ZAENAL SYAMSYUL ARIEF

Data Scientist | Data Engineer | AI Engineer

Garut, West Java | zaenalsyamsyularief@gmail.com | +6281213861662

WhatsApp | LinkedIn | GitHub | Portfolio

PROFESSIONAL SUMMARY

Informatics Engineering graduate with hands-on experience in machine learning, Natural Language Processing (NLP), and Generative AI. Experienced in developing data-processing and AI-assisted workflows using Python, Elasticsearch, TensorFlow, and Large Language Model (LLM) technologies. Interested in building practical AI systems that connect models with real-world data, backend services, and user-facing applications.

WORK EXPERIENCE

Freelance Data & AI Engineer - PURU Software Agency Present

- Developed web-scraping workflows to collect historical Facebook data from 2018 to 2026 and extract articles from websites for client project.
- Applied ETL (Extract, Transform, Load) workflows to clean, normalize, transform, and organize scraped records into structured datasets for further analysis and client use.
- Researched client-related business partners and compiled company logos, addresses, business sectors, and descriptions into structured partner profile

Data Science Intern - NoLimit Indonesia Jul 2025 - Nov 2025

- Developed a Streamlit Validator Panel supporting uploads up to 600 MB and large-scale record processing, with direct data retrieval from APIs and Elasticsearch.
- Enabled date filters spanning multiple years, manual sentiment validation, semantic clustering, LLM-assisted labeling, audit trails, and three user-access roles.
- Processed, cleaned, and prepared social-media datasets ranging from hundreds of thousands to millions of records for sentiment-model development.
- Prototyped Twitter bot detection using eight behavioral indicators and profile-image analysis; supported sentiment routing across four language groups.

Artificial Intelligence Teaching Assistant - Institut Teknologi Garut Feb 2026 - Jul 2026

- Assisted students during Artificial Intelligence practicum sessions covering Python, data preprocessing, machine learning implementation, and model evaluation.
- Helped students troubleshoot code and identify issues in preprocessing and model pipelines, with an emphasis on understanding the cause of errors rather than only providing finished solutions.

PROJECTS

Modnance – AI-Powered Financial Automation System Oct 2026 - Present

- Developing an AI-powered financial assistant using Google Gemini for natural language intent recognition, transaction extraction, and automated financial insights.
- Integrating Discord and WhatsApp through Python, REST APIs, and Evolution API to enable conversational financial tracking.
- Building n8n automation workflows connected to PostgreSQL for transaction recording, balance inquiries, transaction history, and AI-generated financial reports.
- Containerizing application services using Docker Compose to support modular deployment and system integration.

LLM Fine-Tuning, RAG & GRPO Jul 2026

- Built an experimental Large Language Model (LLM) workflow covering dataset preparation, PEFT-based fine-tuning, Retrieval-Augmented Generation (RAG), and GRPO-based optimization using Hugging Face Transformers and TRL.

Implemented model adaptation, retrieval, evaluation, and inference workflows using Python, PyTorch, Hugging Face Transformers, PEFT, and TRL

MyCare AI - Emotion-Aware AI Assistant Jun 2025

- Built a text-based emotion classifier integrated with a Generative AI chatbot, achieving 92% accuracy across four emotion categories.
- Extended the system with CNN/VGG16-based facial emotion recognition using FER-2013 and RAF-DB, then connected the emotion predictions to adaptive chatbot responses.

Tech Stack: Python, TensorFlow, Keras, VGG16, Natural Language Processing (NLP), Computer Vision, Generative AI.

StuntLytics - AI-Based Stunting Risk Prediction System Oct 2025

CiptaMuda AI (APHacton 2025) | *Role: Data Preprocessing & System Integration*

- Built a synthetic dataset containing 100,000 family records using variables derived from BPS and SSGI references.
- Developed a Random Forest workflow with Scikit-learn preprocessing and integrated the results into a Streamlit and Elasticsearch-based decision-support application.

EDUCATION

Bachelor of Informatics Engineering - Institut Teknologi Garut Sep 2022 - Sep 2026

GPA: 3.92/4.00

PROFESSIONAL TRAINING

Machine Learning Engineer Cohort - Coding Camp powered by DBS Foundation Feb 2025 - Jul 2025

Completed project-based training in machine learning, deep learning, Natural Language Processing (NLP), computer vision, ETL, recommendation systems, and model evaluation using Python, TensorFlow, and Scikit-learn.

ORGANIZATIONAL EXPERIENCE

Head of Data - Dicoding Community, Institut Teknologi Garut Nov 2025 – June 2026

- Helped connect community members with Dicoding Learning Paths, coding challenges, and peer-learning opportunities.

Head of Mentorship - Sakode Mei 2022 - Present

- Helped plan a children's coding mentorship program focused on computational thinking, Hour of Code, and Scratch.

TECHNICAL SKILLS

Programming & Data: Python, SQL, Pandas, NumPy

Machine Learning: Scikit-learn, TensorFlow, Keras, Natural Language Processing (NLP), Computer Vision, Model Evaluation

Generative AI: Large Language Models (LLMs), Retrieval-Augmented Generation (RAG), Fine-Tuning, Hugging Face Transformers, PEFT

Backend & Database: FastAPI, PostgreSQL, Elasticsearch, REST APIs

Tools: Git, GitHub, Docker, Jupyter Notebook, Google Colab

CERTIFICATIONS & ACHIEVEMENTS

Pengembangan Generative AI berbasis LLM - Dicoding Indonesia | Jul 2026

Topics: RAG, LLM fine-tuning, PEFT/LoRA, quantization, and model optimization.

Membangun Aplikasi Gen AI dengan Microsoft Azure - Dicoding Indonesia | Apr 2026

Topics: Azure AI Foundry, model deployment, and Responsible AI.

APTIKOM-APHacton 2025 - National Competition Participant | Sep 2025

Machine Learning Engineer - DBS Foundation & Dicoding Academy | Jul 2025

Selected as one of 2,500 participants from 63,000 applicants across 374 universities.

Certificates: View all certificates