC2, S3, Web Scrapping, Spark, Hadoop, etc.
ML Engineer: Python, NumPy, Pandas, TensorFlow, PyTorch, NLP, LSTM, RNN, DNN, Time Series, etc.
Others: Teamwork, Time Management, Teaching, Proactive Learning, Growth Mindsset, etc.

CERTIFICATIONS

Google: IT Automation, Cloud Management, Crash Course on Python, etc. **Certificates** (02/2022) 📇
Imperial London: Mathematics for Machine Learning Certificate, etc. **Certificates** (03/2022) 📇
DeepLearningAI: TensorFlow Developer, TensorFlow Deployment, etc. **Certificates** (04/2022) 📇

LANGUAGE

Indonesia, Melayu **Native**
English **TOEFL 647** (Listening 48/50, Structure&Written 38/40, Vocabulary 47/50) 📇